

Lineare Algebra

Jan Hendrik Bruinier¹

angepasst von Stephan Ehlen zur Vorlesung Lineare Algebra II
im Sommersemester 2018 an der Universität zu Köln

¹Technische Universität Darmstadt, Fachbereich Mathematik, Schlossgartenstraße 7, 64289 Darmstadt, bruinier@mathematik.tu-darmstadt.de

Inhaltsverzeichnis

1	Lineare Gleichungen und Matrizen	1
1.1	Eine Gleichung und eine Variable	1
1.2	m Gleichungen und n Variablen	2
1.2.1	Das Gauß-Verfahren	7
1.2.2	Übungen	9
1.3	Die Struktur der Lösungsmenge	9
1.3.1	Übungen	12
1.4	Lineare Abhängigkeit und Unabhängigkeit	12
1.4.1	Übungen	16
1.5	Der Rang einer Matrix	16
1.6	Matrixmultiplikation	19
1.6.1	Übungen	20
1.7	Etwas analytische Geometrie	21
1.7.1	Längen und Winkel im $\mathbb{R}^n$	21
1.7.2	Das Vektorprodukt	26
1.8	Anhang: Grundbegriffe	28
1.8.1	Mengen	28
1.8.2	Abbildungen	29
1.8.3	Relationen	31
2	Vektorräume	34
2.1	Gruppen	34
2.1.1	Übungen	36
2.2	Permutationen	37
2.3	Gruppenhomomorphismen	40
2.3.1	Übungen	42
2.4	Körper	43
2.5	Vektorräume	45
2.5.1	Übungen	49
3	Lineare Abbildungen	51
3.1	Der Begriff der linearen Abbildung	51
3.2	Kern und Bild	54
3.2.1	Übungen	56
3.3	Isomorphismen	57
3.4	Lineare Abbildungen und Matrizen	58
3.4.1	Übungen	62
3.5	Isomorphismen von Vektorräumen und Basiswechsel	63
3.6	Die allgemeine lineare Gruppe	67
3.6.1	Übungen	70

4	Die Determinante	71
4.1	Die Leibniz-Formel	71
4.2	Weitere Eigenschaften der Determinante	76
4.2.1	Übungen	81
4.3	Die Determinante eines Endomorphismus	82
5	Eigenwerte und Eigenvektoren	85
5.1	Diagonalisierbarkeit	85
5.2	Ringe und Polynome	89
5.2.1	Polynomfunktionen	91
5.2.2	Division mit Rest	92
5.3	Das charakteristische Polynom	94
5.4	Kriterien für Diagonalisierbarkeit	96
5.5	Trigonalisierung	101
6	Die Jordansche Normalform	102
6.1	Der Satz von Cayley-Hamilton	102
6.2	Zerlegung in Haupträume	104
6.3	Die Jordan-Zerlegung	109
6.4	Zyklische Vektoren	111
6.5	Die Jordansche Normalform	116
7	Linear- und Bilinearformen	122
7.1	Linearformen und der Dualraum	122
7.2	Bilinearformen	126
8	Euklidische und unitäre Vektorräume	134
8.1	Euklidische Vektorräume	134
8.2	Unitäre Vektorräume	140
8.3	Normale Endomorphismen	141
8.4	Anwendung: Selbstadjungierte Endomorphismen euklidischer Vektorräume	146
9	Quadratische Formen	147
9.1	Quadratische Formen	147
9.2	Orthogonalbasen und der Kürzungssatz von Witt	149
9.3	Die orthogonale Gruppe	154
9.4	Quadratische Formen über $\mathbb{R}$	159
9.4.1	Matrix-Zerlegungen	166
	Literatur	170

1 Lineare Gleichungen und Matrizen

Eine wichtige Motivation für die Lineare Algebra besteht darin, Lösungen von linearen Gleichungssystemen zu bestimmen, und die Struktur der Gesamtheit all ihrer Lösungen zu verstehen. Dies führt zu einem allgemeineren Studium von „linearen Strukturen“, insbesondere zu den Begriffen des Vektorraums und der linearen Abbildung. Lineare Abbildungen lassen sich mit Hilfe von Matrizen beschreiben.

Die Symmetrien eines Vektorraums sind durch spezielle lineare Abbildungen gegeben. Deren Studium führt zum Begriff der Determinante, sowie zu Eigenwerten und Eigenvektoren. Zur mathematischen Untersuchung von Längen und Winkeln betrachtet man Skalarprodukte auf Vektorräumen. Dies sind die zentralen Begriffe der Vorlesung *Lineare Algebra 1*, die in den folgenden Kapiteln entwickelt werden.

Wir bezeichnen die Menge der natürlichen Zahlen mit $\mathbb{N} = \{1, 2, 3, \dots\}$. Gelegentlich benutzen wir auch die Schreibweise $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$. Für die Mengen der ganzen, rationalen und reellen Zahlen verwenden wir die üblichen Bezeichnungen $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$. In den rationalen und reellen Zahlen haben wir die vier Rechenoperationen Addition, Subtraktion, Multiplikation und Division (durch Zahlen ungleich 0). Dabei gelten die bekannten Rechenregeln. Später werden wir die Gesetzmäßigkeiten dieser Rechenoperationen mit Hilfe der Körperaxiome präzise formulieren.

Im Folgenden betrachten wir lineare Gleichungen zunächst in den reellen Zahlen. Analoge Aussagen gelten für lineare Gleichungen in beliebigen Körpern.

1.1 Eine Gleichung und eine Variable

Die einfachste lineare Gleichung ist

$$ax = b. \tag{1.1}$$

Dabei sind a , b feste gegebene (reelle) Zahlen, und x ist eine Variable. Wir suchen nach Lösungen der Gleichung, das heißt wir suchen nach reellen Zahlen x , für die (1.1) gilt.

Beispiel 1.1. Seien $a = 2$ und $b = 3$. Dann hat die Gleichung die Form

$$2x = 3.$$

Hier gibt es genau eine Lösung für x , nämlich $x = 3/2$.

Allgemein hat man verschiedene Fälle zu unterscheiden.

1. $a \neq 0$. Hier ist $x = b/a$ die einzige Lösung von (1.1).
2. $a = 0$. Hier gibt es zwei Unterfälle:
 - 2.1. $b = 0$. Jede Zahl x löst (1.1).
 - 2.2. $b \neq 0$. Keine Zahl x löst (1.1).

Man kann dies präzise formulieren, indem man die Lösungsmenge

$$M = \{x \in \mathbb{R} \mid ax = b\}$$

angibt. Es gilt in den einzelnen Fällen:

1. $M = \{b/a\}$.

2.1. $M = \mathbb{R}$.

2.2. $M = \emptyset$ (leere Menge).

1.2 m Gleichungen und n Variablen

Ein lineares Gleichungssystem mit m Gleichungen und n Variablen hat die Form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & = & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m \end{array} \quad (1.2)$$

Dabei sind die Koeffizienten a_{ij} und die b_i vorgegeben und wir suchen nach Lösungen $x_1, \dots, x_n$, für die (1.2) richtig ist.

Beispiel 1.2. Das lineare Gleichungssystem

$$\begin{array}{ccccccc} 2x_1 & + & 3x_2 & - & x_3 & = & 2 \\ -x_1 & - & 2x_2 & + & 3x_3 & = & 0 \end{array}$$

besteht aus 2 Gleichungen in 3 Variablen.

Um Schreibarbeit zu sparen gibt man häufig nur die *Koeffizientenmatrix* an:

$$G = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}.$$

Manchmal betrachtet man auch nur die *engere Koeffizientenmatrix*

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

Unter einer *Matrix* versteht man ein rechteckiges Schema von Zahlen. Hat man m Zeilen und n Spalten, so spricht man auch von einer $m \times n$ -*Matrix*.

Zeilen- und *Spaltenvektoren* sind spezielle Matrizen. Ein Zeilenvektor der Länge n ist eine $1 \times n$ -Matrix, man schreibt $(x_1, \dots, x_n)$. Ein Spaltenvektor der Länge m ist eine $m \times 1$ -Matrix und hat die Form $\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$.

Eine $m \times n$ -Matrix besteht aus n Spaltenvektoren der Länge m , bzw. aus m Zeilenvektoren der Länge n .

Weitere spezielle Typen von Matrizen sind *quadratische Matrizen*, *obere Dreiecksmatrizen* und *Diagonalmatrizen*.

Quadratische Matrizen sind $n \times n$ -Matrizen. Sie haben die Form

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Obere Dreiecksmatrizen sind quadratische Matrizen für die

$$a_{ij} = 0 \text{ für alle } i > j$$

gilt. Dies bedeutet, dass bei ihnen alle Einträge unterhalb der Diagonalen Null sind. Diagonalmatrizen sind quadratische Matrizen welche

$$a_{ij} = 0 \text{ für alle } i \neq j$$

erfüllen. Bei ihnen sind alle Einträge außerhalb der Diagonalen Null.

Beispiel 1.3.

- $\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ ist eine obere Dreiecksmatrix,
- $\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ ist eine Diagonalmatrix.

Definition 1.4. Eine $m \times n$ -Matrix A heißt *Normalformmatrix*, falls sie von der Form

$$A = \left(\begin{array}{cccc|ccc} a_{1,1} & a_{1,2} & \dots & a_{1,r} & a_{1,r+1} & \dots & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,r} & a_{2,r+1} & \dots & a_{2,n} \\ \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a_{r,r} & a_{r,r+1} & \dots & a_{r,n} \\ \hline 0 & \dots & 0 & \vdots & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \vdots & 0 & \dots & 0 \end{array} \right) \quad (1.3)$$

ist, wobei $0 \leq r \leq \min(m, n)$ und $a_{i,i} \neq 0$ für $1 \leq i \leq r$.

Falls $r = 0$, so ist A die Nullmatrix.

Falls $r = m$, so hat jede Zeile mindestens einen von Null verschiedenen Eintrag, nämlich $a_{i,i}$.

Schreibt man

$$A = \left(\begin{array}{c|c} A_1 & A_2 \\ \hline A_3 & A_4 \end{array} \right) \quad (1.4)$$

wobei A_1 Format $r \times r$, A_2 Format $r \times (n - r)$, A_3 Format $(m - r) \times r$ und A_4 Format $(m - r) \times (n - r)$ hat, so lauten die Bedingungen:

- i) A_1 ist obere Dreiecksmatrix mit Diagonaleinträgen ungleich Null.
- ii) A_3 und A_4 sind Nullmatrizen (d.h. alle Einträge sind Null).

Beispiel 1.5.

$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 0 \end{pmatrix}$ ist eine Normalformmatrix mit $m = 3$, $n = 3$ und $r = 2$.

Wir betrachten nun wieder das lineare Gleichungssystem (1.2).

Annahme. Die engere Koeffizientenmatrix

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

habe Normalform wie in (1.3) mit $0 \leq r \leq \min(m, n)$.

Die $(r + 1)$ -te bis m -te Zeile von (1.2) lautet dann

$$0 \cdot x_1 + \dots + 0 \cdot x_n = b_i \quad (i = r + 1, \dots, m).$$

Falls eines der b_i mit $i \in \{r + 1, \dots, m\}$ ungleich 0 ist, so hat (1.2) keine Lösung.

Angenommen $b_{r+1} = \dots = b_m = 0$.

Dann ist (1.2) äquivalent zu

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ & & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ & & & & \ddots & & \vdots & = & \vdots \\ & & & & & & a_{rr}x_r & + & \dots & + & a_{rn}x_n & = & b_r \end{array} \quad (1.5)$$

mit $a_{ii} \neq 0$ für $i = 1, \dots, r$ (und $r \leq n$). In diesem Fall existieren Lösungen und man findet alle Lösungen wie folgt:

- gebe $x_{r+1}, \dots, x_n$ willkürlich vor,
- berechne x_r mit der r -ten Gleichung:

$$x_r = \frac{1}{a_{rr}}(b_r - a_{r,r+1}x_{r+1} - \dots - a_{r,n}x_n),$$

- berechne x_{r-1} mit der $(r-1)$ -ten Gleichung:

$$x_{r-1} = \frac{1}{a_{r-1,r-1}}(b_{r-1} - a_{r-1,r}x_r - \dots - a_{r-1,n}x_n),$$

⋮

- berechne x_1 mit der ersten Gleichung.

Also:

Bemerkung 1.6. Wenn die engere Koeffizientenmatrix des linearen Gleichungssystems (1.2) Normalform hat, so existieren Lösungen genau dann, wenn

$$b_{r+1} = \dots = b_m = 0.$$

In diesem Fall erhält man alle Lösungen, indem man $x_{r+1}, \dots, x_n$ frei vorgibt und damit $x_r, \dots, x_1$ wie oben berechnet.

Beobachtung 1.7. Die Lösungsmenge des Gleichungssystems hat dann $n-r$ „freie Parameter“, nämlich $x_{r+1}, \dots, x_n$. Durch diese sind die übrigen Variablen $x_1, \dots, x_r$ eindeutig bestimmt.

Diese Beobachtung werden wir präzisieren. Zunächst zeigen wir jedoch, dass man zu jedem linearen Gleichungssystem (1.2) ein äquivalentes lineares Gleichungssystem in Normalform finden kann. Dieses kann man dann wie oben lösen.

Idee. Gewisse Umformungen ändern die Lösungsmenge einer linearen Gleichung nicht.

Beispiel 1.8.
$$\begin{cases} ax + by = c \\ dx + ey = f \end{cases} \quad \text{mit Koeffizientenmatrix} \quad \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$$

ist äquivalent zu

- $\begin{cases} ax + by = c \\ (a+d)x + (b+e)y = c+f \end{cases}$ Addieren einer Gleichung zu einer anderen,
- $\begin{cases} rax + rby = rc & (r \neq 0) \\ dx + ey = f \end{cases}$ Multiplizieren einer Gleichung mit einer von Null verschiedenen Zahl,
- $\begin{cases} dx + ey = f \\ ax + by = c \end{cases}$ Vertauschen von zwei Gleichungen,

mit jeweiligen Koeffizientenmatrizen:

- $\begin{pmatrix} a & b & c \\ a+d & b+e & c+f \end{pmatrix},$
- $\begin{pmatrix} ra & rb & rc \\ d & e & f \end{pmatrix},$
- $\begin{pmatrix} d & e & f \\ a & b & c \end{pmatrix}.$

Eine analoge Aussage gilt für lineare Gleichungssysteme mit m Gleichungen in n Variablen.

Vereinbarung. Wir definieren die Summe zweier Zeilenvektoren durch

$$(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1 + b_1, \dots, a_n + b_n)$$

und das Produkt einer Zahl λ mit einem Zeilenvektor durch

$$\lambda \cdot (a_1, \dots, a_n) = (\lambda a_1, \dots, \lambda a_n).$$

Wir verallgemeinern Beispiel 1.8 nun auf lineare Gleichungssysteme mit m Gleichungen in n Variablen.

Definition 1.9. *Elementare Zeilenumformungen* einer Matrix sind:

- i) Addition einer Zeile zu einer anderen,
- ii) Multiplizieren einer Zeile mit einer Zahl ungleich Null,
- iii) Vertauschen zweier Zeilen.

Elementare Spaltenumformungen sind analog definiert.

Bemerkung 1.10. Formt man ein lineares Gleichungssystem durch elementare Zeilenumformungen i)-iii) um, so ändert sich die Lösungsmenge nicht.

Achtung.

- Elementare Spaltenumformungen ändern die Lösungsmenge.
- Für Spaltenvertauschungen iii) ist dies lediglich eine Vertauschung der Variablen.
Beispiel: Vertauschen der 1. und 4. Spalte bedeutet Vertauschen von x_1 und x_4 in den Lösungen.

Satz 1.11. *Jede $m \times n$ Matrix lässt sich durch endlich viele elementare Zeilenumformungen i)-iii) und Spaltenvertauschungen in Normalform bringen.*

Beweis. Benutze das Gauß-Verfahren, welches wir im Folgenden erklären. □

1.2.1 Das Gauß-Verfahren

Bemerkung 1.12. Durch Kombination von i) und ii) kann man folgende Umformung durchführen:

- iv) Addieren des r -fachen der i -ten Zeile zur k -ten Zeile für $r \neq 0$ und $1 \leq i, k \leq m$ mit $i \neq k$.

Beweis.

1. Multipliziere die i -te Zeile mit r .
2. Addiere die (neue) i -te Zeile zur k -ten.
3. Multipliziere die (neue) i -te Zeile mit r^{-1} .

□

Beweis von Satz 1.11. Sei A eine $m \times n$ -Matrix, $A = (a_{ij})$.

Ohne Beschränkung der Allgemeinheit sei $A \neq 0_{m,n}$. (Im Fall $A = 0_{m,n}$ sind wir bereits fertig.)

- Falls nötig Spaltenvertauschung, so dass die erste Spalte ungleich $0_{m,1}$ ist.
- Falls nötig Zeilenvertauschung, so dass die erste Zeile der ersten Spalte ungleich 0 ist.

Damit ist $a_{11} \neq 0$.

- Falls $m = 1$, so ist A nun bereits in Normalform.
- Falls $m > 1$:
Addiere für $i = 2, \dots, m$ zur i -ten Zeile das $-a_{i1}/a_{11}$ -fache der ersten Zeile. Wir erhalten somit eine neue Matrix der Form

$$\left(\begin{array}{c|c} a_{11} & * \\ \hline 0 & B \end{array} \right),$$

wobei B eine $(m-1) \times (n-1)$ -Matrix ist. $*$ steht für Einträge, die wir nicht weiter spezifizieren.

- Ab jetzt führen wir nur noch elementare Zeilenumformungen mit den Zeilen $2, \dots, m$ und Spaltenvertauschungen der Spalten $2, \dots, n$ durch. Damit bleibt die erste Spalte $\begin{pmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ unverändert und wir können beliebige elementare Zeilenumformungen und Spaltenvertauschungen für die Matrix B vornehmen.

- Analog zum vorherigen Argument können wir damit die Matrix A auf die Form

$$\left(\begin{array}{c|c|c} a_{11} & a_{12} & * \\ \hline 0 & a_{22} & * \\ \hline 0 & 0 & C \end{array} \right)$$

bringen, wobei $a_{22} \neq 0$ und C eine $(m-2) \times (n-2)$ -Matrix ist.

- Setze das Verfahren fort.
- Erhalten nach endlich vielen Schritten eine Matrix in Normalform.

□

Frage. Was ist die Bedeutung von r in der Normalform? Hängt sie von den Wahlen (bei Spalten bzw. Zeilenvertauschungen) im Gauß-Algorithmus ab?

Beispiel 1.13. Betrachte das lineare Gleichungssystem

$$\begin{array}{rrrrcl} x_1 & + & 2x_2 & + & 3x_3 & = & 1 \\ -x_1 & - & 2x_2 & + & x_3 & = & 1 \\ 2x_1 & + & 3x_2 & + & 4x_3 & = & 1 \end{array}$$

Die engere Koeffizientenmatrix transformiert sich wie folgt:

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

Für die Koeffizientenmatrix bedeutet dies:

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

was folgendem linearen Gleichungssystem entspricht:

$$\begin{array}{rrrrcl} x_1 & + & 2x_2 & + & 3x_3 & = & 1 \\ & & -x_2 & - & 2x_3 & = & -1 \\ & & & & 4x_3 & = & 2. \end{array}$$

Dieses lineare Gleichungssystem wird durch

$$\begin{array}{rcl} x_3 & = & \frac{1}{2} \\ x_2 & = & 0 \\ x_1 & = & -\frac{1}{2} \end{array}$$

eindeutig gelöst.

1.2.2 Übungen

Aufgabe 1.1. Ein lineares Gleichungssystem heißt homogen, wenn die rechte Seite (also $b_1, \dots, b_m$) gleich 0 ist, und andernfalls inhomogen. Geben Sie jeweils ein Beispiel eines Gleichungssystems der folgenden Art an:

1. Ein homogenes lineares Gleichungssystem mit genau einer Lösung.
2. Ein homogenes lineares Gleichungssystem mit unendlich vielen Lösungen.
3. Ein inhomogenes lineares Gleichungssystem ohne Lösung.
4. Ein inhomogenes lineares Gleichungssystem mit genau einer Lösung.
5. Ein inhomogenes lineares Gleichungssystem mit unendlich vielen Lösungen.

Aufgabe 1.2. Bestimmen Sie die Lösungsmengen der folgenden Gleichungssysteme:

$$\begin{array}{lll}
 (a) \begin{array}{rcl} x & +y & =1 \\ 3x & +3y & =3 \end{array} & (b) \begin{array}{rcl} 2x & -y & +2z=2 \\ 3x & +3y & =3 \\ x & & +2z=2 \end{array} & (c) \begin{array}{rcl} x & -y & +z=1 \\ x & +y & -z=1 \\ x & -y & +z=-1 \end{array}
 \end{array}$$

1.3 Die Struktur der Lösungsmenge

Beispiel 1.14. Für $(a, b) \neq (0, 0)$ ist die Lösungsmenge der Gleichung $ax + by = c$ eine Gerade in der „ (x, y) -Ebene“.

Dies werden wir nun verallgemeinern.

Definition 1.15.

- i) Wir definieren

$$\mathbb{R}^n := \{(x_1, \dots, x_n) \mid x_i \in \mathbb{R}, i = 1, \dots, n\},$$

die „Menge der n -Tupel reeller Zahlen“.

Ein n -Tupel reeller Zahlen ist eine Abbildung

$$x : \{1, \dots, n\} \rightarrow \mathbb{R}, i \mapsto x(i).$$

Diese ist durch Angabe der Werte $x_i := x(i)$ für $i = 1, \dots, n$ eindeutig bestimmt und man schreibt $(x_1, \dots, x_n)$. Es gilt weiterhin „ n -Tupel = Zeilenvektor der Länge n “.

- ii) Die Elemente des $\mathbb{R}^n$ heißen auch *Vektoren*.

- iii) Die Summe zweier Vektoren aus $\mathbb{R}^n$ ist gegeben durch

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n).$$

iv) Das Produkt einer reellen Zahl λ mit einem Vektor aus $\mathbb{R}^n$ ist gegeben durch

$$\lambda \cdot (x_1, \dots, x_n) := (\lambda x_1, \dots, \lambda x_n).$$

Definition 1.16. Ein lineares Gleichungssystem

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

heißt *homogen*, falls $b_1 = \dots = b_m = 0$. Andernfalls heißt es *inhomogen*.

Bemerkung 1.17.

i) Homogene lineare Gleichungssysteme sind lösbar. Sie haben stets die „triviale“ Lösung

$$x = (0, \dots, 0).$$

ii) Ist x eine Lösung eines homogenen Gleichungssystems, so ist auch $\lambda \cdot x$ für alle $\lambda \in \mathbb{R}$ eine Lösung des Gleichungssystems.

iii) Sind x, y Lösungen eines homogenen Gleichungssystems, so auch $x + y$.

Definition 1.18. Eine Teilmenge $W \subset \mathbb{R}^n$ heißt *linearer Teilraum*, falls

i) $0 \in W$,

ii) $a, b \in W \Rightarrow a + b \in W$,

iii) $\lambda \in \mathbb{R}, a \in W \Rightarrow \lambda a \in W$.

Bemerkung 1.19. Die Lösungsmenge eines homogenen linearen Gleichungssystems mit m Gleichungen in n Variablen ist ein linearer Teilraum des $\mathbb{R}^n$.

Satz 1.20. Ein homogenes lineares Gleichungssystem mit m Gleichungen in n Variablen besitzt im Falle $m < n$ stets eine nichttriviale Lösung.

Beweis. Dies folgt aus §1.2, Bemerkung 1.6 und Satz 1.11. □

Die Lösungsmenge eines inhomogenen linearen Gleichungssystems ist kein linearer Teilraum, denn 0 ist keine Lösung.

Definition 1.21. Sei $M \subset \mathbb{R}^n$ eine Teilmenge und $a \in \mathbb{R}^n$. Das *Translat* von M um a ist definiert als

$$a + M = \{a + x \mid x \in M\}.$$

Definition 1.22. Eine Teilmenge $A \subset \mathbb{R}^n$ heißt *affiner Teilraum*, falls $A = \emptyset$ oder A das Translat eines linearen Teilraums $W \subset \mathbb{R}^n$ um einen Vektor $a \in \mathbb{R}^n$ ist, also

$$A = a + W.$$

Bemerkung 1.23. Ist $A = a + W \subset \mathbb{R}^n$, wobei $a \in \mathbb{R}^n$ und $W \subset \mathbb{R}^n$ ein linearer Teilraum ist, so gilt:

- i) $a \in A$, da $0 \in W$.
- ii) Ist $b \in A$, so gilt $A = b + W$.
- iii) $W = \{a - b \mid a \in A, b \in A\}$.
- iv) A ist ein linearer Teilraum, falls $0 \in A$.

Beweis. Übung. □

Satz 1.24. Sei

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m \end{array} \quad (1.6)$$

ein beliebiges lineares Gleichungssystem (homogen oder inhomogen). Sei W die Lösungsmenge des zugehörigen homogenen linearen Gleichungssystems

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & 0 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & 0. \end{array} \quad (1.7)$$

Falls (1.6) eine Lösung $u \in \mathbb{R}^n$ besitzt, so ist die Lösungsmenge M von (1.6) gegeben durch

$$M = u + W.$$

Insbesondere ist die Lösungsmenge von (1.6) ein affiner Teilraum des $\mathbb{R}^n$.

Beweis. $M \supset u + W$: Sei $x \in W$, also x Lösung von (1.7). Es folgt, dass $u + x$ eine Lösung von (1.6) ist.

$M \subset u + W$: Sei $v \in M$. Da $u \in M$, gilt $v - u \in W$ und es folgt

$$v = u + (v - u) \in u + W.$$

□

1.3.1 Übungen

Aufgabe 1.3. Für welche $a \in \mathbb{R}$ hat das lineare Gleichungssystem

$$\begin{aligned}x + (3 - a)y &= 1 \\2x + 6y + 2z &= 2 \\2x + 7y + (a + 2)z &= 4\end{aligned}$$

keine, genau eine oder unendliche viele Lösungen in $\mathbb{R}^3$? Geben Sie jeweils die Lösungsmenge an.

Aufgabe 1.4. Sei $a \in \mathbb{R}^n$ und $W \subset \mathbb{R}^n$ ein linearer Teilraum. Zeigen Sie die folgenden Aussagen (siehe Bemerkung 1.23):

- i) $a \in A$, da $0 \in W$.
- ii) Ist $b \in A$, so gilt $A = b + W$.
- iii) $W = \{a - b \mid a \in A, b \in A\}$.
- iv) A ist ein linearer Teilraum, falls $0 \in A$.

Aufgabe 1.5. Bestimmen Sie alle linearen Teilräume des $\mathbb{R}^2$.

1.4 Lineare Abhängigkeit und Unabhängigkeit

Definition 1.25. Ein m -Tupel von Vektoren $a_1, \dots, a_m \in \mathbb{R}^n$ heißt *linear abhängig*, falls es $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ gibt, so dass

- i) $(\lambda_1, \dots, \lambda_m) \neq (0, \dots, 0)$
- ii) $\lambda_1 a_1 + \dots + \lambda_m a_m = 0 \in \mathbb{R}^n$

(„ $0 \in \mathbb{R}^n$ lässt sich nicht-trivial linear kombinieren.“) Falls das m -Tupel nicht linear abhängig ist, so heißt es *linear unabhängig*.

Achtung. Hier sind die a_i nicht Komponenten eines Vektors, sondern selbst Vektoren

$$a_i = (a_{i1}, \dots, a_{in}).$$

Beispiel 1.26.

- i) $(1, 0), (1, 1) \in \mathbb{R}^2$ sind linear unabhängig.
- ii) $(1, 0), (1, 1), (0, 1) \in \mathbb{R}^2$ sind linear abhängig, denn

$$(1, 0) - (1, 1) + (0, 1) = 0.$$

iii) Für $i \in \{1, \dots, n\}$ sei $e_i := (0, \dots, 0, 1, 0, \dots, 0)$, mit der 1 an der i -ten Stelle, der i -te Einheitsvektor im $\mathbb{R}^n$. Ist $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ beliebig, so gilt

$$x = x_1 e_1 + \dots + x_n e_n.$$

Diese Darstellung ist eindeutig, insbesondere sind die Vektoren $e_1, \dots, e_n$ linear unabhängig.

Bemerkung 1.27. $a_1, \dots, a_m$ sind linear abhängig, falls es einen unter ihnen gibt, etwa a_i , der sich aus den übrigen linear kombinieren lässt, d.h.

$$a_i = \sum_{\substack{j=1 \\ j \neq i}}^m \lambda_j a_j \quad \text{mit } \lambda_j \in \mathbb{R} \text{ geeignet.}$$

Beweis. Übung. □

Bemerkung 1.28. Seien $a_1, \dots, a_m \in \mathbb{R}^n$ linear unabhängig. Dann gilt

- i) $a_i \neq 0$ für alle $i = 1, \dots, m$,
- ii) $a_i \neq a_j$ für alle $i \neq j$.

Definition 1.29. Seien $a_1, \dots, a_m \in \mathbb{R}^n$. Das *lineare Erzeugnis* von $a_1, \dots, a_m$ ist die Menge aller Linearkombinationen

$$\text{Lin}(a_1, \dots, a_m) = \left\{ \sum_{i=1}^m \lambda_i a_i; \lambda_i \in \mathbb{R} \right\} = \mathbb{R}a_1 + \dots + \mathbb{R}a_m.$$

Dies ist ein linearer Teilraum des $\mathbb{R}^n$.

Beispiel 1.30.

- $\text{Lin}((1, 1)) \subset \mathbb{R}^2$ ist die Diagonale in der Ebene.
- $\text{Lin}((1, 0, 0), (0, 1, 0)) \subset \mathbb{R}^3$ ist die (x, y) -Ebene in $\mathbb{R}^3$.

Lemma 1.31. Seien $v_1, \dots, v_r \in \mathbb{R}^n$ und $w_1, \dots, w_s \in \mathbb{R}^n$. Es gelte

- i) $r < s$,
- ii) $w_i \in \text{Lin}(v_1, \dots, v_r)$.

Dann sind $w_1, \dots, w_s$ linear abhängig.

Beweis. Wegen ii) existieren $a_{ij} \in \mathbb{R}$ für $1 \leq i \leq r$ und $1 \leq j \leq s$, so dass

$$\begin{array}{ccccccc} a_{11}v_1 & + & \dots & + & a_{r1}v_r & = & w_1 \\ a_{12}v_1 & + & \dots & + & a_{r2}v_r & = & w_2 \\ \vdots & & & & \vdots & & \vdots \\ a_{1s}v_1 & + & \dots & + & a_{rs}v_r & = & w_s. \end{array}$$

Da $r < s$ hat das homogene lineare Gleichungssystem

$$\begin{array}{ccccccc} a_{11}\lambda_1 & + & \dots & + & a_{1s}\lambda_s & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ a_{r1}\lambda_1 & + & \dots & + & a_{rs}\lambda_s & = & 0 \end{array}$$

nach §1.3 Satz 1.20 eine *nichttriviale* Lösung $(\lambda_1, \dots, \lambda_s) \in \mathbb{R}^s$. Nun gilt

$$\begin{aligned} \lambda_1 w_1 + \dots + \lambda_s w_s &= \lambda_1(a_{11}v_1 + \dots + a_{r1}v_r) \\ &\quad + \lambda_2(a_{12}v_1 + \dots + a_{r2}v_r) \\ &\quad + \dots \\ &\quad + \lambda_s(a_{1s}v_1 + \dots + a_{rs}v_r) \\ &= v_1(a_{11}\lambda_1 + a_{12}\lambda_2 + \dots + a_{1s}\lambda_s) \\ &\quad + v_2(a_{21}\lambda_1 + a_{22}\lambda_2 + \dots + a_{2s}\lambda_s) \\ &\quad + \dots \\ &\quad + v_r(a_{r1}\lambda_1 + a_{r2}\lambda_2 + \dots + a_{rs}\lambda_s) \\ &= 0 + \dots + 0 = 0. \end{aligned}$$

□

Folgerung 1.32. Je $n + 1$ Vektoren im $\mathbb{R}^n$ sind linear abhängig.

Beweis. Nach Beispiel 1.26 (iii) ist

$$\mathbb{R}^n = \text{Lin}(e_1, \dots, e_n).$$

Mit Lemma 1.31 folgt die Aussage. □

Definition 1.33. Eine *Basis* eines linearen Teilraums $W \subset \mathbb{R}^n$ ist ein maximales (d.h. nicht mehr vergrößerbare) System linear unabhängiger Vektoren aus W . Dabei vereinbaren wir, dass der Nullraum $\{0\} \subset \mathbb{R}^n$ die leere Menge $\emptyset$ als Basis hat.

Beispiel 1.34. Nach Folgerung 1.32 bilden $e_1, \dots, e_n$ eine Basis des $\mathbb{R}^n$, die „kanonische Basis“ des $\mathbb{R}^n$.

Bemerkung 1.35. Sei $W \subset \mathbb{R}^n$ ein linearer Teilraum und $v_1, \dots, v_r$ eine Basis von W . Dann ist $r \leq n$.

Beweis. Dies folgt aus Folgerung 1.32. □

Satz 1.36. *Jeder lineare Teilraum $W \subset \mathbb{R}^n$ besitzt eine Basis. Je zwei Basen von W sind gleich lang.*

Beweis. Es genügt den Fall $W \neq \{0\}$ zu betrachten. Angenommen W besitzt keine Basis. Dann gibt es zu einem beliebigen System $v_1, \dots, v_r$ linear unabhängiger Vektoren aus W stets einen weiteren Vektor v_{r+1} , so dass $v_1, \dots, v_{r+1}$ linear unabhängig sind. Mit Induktion findet man linear unabhängige Vektoren $v_1, \dots, v_{n+1} \in W \subset \mathbb{R}^n$. Dies steht im Widerspruch zu Folgerung 1.32 und es folgt der erste Teil der Behauptung.

Seien nun $v_1, \dots, v_r$ und $w_1, \dots, w_s$ zwei Basen von W . Angenommen $r \neq s$. Wir können o.B.d.A. annehmen, dass $r < s$. Nach Lemma 1.31 existiert ein i mit $1 \leq i \leq s$, so dass $w_i \notin \text{Lin}(v_1, \dots, v_r)$. Da die Vektoren $v_1, \dots, v_r$ linear unabhängig sind, sind auch die Vektoren $w_i, v_1, \dots, v_r$ linear unabhängig. Dies steht im Widerspruch zur Maximalität von $v_1, \dots, v_r$ und es folgt der zweite Teil der Behauptung. $\square$

Definition 1.37. Unter der *Dimension* eines linearen Teilraums $W \subset \mathbb{R}^n$ versteht man die Länge einer Basis von W .

Bemerkung 1.38. Seien $a_1, \dots, a_n$ linear unabhängig. Dann bilden die Vektoren eine Basis von

$$\text{Lin}(a_1, \dots, a_n).$$

Beweis. Übung. $\square$

Definition 1.39. Sei $W \subset \mathbb{R}^n$ ein linearer Teilraum. Ein m -Tupel $a_1, \dots, a_m$ von Vektoren heißt *Erzeugendensystem* von W , falls

$$W = \text{Lin}(a_1, \dots, a_m).$$

Bemerkung 1.40. Jede Basis von W ist ein Erzeugendensystem von W .

Beweis. Sei $a_1, \dots, a_m$ eine Basis von W . Angenommen $\text{Lin}(a_1, \dots, a_m) \subsetneq W$. Dann gibt es ein $v \in W \setminus \text{Lin}(a_1, \dots, a_m)$. Damit sind die Vektoren $v, a_1, \dots, a_m \in W$ linear unabhängig. Dies steht im Widerspruch zur Maximalität von $a_1, \dots, a_m$ und die Behauptung ist bewiesen. $\square$

Satz 1.41. *Sei $W \subset \mathbb{R}^n$ ein linearer Teilraum und $a_1, \dots, a_m \in W$. Es sind äquivalent:*

- i) $a_1, \dots, a_m$ ist ein maximales System linear unabhängiger Vektoren in W (Basis).
- ii) $a_1, \dots, a_m$ ist ein minimales Erzeugendensystem von W .
- iii) Die Vektoren $a_1, \dots, a_m$ sind linear unabhängig und bilden ein Erzeugendensystem von W .

Beweis. (i) $\Rightarrow$ (ii): Nach Bemerkung 1.40 ist $a_1, \dots, a_m$ ein Erzeugendensystem. Die Minimalität folgt aus der linearen Unabhängigkeit. (Kein a_i ist durch eine Linearkombinationen der anderen darstellbar und darf deshalb nicht weggelassen werden.)

(ii) $\Rightarrow$ (i): Übung.

(i) $\Leftrightarrow$ (iii) ist nun auch klar. $\square$

1.4.1 Übungen

Aufgabe 1.6. Sei $t \in \mathbb{R}$. Bestimmen Sie die Dimension des von

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

erzeugten linearen Unterraums des $\mathbb{R}^3$ in Abhängigkeit von t .

Aufgabe 1.7. Zeigen Sie die folgenden Aussagen:

- (i) Die Vektoren $a_1, \dots, a_m \in \mathbb{R}^n$ sind genau dann linear abhängig, wenn sich einer von ihnen, etwa a_i , aus den übrigen linear kombinieren lässt, also

$$a_i = \sum_{\substack{j=1 \\ j \neq i}}^m \lambda_j a_j$$

mit geeigneten $\lambda_j \in \mathbb{R}$.

- (ii) Sind $a_1, \dots, a_m \in \mathbb{R}^n$ linear unabhängig, so gilt $a_j \neq 0$ für alle j und $a_i \neq -a_j$ für alle i, j .
- (iii) Sind $a_1, \dots, a_m \in \mathbb{R}^n$ linear unabhängig, so lässt sich jedes Element von $\text{Lin}(a_1, \dots, a_m)$ eindeutig als Linearkombination von $a_1, \dots, a_m$ schreiben.

Aufgabe 1.8. Zeigen Sie: Ist $a_1, \dots, a_m \in W$ ein minimales Erzeugendensystem eines linearen Teilraums $W \subset \mathbb{R}^n$, so ist $a_1, \dots, a_m$ ein maximales System linear unabhängiger Vektoren in W .

1.5 Der Rang einer Matrix

Definition 1.42. Der *Zeilenrang* einer Matrix A ist die maximal mögliche Anzahl linear unabhängiger Zeilen von A . Der *Spaltenrang* von A ist analog definiert.

Bemerkung 1.43. Der Zeilenrang von $A \in \mathbb{R}^{m \times n}$ ist die Dimension des von den Zeilen von A aufgespannten linearen Teilraums von $\mathbb{R}^n$. Dies gilt analog für den Spaltenrang.

Beispiel 1.44. Die Matrix

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

hat Zeilenrang = 2 und Spaltenrang = 2.

Satz 1.45. Für jede Matrix gilt Zeilenrang = Spaltenrang.

Beweis. Benutze 3 Lemmata:

Lemma 1.46. *Elementare Zeilenumformungen vom Typ i)-iv) ändern den Zeilen- und Spaltenrang einer Matrix nicht.*

Beweis. Der von den Zeilen aufgespannte lineare Teilraum ändert sich nicht, somit ändert sich auch der Zeilenrang nicht.

Spaltenrang: Übung. (□)

Lemma 1.47. *Elementare Spaltenumformungen vom Typ i)-iv) ändern den Spalten- und Zeilenrang nicht.*

Beweis. Wie Lemma 1.46 (□)

Lemma 1.48. *Jede $m \times n$ -Matrix lässt sich durch elementare Zeilen- und Spaltenumformungen in die Form*

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

überführen.

Beweis. Durch Zeilenumformungen von Typ i)-iii) und Spaltenvertauschungen erreichen wir Normalform

$$\begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix} \text{ mit } A_1 = \begin{pmatrix} a_{11} & * & * & * \\ 0 & a_{22} & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & a_{rr} \end{pmatrix} \text{ und } a_{ii} \neq 0.$$

Wir transformieren die Matrix weiter wie folgt:

- Addiere ein Vielfaches der 1. Spalte zu den anderen Spalten

$$1. \text{ Zeile } \rightsquigarrow (a_{11}, 0, \dots, 0)$$

- Addiere ein Vielfaches der 2. Spalte zu den anderen Spalten

$$2. \text{ Zeile } \rightsquigarrow (0, a_{22}, 0, \dots, 0)$$

$\vdots$

- Wir erhalten eine Matrix der Form

$$\left(\begin{array}{ccc|c} a_{11} & & 0 & \\ & \ddots & & 0 \\ 0 & & a_{rr} & \\ \hline & 0 & & 0 \end{array} \right), \quad a_{ii} \neq 0 \text{ für } i = 1, \dots, r.$$

- Multipliziere die i -te Zeile mit $\frac{1}{a_{ii}}$ für $i = 1, \dots, r$.
- Wir erhalten eine Matrix der Form

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

(□)

Beweis von Satz 1.45. Nach den Lemmata 1.46-1.48 genügt es, die Behauptung für den Fall, dass die Matrix die Form

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

hat, zu zeigen. In diesem Fall ist offenbar

$$\text{Zeilenrang} = r = \text{Spaltenrang}.$$

□

Definition 1.49. Der *Rang* einer Matrix ist definiert als ihr Zeilenrang (oder Spaltenrang).

Bemerkung 1.50. Der Rang einer Matrix in Normalform (vergleiche §1.2) ist gleich der dort auftretenden Zahl r .

Damit können wir unsere Ergebnisse zur Lösungsmenge des linearen Gleichungssystems

$$\begin{array}{ccccccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n & = & b_m \end{array} \quad (1.8)$$

wie folgt zusammenfassen. Sei

$$G = \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix} \text{ und } A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}.$$

Satz 1.51. Das lineare Gleichungssystem (1.8) hat genau dann mindestens eine Lösung, wenn

$$\text{Rang}(G) = \text{Rang}(A).$$

Die Lösungsmenge des zugehörigen homogenen linearen Gleichungssystems ist ein linearer Teilraum der Dimension $n - r$, wobei r für den Rang steht. Die Lösungsmenge des (in-)homogenen Systems ist ein affiner Teilraum des $\mathbb{R}^n$. Dieser ist entweder leer, oder ein Translat des Lösungsraums des homogenen Systems.

1.6 Matrixmultiplikation

Sei $\mathbb{R}^{m \times n}$ die Menge der $m \times n$ -Matrizen. Zunächst: Sind $A, B \in \mathbb{R}^{m \times n}$ mit Einträgen a_{ij} , beziehungsweise b_{ij} , so definiert man die *Summe*

$$C = A + B \in \mathbb{R}^{m \times n}$$

durch $C = (c_{ij})$ mit $c_{ij} = a_{ij} + b_{ij}$. Ist $\lambda \in \mathbb{R}$ ein Skalar, so definiert man

$$\lambda \cdot A \in \mathbb{R}^{m \times n}$$

als die Matrix mit Einträgen $\lambda \cdot a_{ij}$.

Bemerkung 1.52. Sind $A, B \in \mathbb{R}^{m \times n}$ und $\lambda, \mu \in \mathbb{R}$, so gilt:

$$\begin{aligned}\lambda(A + B) &= \lambda A + \lambda B \\ (\lambda + \mu)A &= \lambda A + \mu A \\ (\lambda\mu)A &= \lambda(\mu A).\end{aligned}$$

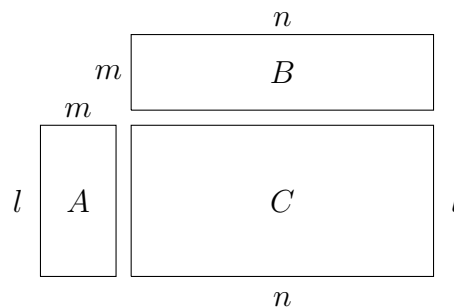
Matrixmultiplikation ordnet einer $l \times m$ -Matrix $A = (a_{ij})$ und einer $m \times n$ -Matrix $B = (b_{jk})$ eine $l \times n$ -Matrix

$$C = A \cdot B$$

zu, wobei $C = (c_{ik})$ mit

$$c_{ik} = \sum_{j=1}^m a_{ij} b_{jk}, \text{ für } i = 1, \dots, l \text{ und } k = 1, \dots, n.$$

Wichtig. Die Formate von A und B müssen „passen“.



Beispiel 1.53. Matrixmultiplikation ist im Allgemeinen nicht kommutativ. Es gilt zum Beispiel

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 1 & 2 \end{pmatrix}.$$

Bemerkung 1.54. Für $A \in \mathbb{R}^{l \times m}$, $B \in \mathbb{R}^{m \times n}$ und $C \in \mathbb{R}^{n \times p}$ gilt das Assoziativgesetz

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

Beweis. Übung. □

Bemerkung 1.55. Für $A \in \mathbb{R}^{l \times m}$, $B \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{m \times n}$ und $D \in \mathbb{R}^{n \times p}$ gilt das Distributivgesetz

$$\begin{aligned} A(B + C) &= AB + AC, \\ (B + C)D &= BD + CD. \end{aligned}$$

Beweis. Übung. □

Lineare Gleichungssysteme

$$\begin{array}{ccccccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n & = & b_m \end{array} \quad (1.8)$$

lassen sich elegant mit Hilfe des Matrixprodukts schreiben. Sei $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ die engere Koeffizientenmatrix. Fasse $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^{n \times 1}$ und $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^{m \times 1}$ als Spaltenvektoren auf. Damit wird (1.8) zu

$$Ax = b.$$

1.6.1 Übungen

Aufgabe 1.9. Zeigen Sie die folgenden Rechenregeln für die Matrixmultiplikation:

1. Seien $A \in \mathbb{R}^{m \times n}$ und $B \in \mathbb{R}^{n \times p}$. Dann gilt $(AB)^t = B^t A^t$.
2. Seien $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{n \times p}$ und $D \in \mathbb{R}^{l \times m}$. Dann gilt: $(A + B)C = AC + BC$ und $D(A + B) = DA + DB$.

Aufgabe 1.10. Es sei E_n die $n \times n$ Einheitsmatrix und E_{ij} die $n \times n$ -Matrix, die an der (i, j) -ten Stelle den Eintrag 1 hat und sonst nur 0. Wir betrachten folgende drei spezielle Typen reeller $n \times n$ Matrizen, sogenannte *Elementarmatrizen*:

- Für $i \neq j$ sei P_{ij} die Matrix, die aus der Einheitsmatrix durch Vertauschen der i -ten und j -ten Spalte entsteht.
- Für $i \neq j$ und $c \neq 0$ sei $Q_{ij}(c) = E_n + cE_{ij}$ die Matrix mit 1 auf der Diagonale, c bei (i, j) , und sonst 0.
- Für $1 \leq i \leq n$ und $\lambda \neq 0$ sei $S_i(\lambda)$ die Diagonalmatrix, deren i -ter Diagonaleintrag λ ist, und die sonst 1 auf der Diagonale hat.

Machen Sie sich für $n = 3$ klar, wie die Elementarmatrizen aussehen. Was bedeutet es anschaulich eine Matrix $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ von links oder rechts mit P_{ij} , Q_{ij} oder $S_i(\lambda)$ zu multiplizieren?

Aufgabe 1.11. Zeigen Sie: Kommutiert eine Matrix $A \in \mathbb{R}^{n \times n}$ mit allen reellen $n \times n$ -Matrizen (also $AB = BA$ für alle $B \in \mathbb{R}^{n \times n}$), so ist A ein skalares Vielfaches der Einheitsmatrix E_n , d.h. $A = \lambda E_n$ für ein $\lambda \in \mathbb{R}$.

Hinweis: Betrachten Sie Matrizen $E_{i,j}$, die an der Stelle (i,j) den Eintrag 1, und sonst 0 haben.

Aufgabe 1.12. Sei $A \in \mathbb{R}^{n \times n}$ mit

$$a_{ij} = \begin{cases} 1, & \text{für } j = i + 1, \\ 0, & \text{sonst.} \end{cases}$$

Berechnen Sie $A^k = \underbrace{A \cdot A \cdots A}_{k \text{ mal}}$ für alle $1 \leq k \leq n$. Benutzen Sie vollständige Induktion nach k .

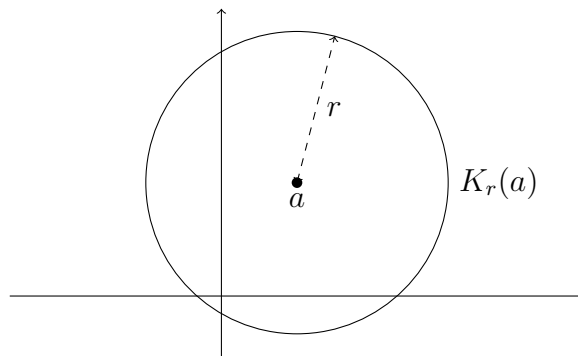
1.7 Etwas analytische Geometrie

1.7.1 Längen und Winkel im $\mathbb{R}^n$

Wir können bereits Geraden im $\mathbb{R}^n$ und allgemeiner auch affine Teilräume beschreiben. Um kompliziertere geometrische Objekte und die Lage von geometrischen Objekten zueinander zu beschreiben, müssen wir Längen und Winkel messen können.

Beispiel 1.56. Um einen Kreis in $\mathbb{R}^2$ mit Mittelpunkt a und Radius r ,

$$K_r(a) = \{x \in \mathbb{R}^2 \mid \text{Abstand von } x \text{ zu } a \text{ ist } r\},$$



zu definieren, benötigen wir einen Abstands begriff. Diesen können wir durch ein Skalarprodukt erhalten.

Definition 1.57. Das *Standardskalarprodukt* zweier Vektoren $x = (x_1, \dots, x_n), y = (x_1, \dots, x_n) \in \mathbb{R}^n$ ist definiert als

$$\langle x, y \rangle = x_1 y_1 + \cdots + x_n y_n = \sum_{k=1}^n x_k y_k.$$

Es besitzt die folgenden Eigenschaften:

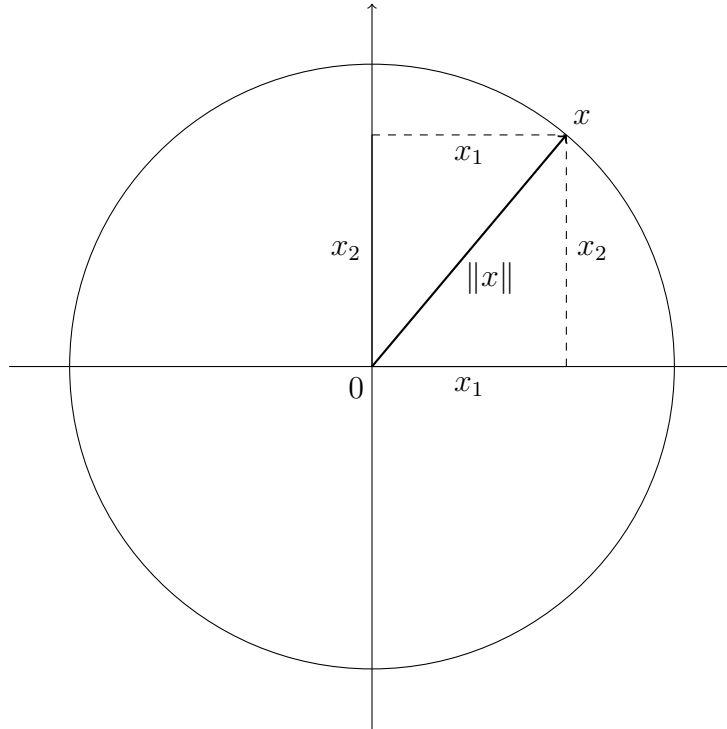
- Symmetrie: $\langle x, y \rangle = \langle y, x \rangle$ für alle $x, y \in \mathbb{R}^n$.
- Linearität in der ersten Variable: $\langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle$ und $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$ für alle $\alpha \in \mathbb{R}$ und $x, x', y \in \mathbb{R}^n$.
- Linearität in der zweiten Variable: $\langle x, y + y' \rangle = \langle x, y \rangle + \langle x, y' \rangle$ und $\langle x, \alpha y \rangle = \alpha \langle x, y \rangle$ für alle $\alpha \in \mathbb{R}$ und $x, y, y' \in \mathbb{R}^n$.
- Positive Definitheit: $\langle x, x \rangle \geq 0$ für alle $x \in \mathbb{R}^n$ und $\langle x, x \rangle = 0 \Leftrightarrow x = 0$.

Die *euklidische Norm* von $x \in \mathbb{R}^n$ ist definiert als

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

Die Größe $\|x\|$ gibt den Abstand von x als Punkt im $\mathbb{R}^n$ zum Ursprung an und bezeichnet auch die *Länge* des Vektors x .

Beispiel 1.58. Für $x = (x_1, x_2) \in \mathbb{R}^2$ ist $\|x\| = \sqrt{x_1^2 + x_2^2}$



Damit ist $K_r(a) = \{x \in \mathbb{R}^2 \mid \|x - a\| = r\} \subset \mathbb{R}^2$. Der *Einheitskreis* (mit Mittelpunkt 0) im $\mathbb{R}^2$ ist

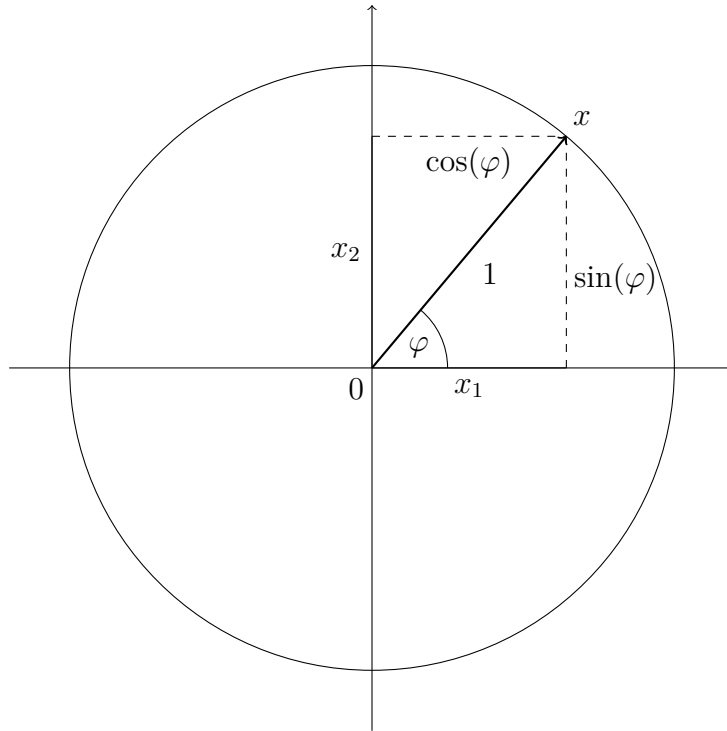
$$\begin{aligned} K_1(0) &= \{x \in \mathbb{R}^2 \mid \|x\| = 1\} \\ &= \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}. \end{aligned}$$

Der Einheitskreis $K_1(0)$ lässt sich auch mit Hilfe des Winkels φ eindeutig parametrisieren. Für $\phi \in \mathbb{R}$ sei

$$e(\phi) = (\cos(\phi), \sin(\phi)).$$

Dann gilt

$$K_1(0) = \{e(\varphi) \mid \varphi \in [0, 2\pi)\}.$$



Polarkoordinaten in $\mathbb{R}^2$

Für $x \in \mathbb{R}^2 \setminus \{0\}$ ist $\frac{1}{\|x\|}x \in K_1(0)$. Dazu gibt es genau ein $\varphi \in [0, \pi)$, so dass $\frac{1}{\|x\|}x = e(\phi)$. Daraus folgt $x = \|x\|e(\phi)$. Zu $x \in \mathbb{R}^2 \setminus \{0\}$ gibt es also genau ein Paar von Zahlen $r, \varphi \in \mathbb{R}_{>0} \times [0, \pi)$, so dass

$$x = r \cdot e(\varphi).$$

Die Zahl r gibt hierbei den Abstand von 0 an. Die Zahl φ den Winkel zwischen der der x_1 -Achse und dem Vektor x .

Winkel zwischen Vektoren

Der Winkel zwischen $e(\varphi)$ und $e(\psi)$ ist $\psi - \varphi$. Das Additionstheorem für den Kosinus und Sinus besagt

$$\begin{aligned}\sin(\varphi \pm \psi) &= \sin(\varphi) \cos(\psi) \pm \cos(\varphi) \sin(\psi), \\ \cos(\varphi \pm \psi) &= \cos(\varphi) \cos(\psi) \mp \sin(\varphi) \sin(\psi).\end{aligned}$$

Folglich gilt für den Winkel $\theta = \psi - \varphi$ zwischen $e(\varphi)$ und $e(\psi)$ die Gleichung

$$\cos(\theta) = \cos(\varphi)\cos(\psi) + \sin(\varphi)\sin(\psi) = \langle e(\varphi), e(\psi) \rangle.$$

Sind nun $x, y \in \mathbb{R}^2 \setminus \{0\}$ beliebige Vektoren, so sind $\frac{1}{\|x\|}x, \frac{1}{\|y\|}y \in K_1(0)$, und damit gilt für den Winkel θ zwischen x und y die Gleichung

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

Also:

- $\langle x, y \rangle$ ist ein Maß für den Winkel zwischen x und y .
- $\|x\| = \sqrt{\langle x, x \rangle}$ ist ein Maß für die Länge von x .

Dies gilt analog im $\mathbb{R}^n$, da die Messung eines Winkels zwischen x und y sich in der von x und y aufgespannten Ebene $\text{Lin}(x, y)$ abspielt. Viele elementargeometrische Sätze lassen sich nun einfach beweisen.

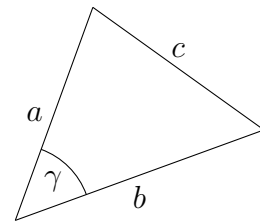
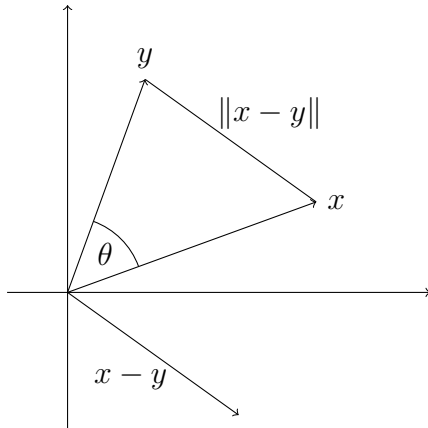
Satz 1.59 (Kosinussatz). *Für $x, y \in \mathbb{R}^2$ gilt*

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\|\cos(\theta),$$

wobei θ den Winkel zwischen x und y bezeichnet. Eine alternative Formulierung für Dreiecke ist durch

$$c^2 = a^2 + b^2 - 2ab\cos(\gamma)$$

gegeben, wobei a, b, c die (Längen der) Seiten des Dreiecks bezeichnen und der Winkel γ der Seite (mit Länge) c gegenüberliegt.



Beweis. Es ist

$$\begin{aligned} \|x - y\|^2 &= \langle x - y, x - y \rangle \\ &= \langle x, x \rangle - \langle x, y \rangle - \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \|y\|^2 - 2\langle x, y \rangle \\ &= \|x\|^2 + \|y\|^2 - 2\|x\|\|y\|\cos(\theta). \end{aligned}$$

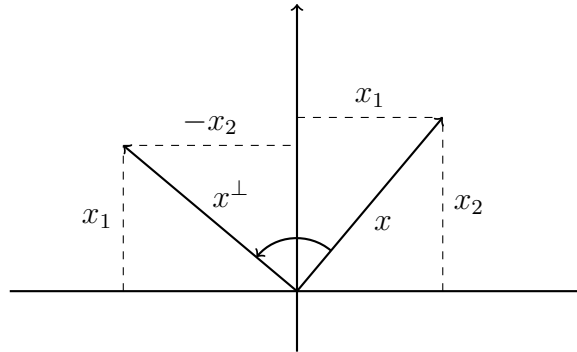
□

Drehung um $\frac{\pi}{2}$

Für $x = (x_1, x_2) \in \mathbb{R}^2$ definieren wir

$$x^\perp = (-x_2, x_1).$$

Der Vektor $x^\perp$ entsteht aus x durch Drehung um $\frac{\pi}{2}$ (Übung).



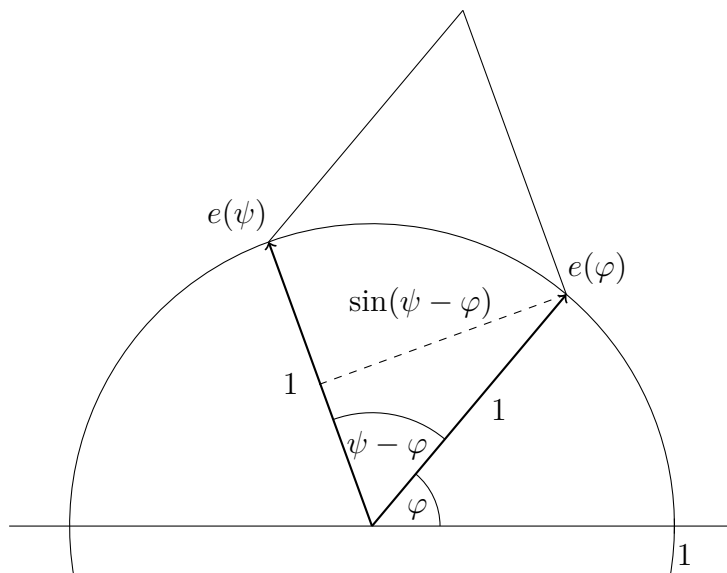
Es gilt $\langle x, x^\perp \rangle = 0$, die Vektoren stehen *orthogonal* zueinander.

Definition 1.60. Für $x, y \in \mathbb{R}^2$ definiert man

$$[x, y] = \langle x^\perp, y \rangle = x_1 y_2 - x_2 y_1 = \det(x, y).$$

Für $e(\varphi), e(\psi) \in K_1(0)$ ist

$$\begin{aligned} [e(\varphi), e(\psi)] &= \cos(\varphi) \sin(\psi) - \sin(\varphi) \cos(\psi) \\ &= \sin(\psi - \varphi) \\ &= \pm \text{Länge der Höhe im Dreieck mit Eckpunkten } 0, e(\varphi), e(\psi) \\ &= \pm \text{Fläche des Parallelogramms, das von } e(\varphi), e(\psi) \text{ aufgespannt wird.} \end{aligned}$$



Daraus folgt der

Satz 1.61. Für $x, y \in \mathbb{R}^2$ ist die Fläche des Parallelogramms mit den Ecken $0, x, y, x + y$ gegeben durch

$$|[x, y]| = |\langle x^\perp, y \rangle|.$$

Bemerkung 1.62. Für $x, y \in \mathbb{R}^2$ gilt

$$\langle x, y \rangle^2 + [x, y]^2 = \|x\|^2 \|y\|^2.$$

Geometrie im Raum

Wie im $\mathbb{R}^2$ ist die Länge eines Vektors x im $\mathbb{R}^3$ durch $\|x\| = \sqrt{\langle x, x \rangle}$ gegeben. Der Kosinus des Winkels θ zwischen $x, y \in \mathbb{R}^3 \setminus \{0\}$ erfüllt weiterhin die Gleichung

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

1.7.2 Das Vektorprodukt

Im $\mathbb{R}^3$ gibt es neben dem Skalarprodukt $\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ auch das *Vektorprodukt (äußere Produkt)* zweier Vektoren, geschrieben

$$\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}, \quad (a, b) \mapsto a \times b.$$

Das Vektorprodukt $a \times b$ ist wieder ein Vektor in $\mathbb{R}^3$. Dieser ist durch

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} = \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ e_1 & e_2 & e_3 \end{pmatrix},$$

wobei die letzte Formulierung symbolisch zu verstehen ist und e_1, e_2, e_3 die Standardbasisvektoren im $\mathbb{R}^3$ bezeichnen.

Bemerkung 1.63. Aus der Definition folgt direkt, dass für die Standardbasisvektoren im $\mathbb{R}^3$ die Gleichungen

$$e_i \times e_i = 0, \quad e_1 \times e_2 = e_3, \quad e_2 \times e_3 = e_1, \quad e_3 \times e_1 = e_2$$

mit $i = 1, 2, 3$ gelten.

Bemerkung 1.64. Das Vektorprodukt hat folgende Eigenschaften:

- $a \times b = -b \times a$,
- $(a + a') \times b = a \times b + a' \times b$, $(\alpha \cdot a) \times b = \alpha \cdot (a \times b)$,
- $a \times (b + b') = a \times b + a \times b'$, $a \times (\alpha \cdot b) = \alpha \cdot (a \times b)$

für alle $\alpha \in \mathbb{R}$ und $a, a', b, b' \in \mathbb{R}^3$. Weiter gilt die Grassmann-Identität

$$a \times (b \times c) = \langle a, c \rangle b - \langle a, b \rangle c$$

und die Jacobi-Identität

$$a \times (b \times c) + b \times (c \times a) + c \times (a \times b) = 0.$$

Das Vektorprodukt ist nicht assoziativ, denn es gilt zum Beispiel

$$(e_1 \times e_1) \times e_2 = 0 \times e_2 = 0 \neq -e_2 = e_1 \times e_3 = e_1 \times (e_1 \times e_2).$$

Bemerkung 1.65. Eine Algebra mit einem antikommutativen Produkt, welches der Jacobi-Identität genügt, nennt man Lie-Algebra. Die Algebra $(\mathbb{R}^3, \times)$ ist eine Lie-Algebra.

Bemerkung 1.66. Für $a, b \in \mathbb{R}^3$ gilt

$$\begin{aligned}\langle a \times b, a \rangle &= 0, \\ \langle a \times b, b \rangle &= 0.\end{aligned}$$

Das Vektorprodukt $a \times b$ steht senkrecht auf der von a und b aufgespannten Ebene.

Beweis. Dies folgt direkt aus $\langle a \times b, c \rangle = \det(a, b, c)$. □

Bemerkung 1.67. Für $a, b \in \mathbb{R}^3$ gilt

$$\langle a, b \rangle^2 + \|a \times b\|^2 = \|a\|^2 \|b\|^2.$$

Beweis. Nachrechnen. □

Folgerung 1.68. Für $a, b \in \mathbb{R}^3$ gilt

$$\|a \times b\| = \|a\| \cdot \|b\| \cdot |\sin(\theta)|,$$

wobei θ den Winkel zwischen a und b bezeichnet.

Beweis. Mit

$$\langle a, b \rangle^2 = \cos^2(\theta) \cdot \|a\|^2 \cdot \|b\|^2$$

gilt

$$\cos^2(\theta) \cdot \|a\|^2 \cdot \|b\|^2 + \|a \times b\|^2 = \|a\|^2 \cdot \|b\|^2.$$

Daraus folgt

$$\|a \times b\|^2 = \|a\|^2 \cdot \|b\|^2 \cdot (1 - \cos^2(\theta)) = \|a\|^2 \cdot \|b\|^2 \cdot \sin^2(\theta)$$

und durch Wurzelziehen die Behauptung. □

Die Länge von $a \times b$ ist durch die Fläche des von a und b aufgespannten Parallelogramms gegeben. Die Orientierung von $a \times b$ ist durch die „Rechte-Hand-Regel“ gegeben (a =Daumen, b =Zeigefinger, $a \times b$ =Mittelfinger, der senkrecht auf der Handfläche steht).

1.8 Anhang: Grundbegriffe

Hier sollen einige Grundlagen zu Mengen, Abbildungen und Relationen zusammengestellt werden.

1.8.1 Mengen

Georg Cantor: „Eine *Menge* ist eine Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.“

Veranschaulichung nach Richard Dedekind: „wie ein Sack, der gewisse Dinge enthält.“

Moderne Mengenlehre: Zermelo-Fraenkel Axiome.

Von einem Objekt muss klar sein, ob es einer Menge M angehört. Jedes Objekt kann in M nur einmal vorkommen. Man spricht in der Mathematik eher von den *Elementen* einer Menge, als von den Objekten. Häufig sind die Elemente Zahlen oder Punkte.

Beispiel 1.69.

- Endliche Mengen, zum Beispiel $\{1, 2, 3, 4\}$.
- Die leere Menge, $\emptyset = \{\}$.
- $\mathbb{N} = \{1, 2, 3, \dots\}$, die Menge der natürlichen Zahlen.
- $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots\}$, die Menge der ganzen Zahlen.
- $\mathbb{Q} = \{\frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0\}$, die Menge der rationalen Zahlen.
- $\mathbb{R}$, die Menge der reellen Zahlen.

Ist a ein Element der Menge M , so schreibt man

$$a \in M.$$

Sind M und N Mengen und ist jedes Element der Menge M auch ein Element der Menge N , so heißt M *Teilmenge* von N . Man schreibt

$$M \subset N.$$

Aus gegebenen Mengen kann man Teilmengen auswählen, die durch eine gewisse Eigenschaft charakterisiert sind.

Beispiel 1.70.

$$\begin{aligned} M &= \{n \in \mathbb{N} \mid n \text{ ist gerade}\} \\ &= \{2, 4, 6, \dots\}. \end{aligned}$$

Zwei Mengen M und N heißen gleich, falls

$$M \subset N \text{ und } N \subset M$$

gilt.

Sind zwei Mengen M und N gegeben, so ist die *Vereinigung*

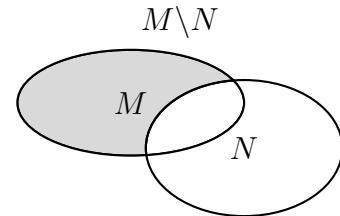
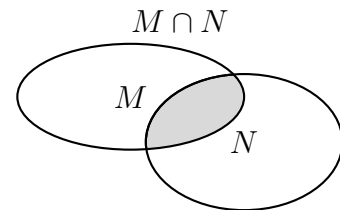
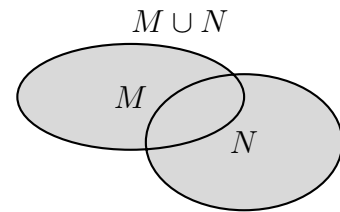
$$M \cup N$$

definiert als die Menge aller x , die in mindestens einer der Mengen M und N enthalten sind. Der *Durchschnitt*

$$M \cap N$$

ist definiert als die Menge aller x , die in sowohl in M als auch in N enthalten sind. Die *Differenzmenge* $M \setminus N$ ist definiert als

$$\{x \in M \mid x \notin N\}.$$



Das *kartesische Produkt* $M \times N$ der Mengen M und N ist die Menge

$$M \times N = \{(m, n) \mid m \in M, n \in N\}$$

aller geordneten Paare von Elementen $m \in M$ und $n \in N$.

Beispiel 1.71.

$$\{1, 2\} \times \{4, 5\} = \{(1, 4), (1, 5), (2, 4), (2, 5)\}.$$

1.8.2 Abbildungen

Seien M und N Mengen.

Definition 1.72. Eine *Abbildung* f von M nach N ist eine Vorschrift, welche *jedem* Element $m \in M$ *genau ein* Element $f(m)$ aus N zuordnet. Man schreibt

$$f : M \rightarrow N, m \mapsto f(m).$$

M heißt *Definitionsbereich* von f , und N heißt *Wertebereich* von f . Das *Bild* von f ist die Menge

$$f(M) := \{n \in N \mid \text{es existiert ein } m \in M \text{ mit } f(m) = n\}.$$

Das *Urbild* einer Teilmenge $A \subset N$ ist die Teilmenge

$$f^{-1}(A) := \{m \in M \mid f(m) \in A\} \subset M.$$

Der *Graph* von f ist die Menge

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

Beispiel 1.73.

- „Multiplikation mit 2“ $f : \mathbb{N} \rightarrow \mathbb{Z}, x \mapsto f(x) = 2x$,
- $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x) = x^2$.

Definition 1.74. Eine Abbildung $f : M \rightarrow N$ heißt

- surjektiv*, falls es zu jedem $n \in N$ ein $m \in M$ mit $f(m) = n$ gibt. Dies kann man als $f(M) = N$ schreiben. Eine Abbildung ist surjektiv, wenn jedes Element (mindestens) ein Urbild besitzt.
- injektiv*, falls für alle $m, m' \in M$ mit $f(m) = f(m')$ gilt, dass $m = m'$. Eine Abbildung ist injektiv, wenn jedes Element höchstens ein Urbild hat.
- bijektiv*, falls f injektiv und surjektiv ist. Eine Abbildung ist bijektiv, wenn jedes Element genau ein Urbild besitzt.

Bemerkung 1.75. Ist $f : M \rightarrow N$ bijektiv, so besteht für jedes $n \in N$ das Urbild $f^{-1}(\{n\})$ aus genau einem Element. Bezeichnen wir dies mit $g(n)$, so definiert

$$g : N \rightarrow M, \quad n \mapsto g(n)$$

eine Abbildung, die *Umkehrabbildung* von f . Es gilt

$$\begin{aligned} g(f(m)) &= m \text{ für alle } m \in M, \\ f(g(n)) &= n \text{ für alle } n \in N. \end{aligned}$$

Für die Umkehrabbildung g schreiben wir auch f^{-1} .

Definition 1.76. Seien A, B und C Mengen, $f : A \rightarrow B$ und $g : B \rightarrow C$ Abbildungen. Die Abbildung

$$g \circ f : A \rightarrow C, \quad a \mapsto g(f(a))$$

heißt *Komposition* (Hintereinanderausführung) der Abbildungen f und g .

Definition 1.77. Sei M eine Menge. Die Abbildung

$$\text{id}_M : M \rightarrow M, \quad \text{id}_M(m) = m$$

heißt *Identität* auf M .

Bemerkung 1.78. Sei $f : M \rightarrow N$ eine bijektive Abbildung mit Umkehrabbildung $f^{-1} : N \rightarrow M$. Es gilt

$$\begin{aligned} f^{-1} \circ f &= \text{id}_M, \\ f \circ f^{-1} &= \text{id}_N. \end{aligned}$$

Beispiel 1.79.

- Die Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^2$ ist weder injektiv, noch surjektiv.
- Die Abbildung $f : \mathbb{R}_{>0} \rightarrow \mathbb{R}, f(x) = x^2$ ist injektiv, aber nicht surjektiv.
- Die Abbildung $f : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}, f(x) = x^2$ ist bijektiv mit Umkehrabbildung $y \mapsto \sqrt{y}$.

1.8.3 Relationen

Definition 1.80. Eine *Relation* auf einer Menge M ist eine Teilmenge des kartesischen Produkts

$$M \times M.$$

Man schreibt aRb für $(a, b) \in R$.

Anstelle von R verwenden wir oft ein intuitiveres Zeichen.

Beispiel 1.81.

1. Die „Gleichheitsrelation“ ist gegeben durch

$$R = \{(a, a) \mid a \in M\} \subset M \times M,$$

es gilt

$$(a, b) \in R \Leftrightarrow a = b.$$

2. Die „Größerrelation auf $\mathbb{R}$ “ ist gegeben durch

$$R = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x > y\},$$

es gilt

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

3. Jede Abbildung $f : M \rightarrow M$ kann als Relation

$$R = \{(a, f(a)) \mid a \in M\}$$

interpretiert werden. Es gilt

$$(a, b) \in R \Leftrightarrow f(a) = b.$$

4. Für die Relation

$$R = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid a - b \text{ ist gerade}\} \subset \mathbb{Z} \times \mathbb{Z}$$

gilt

$$(a, b) \in R \Leftrightarrow a \text{ und } b \text{ sind beide gerade oder beide ungerade.}$$

Definition 1.82. Eine Relation $R \subset M \times M$ auf einer Menge M heißt *Äquivalenzrelation*, falls mit der Schreibweise

$$a \sim b \Leftrightarrow (a, b) \in R$$

gilt:

- i) Reflexivität: $a \sim a$ für alle $a \in M$.
- ii) Symmetrie: $a \sim b \Rightarrow b \sim a$.
- iii) Transitivität: $a \sim b$ und $b \sim c \Rightarrow a \sim c$.

Beispiel 1.83. Die „Gleichheitsrelation“ und die „gerade/ungerade Relation“ (Beispiel 1.81 (1) und (4)) sind Äquivalenzrelationen.

Definition 1.84 (Äquivalenzklassen). Sei $\sim$ eine Äquivalenzrelation auf einer Menge M und $a \in M$. Die Menge

$$[a] = \{b \in M \mid b \sim a\} \subset M$$

heißt *Äquivalenzklasse* von a . Eine Teilmenge $N \subset M$ heißt Äquivalenzklasse, falls es ein $c \in M$ mit $[c] = N$ gibt. Dann heißt c *Repräsentant* der Äquivalenzklasse N .

Beispiel 1.85. In Beispiel 1.81 4 gilt:

$$\begin{aligned} [0] &= \{b \in \mathbb{Z} \mid b \text{ ist gerade}\}, \\ [1] &= \{b \in \mathbb{Z} \mid b \text{ ist ungerade}\}, \\ [2] &= [0], \\ [3] &= [1]. \end{aligned}$$

Bemerkung 1.86. Sei $\sim$ eine Äquivalenzrelation auf einer Menge M . Zwei Äquivalenzklassen $A, B \subset M$ sind entweder gleich ($A = B$) oder disjunkt ($A \cap B = \emptyset$). Insbesondere gilt für $a, b \in M$:

$$[a] = [b] \Leftrightarrow [a] \cap [b] \neq \emptyset \Leftrightarrow a \sim b.$$

M ist die disjunkte Vereinigung der Äquivalenzklassen von $\sim$.

Zur letzten Aussage gilt die folgende Umkehrung:

Bemerkung 1.87. Sei N eine Menge und

$$N = \dot{\bigcup}_{i \in I} A_i$$

die disjunkte Vereinigung von Teilmengen $A_i \subset N$. Dann definiert

$$a \sim b :\Leftrightarrow \text{es gibt ein } i \in I \text{ mit } a, b \in A_i$$

eine Äquivalenzrelation auf N .

Beweis. Übung. □

Definition 1.88. Sei $\sim$ eine Äquivalenzrelation auf einer Menge M . Die *Faktormenge* $M/\sim$ ist die Menge der Äquivalenzklassen in M bezüglich $\sim$.

Definition 1.89. Die *Potenzmenge* einer Menge M ist die Menge aller Teilmengen von M .

$M/\sim$ ist eine Teilmenge der Potenzmenge von M .

Beispiel 1.90. In Beispiel 1.81 (4) gilt:

$$\mathbb{Z}/\sim = \{[0], [1]\}.$$

Dieses Beispiel verallgemeinern wir.

Definition 1.91. Sei $n \in \mathbb{N}$ fest. $a, b \in \mathbb{Z}$ heißen *kongruent modulo n* , falls $a - b$ durch n teilbar ist. Wir schreiben dann

$$a \equiv b \pmod{n}.$$

Dies definiert eine Äquivalenzrelation auf $\mathbb{Z}$. Die Äquivalenzklasse von $a \in \mathbb{Z}$ bezeichnen wir mit $[a]_n$ oder mit $\bar{a}$, falls klar ist, auf welches n sich dies bezieht.

Bemerkung 1.92. Es gilt

$$[a]_n = a + n\mathbb{Z} := \{a + nt \mid t \in \mathbb{Z}\}.$$

Die Menge $\mathbb{Z}/\sim$ der Äquivalenzklassen bezeichnen wir in diesem Fall auch mit $\mathbb{Z}/n\mathbb{Z}$. Mit Division mit Rest sieht man

$$\mathbb{Z}/n\mathbb{Z} = \{[0]_n, [1]_n, \dots, [n-1]_n\}.$$

(Übung.) Dies ist eine Menge mit n Elementen.

Beispiel 1.93. Es gilt

$$4 \equiv 7 \pmod{3},$$

da

$$4 - 7 = -3$$

durch 3 teilbar ist.

Beispiel 1.94. Im Fall $n = 2$ gilt

$$a \equiv b \pmod{2} \Leftrightarrow a - b \text{ ist gerade}$$

und

$$\begin{aligned} [0]_2 &= \{b \in \mathbb{Z} \mid b \text{ ist gerade}\}, \\ [1]_2 &= \{b \in \mathbb{Z} \mid b \text{ ist ungerade}\}. \end{aligned}$$

Es ist

$$\mathbb{Z}/2\mathbb{Z} = \{[0]_2, [1]_2\}.$$

2 Vektorräume

2.1 Gruppen

Definition 2.1. Sei M eine Menge. Eine innere *Verknüpfung* „ $\circ$ “ von M ist eine Vorschrift, welche je zwei Elementen $a, b \in M$ in eindeutiger Weise ein Weiteres zuordnet. Wir können die innere Verknüpfung als Abbildung

$$\circ : M \times M \rightarrow M, \quad (a, b) \mapsto a \circ b$$

auffassen.

Definition 2.2. Eine *Gruppe* ist ein Paar $(G, \circ)$ bestehend aus einer Menge G und einer inneren Verknüpfung $\circ : G \times G \rightarrow G$, $(a, b) \mapsto a \circ b$, mit den Eigenschaften

- G1) Assoziativgesetz: $(a \circ b) \circ c = a \circ (b \circ c)$ für alle $a, b, c \in G$.
- G2) Neutrales Element: Es gibt ein $e \in G$ so dass $e \circ a = a$ für alle $a \in G$ gilt.
- G3) Inverses Element: Für alle $a \in G$ existiert ein $a' \in G$ mit $a' \circ a = e$.

Falls zusätzlich das Kommutativgesetz,

$$a \circ b = b \circ a \quad \text{für alle } a, b \in G,$$

gilt, so heißt die Gruppe *abelsch*.

Bemerkung 2.3. Sei $(G, \circ)$ eine Gruppe. Es gilt:

- i) Sei $e \in G$ ein neutrales Element. Dann gilt $a \circ e = a$ für alle $a \in G$.
- ii) Es gibt genau ein neutrales Element $e \in G$.
- iii) Ist a' invers zu $a \in G$, dann gilt $a \circ a' = e$.
- iv) Zu jedem Element a gibt es genau ein inverses Element a' , welches man auch mit a^{-1} bezeichnet.

Beweis. iii) Zu $a' \in G$ existiert nach G3 ein $a'' \in G$ mit $a'' \circ a' = e$. Es folgt

$$\begin{aligned} a \circ a' &= e \circ (a \circ a') \\ &= (a'' \circ a') \circ (a \circ a') \\ &\stackrel{\text{G1}}{=} a'' \circ ((a' \circ a) \circ a') \\ &\stackrel{\text{G3}}{=} a'' \circ (e \circ a') \\ &\stackrel{\text{G2}}{=} a'' \circ a' \\ &= e. \end{aligned}$$

i) Es gilt

$$\begin{aligned} a \circ e &\stackrel{\text{G3}}{=} a \circ (a' \circ a) \\ &\stackrel{\text{G1}}{=} (a \circ a') \circ a \\ &\stackrel{\text{(iii)}}{=} e \circ a \\ &\stackrel{\text{G2}}{=} a. \end{aligned}$$

ii) Sei e' ein weiteres neutrales Element. Es folgt

$$e' \stackrel{\text{G2}}{=} e \circ e' \stackrel{\text{(i)}}{=} e.$$

iv) Übung. □

Bemerkung 2.4. Sei $(G, \circ)$ eine Gruppe und $a, b \in G$. Es gilt:

- i) $(a^{-1})^{-1} = a$,
- ii) $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$.

Beweis. i) Nach Bemerkung 2.3 (iv) existiert zu a^{-1} genau ein inverses Element. Da $a \circ a^{-1} = e$, ist a das eindeutige inverse Element zu a^{-1} . Es folgt $(a^{-1})^{-1} = a$.

ii) Übung. □

Beispiel 2.5 (Beispiele für Gruppen).

- 1) $(\mathbb{Z}, +)$, die Menge der ganzen Zahlen mit Addition.
Neutrales Element ist 0,
inverses Element zu n ist $-n$.
Dies gilt auch für $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$.
- 2) $(\mathbb{R}^\times, \cdot)$, wobei $\mathbb{R}^\times = \mathbb{R} \setminus \{0\}$.
Neutrales Element ist 1,
inverses Element zu a ist $\frac{1}{a}$.
Dies gilt auch für $(\mathbb{Q}^\times, \cdot)$ und $(\mathbb{R}_{>0}, \cdot)$.
- 3) $(\{\pm 1\}, \cdot)$ ist eine Gruppe mit 2 Elementen. Die Gruppentafel lautet

$\cdot$	1	-1
1	1	-1
-1	-1	1

- 4) Hier betrachten wir für eine natürliche Zahl n die Menge $\mathbb{Z}/n\mathbb{Z}$ der Äquivalenzklassen von $\mathbb{Z}$ bezüglich der Äquivalenzrelation „Kongruenz modulo n “, aus Kapitel 1.8.3. Wir definieren eine innere Verknüpfung auf $\mathbb{Z}/n\mathbb{Z}$ durch

$$[a]_n + [b]_n := [a + b]_n.$$

Dabei sind $a, b \in \mathbb{Z}$ Repräsentanten der entsprechenden Äquivalenzklassen. Wir nennen $[a+b]_n$ auch die *Summe* der Äquivalenzklassen $[a]_n$ und $[b]_n$. Die Definition dieser Summe ist unabhängig von der Wahl der Repräsentanten und damit wohldefiniert. Denn sind $a' \in [a]_n$ und $b' \in [b]_n$ andere Repräsentanten, so gilt

$$[a' + b']_n = [a + b]_n.$$

Die Menge $\mathbb{Z}/n\mathbb{Z}$ wird mit dieser Summe von Äquivalenzklassen zu einer abelschen Gruppe (Übung).

5) Sei M eine Menge. Definiere

$$S(M) = \{f : M \rightarrow M \mid f \text{ bijektiv}\},$$

die Menge der bijektiven Abbildungen von M nach M , und

$$\circ : S(M) \times S(M) \rightarrow S(M), (f, g) \mapsto f \circ g,$$

die Komposition von Abbildungen. Damit ist $(S(M), \circ)$ eine Gruppe.

Neutrales Element ist die Identitätsabbildung, $\text{id}_M : M \rightarrow M, m \mapsto m$.

Inverses Element zu f ist f^{-1} , die Umkehrabbildung zu f .

Die Gruppen in Beispiel 2.5 (1)–(4) sind abelsch (kommutativ), für (5) stimmt dies im Allgemeinen nicht. (Beweis in der Übung.)

2.1.1 Übungen

Aufgabe 2.1. Sei $(G, \circ)$ eine Gruppe mit neutralem Element e . Zeigen Sie:

- (a) Für alle $a, b \in G$ gilt $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$.
- (b) Sei H eine Teilmenge von G . Zeigen Sie, dass $(H, \circ)$ genau dann eine Untergruppe von $(G, \circ)$ ist, wenn H nicht die leere Menge ist und wenn $a \circ b^{-1} \in H$ für alle $a, b \in H$ gilt.

Nun sei G' eine weitere Gruppe mit neutralem Element e' und es sei $f : G \rightarrow G'$ ein Gruppenhomomorphismus. Zeigen Sie:

- (c) Es gilt $f(e) = e'$.
- (d) Für alle $a \in G$ gilt $f(a^{-1}) = f(a)^{-1}$.

Aufgabe 2.2. Sei $n \in \mathbb{N}$ fest. Für $x, y \in \mathbb{Z}$ schreiben wir

$$x \equiv y \pmod{n} \quad :\Leftrightarrow \quad n \text{ teilt } x - y.$$

- (a) Zeigen Sie, dass $\equiv \pmod{n}$ eine Äquivalenzrelation ist.

(b) Wir betrachten nun die sogenannte *Restklasse* einer Zahl $k \in \mathbb{Z}$, definiert durch

$$[k]_n = k + n\mathbb{Z} = \{k + nz : z \in \mathbb{Z}\} \subset \mathbb{Z}.$$

In einer solchen Restklasse liegen also alle Zahlen m , die bei Division durch n den gleichen Rest wie k ergeben. Die Menge aller Restklassen schreibt man als $\mathbb{Z}/n\mathbb{Z}$.

Zeigen Sie $\mathbb{Z}/n\mathbb{Z} = \{[0]_n, [1]_n, \dots, [n-1]_n\}$.

(c) Wir definieren Addition und Multiplikation von Restklassen mod n durch

$$\begin{aligned} [a]_n + [b]_n &:= [a + b]_n, \\ [a]_n \cdot [b]_n &:= [a \cdot b]_n. \end{aligned}$$

Zeigen Sie, dass diese Verknüpfungen wohldefiniert sind, dass sie assoziativ und kommutativ sind, und dass $(\mathbb{Z}/n\mathbb{Z}, +)$ eine abelsche Gruppe ist. Zeigen Sie, dass $(\mathbb{Z}/n\mathbb{Z}, \cdot)$ im Allgemeinen keine Gruppe ist.

Aufgabe 2.3. Zeigen Sie, dass $\mathbb{Z}/n\mathbb{Z}$ genau dann ein Körper ist, wenn n eine Primzahl ist.

Hinweis: Benutzen Sie das *Lemma von Bézout*: Sind $a, b \in \mathbb{Z}$ teilerfremd, so gibt es $x, y \in \mathbb{Z}$ mit $ax + by = 1$.

2.2 Permutationen

Beispiel 2.5 (5) ist insbesondere interessant, wenn

$$M = M_n = \{1, \dots, n\}$$

für ein $n \in \mathbb{N}$.

Definition 2.6. Mit

$$S_n := S(M_n)$$

bezeichnen wir die Menge der bijektiven Selbstabbildungen (*Permutationen*) von M_n . Mit der Komposition als Verknüpfung heißt diese Menge *Symmetrische Gruppe n -ten Grades*.

Notation. Eine Permutation $\sigma \in S_n$ stellt man häufig durch eine Wertetabelle, d.h. durch ein Schema der Form

$$\begin{pmatrix} 1 & 2 & 3 & \dots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \dots & \sigma(n) \end{pmatrix},$$

dar.

Beispiel 2.7. $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \in S_3$ ist die Permutation von $\{1, 2, 3\}$, welche 1 auf 3, 2 auf 2 und 3 auf 1 abbildet.

Beispiel 2.8.

$$\begin{aligned}
S_1 &= \{\text{id}_{\{1\}}\}, \\
S_2 &= \left\{ \text{id}_{\{1,2\}}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\}, \\
S_3 &= \left\{ \text{id}_{\{1,2,3\}}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}.
\end{aligned}$$

Bemerkung 2.9. Die Anzahl der Elemente von S_n ist $n! = n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1$.

Beweis. Übung. □

Definition 2.10. Eine Permutation $\tau \in S_n$ heißt *Transposition*, falls es $i, j \in \{1, \dots, n\}$ mit $i \neq j$ gibt, so dass

$$\tau(i) = j, \quad \tau(j) = i$$

und

$$\tau(k) = k$$

für alle $k \neq i, j$.

Bemerkung 2.11. Es gibt $\binom{n}{2}$ Transpositionen in S_n .

Lemma 2.12. Sei $n \in \mathbb{N}$, $n \geq 2$. Jede Permutation $\sigma \in S_n$ lässt sich als Produkt von endlich vielen Transpositionen schreiben,

$$\sigma = \tau_1 \circ \tau_2 \circ \dots \circ \tau_m.$$

Beweis. Induktion nach n .

$n = 2$: Es gilt $S_2 = \{\text{id}_{\{1,2\}}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\}$. Die Identität entspricht einem leeren Produkt, die Behauptung ist klar.

$n > 2$: Die Behauptung sei für ein n bewiesen. Wir zeigen, dass sie dann auch für $n+1$ gilt. Sei $\sigma \in S_{n+1}$.

1. Fall: $\sigma(n+1) = n+1$.

Dann permutiert σ die Ziffern $1, \dots, n$ und definiert damit ein Element σ' von S_n durch $\sigma' = \sigma|_{M_n}$. Nach Induktionsvoraussetzung ist σ' ein Produkt von Transpositionen aus S_n . Dies liefert eine Darstellung von σ als Produkt von Transpositionen aus S_{n+1} .

2. Fall: $\sigma(n+1) \neq n+1$.

Sei $\tau \in S_{n+1}$ die Transposition, die $n+1$ und $\sigma(n+1)$ vertauscht. Sei $\tilde{\sigma} = \tau \circ \sigma \in S_{n+1}$. Es gilt $\tilde{\sigma}(n+1) = n+1$. Wie im 1. Fall folgt, dass $\tilde{\sigma}$ ein Produkt von Transpositionen ist. Mit $\sigma = \tau \circ \tilde{\sigma}$ folgt die Behauptung. □

Die Darstellung von σ als Produkt von Transpositionen ist *nicht* eindeutig.

Beispiel 2.13. Ist $\tau \in S_n$ eine Transposition, so gilt $\tau^2 = \text{id}_{M_n}$. Daraus folgt auch $\tau^{2n} = \text{id}_{M_n}$ für alle $n \in \mathbb{N}$.

Es gilt aber:

Satz 2.14. *Sei $n \geq 2$ und $\sigma \in S_n$. Sei*

$$\sigma = \tau_1 \circ \dots \circ \tau_m \quad (2.1)$$

eine Darstellung von σ als Produkt von Transpositionen. Setze

$$\operatorname{sgn}(\sigma) := (-1)^m \in \{\pm 1\}.$$

Dann ist $\operatorname{sgn}(\sigma)$ unabhängig von der Wahl der Darstellung in (2.1).

Folgerung 2.15. *Es gibt eine eindeutig bestimmte Abbildung $\operatorname{sgn} : S_n \rightarrow \{\pm 1\}$, so dass*

$$i) \operatorname{sgn}(\sigma \circ \sigma') = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\sigma') \text{ für alle } \sigma, \sigma' \in S_n.$$

$$ii) \operatorname{sgn}(\tau) = -1 \text{ für alle Transpositionen } \tau \in S_n.$$

Beweis von Satz 2.14.

Idee. Finde eine Formel für $\operatorname{sgn}(\sigma)$, die nicht von der Darstellung in (2.1) abhängt. Definiere dazu

$$\Delta : \mathbb{Z}^n \rightarrow \mathbb{Z}, \quad \Delta(x_1, \dots, x_n) = \prod_{1 \leq i < j \leq n} (x_i - x_j).$$

Ist τ eine Transposition, so gilt

$$\Delta(x_1, \dots, x_n) = -\Delta(x_{\tau(1)}, \dots, x_{\tau(n)}).$$

Für ein Produkt von Transpositionen $\sigma = \tau_1 \circ \dots \circ \tau_m$ folgt

$$\Delta(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = (-1)^m \Delta(x_1, \dots, x_n).$$

Es gilt daher

$$\operatorname{sgn}(\sigma) = (-1)^m = \frac{\Delta(x_{\sigma(1)}, \dots, x_{\sigma(n)})}{\Delta(x_1, \dots, x_n)}, \quad (2.2)$$

wobei die rechte Seite unabhängig von der Darstellung von σ als Produkt von Transpositionen ist. $\square$

Bemerkung 2.16. Die Formel (2.2) für $\operatorname{sgn}(\sigma)$ gilt für alle n -Tupel $x = (x_1, \dots, x_n)$ mit paarweise verschiedenen Komponenten $x_1, \dots, x_n$. Insbesondere können wir $x = (1, \dots, n)$ setzen und erhalten die nützliche Formel

$$\begin{aligned} \operatorname{sgn}(\sigma) &= \frac{\Delta(\sigma(1), \dots, \sigma(n))}{\Delta(1, \dots, n)} \\ &= \prod_{1 \leq i < j \leq n} \frac{\sigma(i) - \sigma(j)}{i - j}. \end{aligned}$$

Definition 2.17. Eine Permutation σ heißt *gerade*, falls $\text{sgn}(\sigma) = +1$ und *ungerade*, falls $\text{sgn}(\sigma) = -1$.

Beispiel 2.18. Es gilt

$$\begin{aligned} \text{sgn} \left(\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \right) &= \text{sgn} \left(\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right) \\ &= \text{sgn} \left(\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \right) \cdot \text{sgn} \left(\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right) \\ &= (-1) \cdot (-1) = +1, \end{aligned}$$

die Permutation $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ ist gerade.

2.3 Gruppenhomomorphismen

Gruppenhomomorphismen sind Abbildungen zwischen zwei Gruppen, die mit der Gruppenstruktur verträglich sind. Analog werden später Körperhomomorphismen und Vektorraumhomomorphismen betrachtet.

Definition 2.19. Seien $(G, \circ_G)$ und $(H, \circ_H)$ Gruppen. Ein *Gruppenhomomorphismus* ist eine Abbildung $f : G \rightarrow H$, so dass

$$f(g \circ_G g') = f(g) \circ_H f(g')$$

für alle $g, g' \in G$.

Beispiel 2.20.

- $\exp : (\mathbb{R}, +) \rightarrow (\mathbb{R}_{>0}, \cdot), x \mapsto e^x$.
- $\text{sgn} : S_n \rightarrow \{\pm 1\}, \sigma \mapsto \text{sgn}(\sigma)$.
- Sei $M = \{g, u\}$ eine zweielementige Menge. Definiere darauf eine Addition $\oplus$ durch

$$\begin{array}{c|cc} \oplus & g & u \\ \hline g & g & u \cdot \\ u & u & g \end{array}$$

Damit ist $(M, \oplus)$ eine Gruppe und die Abbildung

$$\begin{aligned} \pi : (\mathbb{Z}, +) &\rightarrow (M, \oplus), \\ a &\mapsto \begin{cases} g, & \text{falls } a \text{ gerade ist,} \\ u, & \text{falls } a \text{ ungerade ist,} \end{cases} \end{aligned}$$

ein Gruppenhomomorphismus.

- Für $n \in \mathbb{N}$ ist die Abbildung

$$\begin{aligned} (\mathbb{Z}, +) &\rightarrow (\mathbb{Z}/n\mathbb{Z}, +), \\ a &\mapsto [a]_n \end{aligned}$$

ein Gruppenhomomorphismus.

Bemerkung 2.21. Ein Gruppenhomomorphismus bildet das neutrale Element von G auf das neutrale Element von H ab.

Beweis. Sei e_G das neutrale Element von G und e_H das neutrale Element von H . Es gilt $f(e_G) = f(e_G \circ_G e_G) = f(e_G) \circ_H f(e_G)$. Multiplizieren wir diese Gleichung mit $(f(e_G))^{-1}$, so folgt $e_H = f(e_G)$. $\square$

Definition 2.22. Seien G und H Gruppen. Ein *Gruppenisomorphismus* ist ein bijektiver Gruppenhomomorphismus.

Beispiel 2.23.

- Die Exponentialfunktion $\exp : (\mathbb{R}, +) \rightarrow (\mathbb{R}_{>0}, \cdot)$ ist ein Gruppenisomorphismus.
- Sei $(M, \oplus)$ die Gruppe aus Beispiel 2.20. Die Abbildung

$$\begin{aligned} M &\rightarrow \mathbb{Z}/2\mathbb{Z}, \\ g &\mapsto [0]_2, \\ u &\mapsto [1]_2, \end{aligned}$$

ist ein Gruppenisomorphismus.

Isomorphe Gruppen sind von ihrer Struktur als gleich anzusehen.

Definition 2.24. Eine Teilmenge $H \subset G$ einer Gruppe G heißt Untergruppe, falls sie die folgenden Bedingungen erfüllt:

- i) $e \in H$,
- ii) $a, b \in H \Rightarrow ab \in H$,
- iii) $a \in H \Rightarrow a^{-1} \in H$.

Bemerkung 2.25. Eine Teilmenge H einer Gruppe G ist genau dann eine Untergruppe, wenn $\emptyset \neq H$ und

$$a, b \in H \Rightarrow ab^{-1} \in H.$$

Beweis. Übung. $\square$

Bemerkung 2.26. Es gilt:

- H ist eine Gruppe.
- Die natürliche Inklusion $i : H \rightarrow G$ ist ein Gruppenhomomorphismus.

Beispiel 2.27.

- $(\mathbb{Z}, +) \subset (\mathbb{R}, +)$.
- Die Menge der geraden Zahlen, $(2\mathbb{Z}, +) \subset (\mathbb{Z}, +)$.

2.3.1 Übungen

Aufgabe 2.4. Entscheiden Sie jeweils, ob es sich bei den folgenden Abbildungen um Gruppenhomomorphismen handelt.

- (a) $(\mathbb{R}, +) \rightarrow (\mathbb{R}^\times, \cdot), x \mapsto e^x.$
- (b) $(\mathbb{R}^\times, \cdot) \rightarrow (\mathbb{R}, +), x \mapsto x^{-1}.$
- (c) $(\mathbb{R}, +) \rightarrow (\mathbb{R}, +), x \mapsto x^2.$
- (d) $(\{\pm 1\}, \cdot) \rightarrow (\{\pm 1\}, \cdot), x \mapsto x^2.$
- (e) $(\mathbb{R}^2, +) \rightarrow (\mathbb{R}, +), (x, y) \mapsto 17x - 3y.$
- (f) $(\mathbb{Z}, +) \rightarrow (\mathbb{Z}/n\mathbb{Z}, +), x \mapsto x + n\mathbb{Z}$ für $n \in \mathbb{N}.$
- (g) $(\mathbb{C}, +) \rightarrow (\mathbb{C}, +), z \mapsto z + 1.$
- (h) $(\mathbb{C}^\times, \cdot) \rightarrow (\mathbb{C}^\times, \cdot), z \mapsto iz.$

Aufgabe 2.5. Sei $(G, \circ)$ eine abelsche Gruppe und $(H, \circ)$ eine Untergruppe. Wir definieren eine Relation auf G mittels

$$x \equiv y \pmod{H} \quad :\Leftrightarrow \quad x^{-1} \circ y \in H.$$

Zeigen Sie die folgenden Aussagen:

- (a) Die Relation $\equiv \pmod{H}$ definiert eine Äquivalenzrelation auf G .
- (b) Die Äquivalenzklassen haben die Form $x \circ H = \{x \circ h : h \in H\}$ mit $x \in G$. Man schreibt G/H für die Menge aller Äquivalenzklassen.
- (c) Die Vorschrift $(x \circ H) \circ (y \circ H) := (x \circ y) \circ H$ liefert eine wohldefinierte innere Verknüpfung auf G/H mit der $(G/H, \circ)$ zu einer abelschen Gruppe wird.
- (d) Die Abbildung

$$G \rightarrow G/H, \quad x \mapsto x \circ H$$

definiert einen surjektiven Gruppenhomomorphismus.

Aufgabe 2.6. Für $n \in \mathbb{N}$ sei $K_n[x] = \left\{ \sum_{j=0}^n a_j x^j : a_0, \dots, a_n \in K \right\}$ der K -Vektorraum der Polynome vom Grad $\leq n$ mit Koeffizienten in K . Wir betrachten die Abbildung

$$D : K_n[x] \rightarrow K_{n-1}[x], \quad \sum_{j=0}^n a_j x^j \mapsto \sum_{j=0}^n j a_j x^{j-1}.$$

Zeigen Sie:

- (a) Die Abbildung $D : (K_n[X], +) \rightarrow (K_{n-1}[X], +)$ ist ein surjektiver Gruppenhomomorphismus.
- (b) Es gilt die Produktregel $D(p \cdot q) = (Dp) \cdot q + p \cdot (Dq)$ für $p, q \in K_n[x]$.

2.4 Körper

Definition 2.28. Ein *Körper* ist ein Tripel $(K, +, \cdot)$ bestehend aus einer Menge K und zwei inneren Verknüpfungen $+$ (Addition) und $\cdot$ (Multiplikation), so dass gilt:

- K1) $(K, +)$ ist eine abelsche Gruppe. Ihr neutrales Element wird mit 0 bezeichnet.
- K2) $(K^\times, \cdot)$ ist eine abelsche Gruppe, wobei $K^\times = K \setminus \{0\}$. Ihr neutrales Element wird mit 1 bezeichnet. Insbesondere ist für $a, b \in K^\times$ auch $a \cdot b \in K^\times$.
- K3) Es gelten die Distributivgesetze: $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ und $(b + c) \cdot a = (b \cdot a) + (c \cdot a)$ für alle $a, b, c \in K$.

Notation.

- Für $a \cdot b$ schreibt man häufig ab .
- Man vereinbart üblicherweise, dass die Multiplikation stärker bindet als die Addition. Damit ist $(a \cdot b) + (a \cdot c) = ab + ac$.

Bemerkung 2.29. Sei $(K, +, \cdot)$ ein Körper. Es gilt:

- i) $0 \cdot a = a \cdot 0 = 0$ für alle $a \in K$.
- ii) Sind $a, b \in K$ und $a \cdot b = 0$, so folgt $a = 0$ oder $b = 0$.
- iii) Für alle $a, b \in K$ gilt $a \cdot (-b) = (-a) \cdot b = -(a \cdot b)$.
- iv) Für alle $a, b \in K$ gilt $(-a) \cdot (-b) = a \cdot b$.

Beweis. Übung. □

Beispiel 2.30.

- 1) $(\mathbb{Q}, +, \cdot)$, der Körper der rationalen Zahlen.
- 2) $(\mathbb{R}, +, \cdot)$, der Körper der reellen Zahlen.
- 3) Sei $M = \{g, u\}$ eine zweielementige Menge. Definiere

$$\begin{array}{c|cc} \oplus & g & u \\ \hline g & g & u \\ u & u & g \end{array} \quad \text{und} \quad \begin{array}{c|cc} \odot & g & u \\ \hline g & g & g \\ u & g & u \end{array}$$

Dann ist $(M, \oplus, \odot)$ ein Körper.

- 4) Für ein $n \in \mathbb{N}$ definieren wir eine Multiplikation auf $\mathbb{Z}/n\mathbb{Z}$ durch $[a]_n \cdot [b]_n = [a \cdot b]_n$. Es gilt dann

$$(\mathbb{Z}/n\mathbb{Z}, +, \cdot) \text{ ist ein Körper} \iff n \text{ ist eine Primzahl.}$$

Für eine Primzahl p wird der Körper $\mathbb{Z}/p\mathbb{Z}$ mit $\mathbb{F}_p$ bezeichnet. (Beweis: Übung.)

- 5) $(\mathbb{R} \times \mathbb{R}, +, \cdot)$ ist ein Körper, wobei

$$(a, b) + (a', b') = (a + a', b + b')$$

und

$$(a, b) \cdot (a', b') = (aa' - bb', ab' + a'b).$$

Das neutrale Element der Addition ist $(0, 0)$, das Inverse von (a, b) ist $(-a, -b)$. Das neutrale Element der Multiplikation ist $(1, 0)$, das Inverse von $(a, b) \neq (0, 0)$ ist $(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2})$. Dieser Körper heißt Körper der *komplexen Zahlen*, er wird mit $\mathbb{C}$ bezeichnet.

Die Abbildung $\mathbb{R} \rightarrow \mathbb{C}$, $r \mapsto (r, 0)$ ist injektiv. Es gilt

$$(r, 0) + (r', 0) = (r + r', 0)$$

und

$$(r, 0) \cdot (r', 0) = (rr', 0).$$

Daher können wir $\mathbb{R}$ mit dem Teilkörper $\mathbb{R} \times \{0\} = \{(a, b) \in \mathbb{C} \mid b = 0\}$ von $\mathbb{C}$ identifizieren und $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ auffassen.

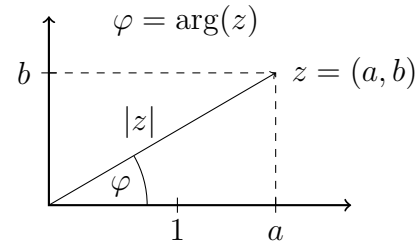
Notation.

- $i := (0, 1) \in \mathbb{C}$ heißt imaginäre Einheit.
- Es gilt $i^2 = -1$.
- Jedes $z = (a, b) \in \mathbb{C}$ hat eine eindeutige Darstellung $(a, b) = (a, 0) + (0, b) = a + bi$.
- $\Re(z) := a$ heißt *Realteil* von z .
- $\Im(z) := b$ heißt *Imaginärteil* von z .
- $\bar{z} = a - bi = (a, -b)$ heißt zu z *komplex konjugierte Zahl*.
- $|z| = \sqrt{z \cdot \bar{z}} = \sqrt{a^2 + b^2}$ heißt *Betrag* von z .

Bemerkung 2.31. Für $0 \neq z \in \mathbb{C}$ gilt $z^{-1} = \frac{\bar{z}}{|z|^2}$.

Zur geometrischen Veranschaulichung kann man die komplexen Zahlen in der Gaußschen Zahlenebene darstellen.

- Addition in $\mathbb{C}$: Vektoraddition in $\mathbb{R}^2$.
- Multiplikation in $\mathbb{C}$: Beträge werden multipliziert, Argumente (Winkel) werden addiert.



2.5 Vektorräume

Erinnerung: Für $\mathbb{R}^n$ gibt es die Verknüpfungen

$$\begin{aligned} + : \mathbb{R}^n \times \mathbb{R}^n &\rightarrow \mathbb{R}^n && \text{Vektoraddition,} \\ \cdot : \mathbb{R} \times \mathbb{R}^n &\rightarrow \mathbb{R}^n && \text{Multiplikation mit einem Skalar.} \end{aligned}$$

Wir haben in Kapitel 1 lineare Teilräume $W \subset \mathbb{R}^n$ betrachtet. Wir wollen dies verallgemeinern.

Definition 2.32. Sei K ein Körper. Ein *Vektorraum* über K (auch K -Vektorraum) ist ein Tripel $(V, +, \cdot)$ bestehend aus einer Menge V und zwei Verknüpfungen

$$\begin{aligned} + : V \times V &\rightarrow V, & (a, b) &\mapsto a + b, \\ \cdot : K \times V &\rightarrow V, & (\lambda, a) &\mapsto \lambda \cdot a, \end{aligned}$$

so dass gilt:

V1) $(V, +)$ ist eine abelsche Gruppe.

V2) Für die Multiplikation mit Skalaren gilt:

- a) $(\lambda + \mu) \cdot a = \lambda \cdot a + \mu \cdot a,$
- b) $\lambda \cdot (a + b) = \lambda a + \lambda b,$
- c) $(\lambda \mu) \cdot a = \lambda \cdot (\mu \cdot a),$
- d) $1 \cdot a = a,$

für alle $\lambda, \mu \in K$ und $a, b \in V$.

Das neutrale Element von $(V, +)$ heißt *Nullvektor*, man bezeichnet es mit 0 oder 0_V .

Beispiel 2.33.

- 1) Die Menge der n -Tupel von Zahlen aus K ,

$$K^n = \{(x_1, \dots, x_n) \mid x_1, \dots, x_n \in K\},$$

ist ein K -Vektorraum.

2) $\mathbb{C}$ ist ein $\mathbb{R}$ -Vektorraum.

3) Sei X eine Menge und K ein Körper. Sei

$$\text{Abb}(X, K) = \text{Menge aller Abbildungen } f : X \rightarrow K.$$

(Man schreibt auch K^X für $\text{Abb}(X, K)$.) Dann ist $\text{Abb}(X, K)$ ein K -Vektorraum mit Addition gegeben durch

$$f + g : X \rightarrow K, \quad (f + g)(x) = f(x) + g(x)$$

und Skalarmultiplikation gegeben durch

$$\lambda \cdot f : X \rightarrow K, \quad (\lambda \cdot f)(x) = \lambda f(x)$$

für $\lambda \in K$ und $f, g \in \text{Abb}(X, K)$. Das neutrale Element ist die Nullabbildung

$$0 : X \rightarrow K, \quad x \mapsto 0.$$

4) Lineare Teilräume des $\mathbb{R}^n$ sind $\mathbb{R}$ -Vektorräume.

5) Beispiele aus der Analysis:

- $\text{Abb}(\mathbb{R}, \mathbb{R})$, der Vektorraum der Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$.
- Der Vektorraum der stetigen Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$.
- Der Vektorraum der differenzierbaren Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$.

Definition 2.34. Sei V ein K -Vektorraum. Eine Teilmenge $W \subset V$ heißt *linearer Teilraum* oder *Untervektorraum*, falls die folgenden Bedingungen erfüllt sind:

- $0 \in W$,
- $a, b \in W \Rightarrow a + b \in W$,
- $\lambda \in K, a \in W \Rightarrow \lambda \cdot a \in W$.

Solch ein W ist ein K -Vektorraum.

Bemerkung 2.35. Sei V ein K -Vektorraum und $a_1, \dots, a_m \in V$. Dann ist das *lineare Erzeugnis*

$$\begin{aligned} \text{Lin}(a_1, \dots, a_n) &= Ka_1 + \dots + Ka_n \\ &= \left\{ \sum_{i=1}^n \lambda_i a_i \mid \lambda_i \in K \right\} \end{aligned}$$

ein Untervektorraum von V .

Definition 2.36. Ein K -Vektorraum heißt *endlich erzeugt*, falls es $a_1, \dots, a_m \in V$ gibt, so dass

$$V = \text{Lin}(a_1, \dots, a_m).$$

Dann heißt $a_1, \dots, a_m$ *Erzeugendensystem* von V .

Beispiel 2.37.

- K^n ist endlich erzeugt.
- Der Vektorraum der stetigen Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ ist nicht endlich erzeugt.

Definition 2.38. Ein m -Tupel von Vektoren $a_1, \dots, a_m$ eines K -Vektorraums V heißt *linear abhängig*, falls es $\lambda_1, \dots, \lambda_m \in K$ gibt, so dass

$$(\lambda_1, \dots, \lambda_m) \neq (0, \dots, 0)$$

und

$$\lambda_1 a_1 + \dots + \lambda_m a_m = 0.$$

Andernfalls heißt das m -Tupel *linear unabhängig*.

Lemma 2.39. Sei V ein K -Vektorraum. Seien $v_1, \dots, v_r \in V$ und $w_1, \dots, w_s \in V$. Es gelte

- i) $r < s$,
- ii) $w_i \in \text{Lin}(v_1, \dots, v_r)$ für alle $i = 1, \dots, s$.

Dann sind $w_1, \dots, w_s \in V$ linear abhängig.

Beweis. Wie in Kapitel 1, §1.4 Lemma 1.31. □

Korollar 2.40. Sei V ein endlich erzeugter K -Vektorraum. Dann existiert ein $d \in \mathbb{N}$, so dass je $d + 1$ Vektoren in V linear abhängig sind.

Definition 2.41. Eine *Basis* eines K -Vektorraums ist ein maximales System linear unabhängiger Vektoren.

Bemerkung 2.42. Eine Basis eines Vektorraums ist ein Erzeugendensystem.

Satz 2.43 (Basisergänzungssatz). Sei V ein endlich erzeugter K -Vektorraum. Jedes System linear unabhängiger Vektoren kann zu einer Basis ergänzt werden. Insbesondere besitzt V eine Basis.

Beweis. Seien $v_1, \dots, v_r \in V$ linear unabhängig. Falls $\text{Lin}(v_1, \dots, v_r) = V$, sind wir fertig. Falls nicht, so existiert ein $v_{r+1} \in V \setminus \text{Lin}(v_1, \dots, v_r)$. Es folgt, dass $v_1, \dots, v_{r+1}$ linear unabhängig sind. Dieses Verfahren setzt man fort, wegen Korollar 2.40 bricht es mit einer Basis $v_1, \dots, v_{d'}$ ab, wobei $r \leq d' \leq d$. □

Satz 2.44. *Sei V ein endlich erzeugter Vektorraum. Je zwei Basen von V sind gleich lang.*

Beweis. Dies folgt aus Lemma 2.39 und Bemerkung 2.42. $\square$

Definition 2.45. Die *Dimension* eines Vektorraumes ist die Länge einer Basis.

Satz 2.46. *Seien $a_1, \dots, a_m \in V$. Es sind äquivalent:*

- i) $a_1, \dots, a_m$ ist ein maximales System linear unabhängiger Vektoren in V (Basis).*
- ii) $a_1, \dots, a_m$ ist ein minimales Erzeugendensystem von V .*
- iii) Die Vektoren $a_1, \dots, a_m$ sind linear unabhängig und bilden ein Erzeugendensystem von V .*

Beweis. Wie in Kapitel 1 §1.4 Satz 1.41. $\square$

Bemerkung 2.47. Sei V ein endlich erzeugter Vektorraum. Aus jedem Erzeugendensystem von V kann eine Basis ausgewählt werden.

Beweis. Übung. $\square$

Bemerkung 2.48. Sei V ein endlich erzeugter Vektorraum und $W \subset V$ ein Untervektorraum. Es gilt:

- i) W ist endlich erzeugt.
- ii) Es gilt $\dim(W) \leq \dim(V)$.
- iii) Falls $\dim(W) = \dim(V)$, so ist $W = V$.

Beweis. i) Angenommen W ist nicht endlich erzeugt. Dann gibt es zu jedem $d \in \mathbb{N}$ Vektoren $a_1, \dots, a_d \in W$ die linear unabhängig sind. Da diese Vektoren auch Elemente von V sind, widerspricht dies Korollar 2.40.

ii) und iii) folgen aus dem Basisergänzungssatz (Satz 2.43). $\square$

Anhang: beliebige Vektorräume

Bemerkung 2.49. Sei V ein beliebiger K -Vektorraum (nicht notwendigerweise endlich erzeugt) und M eine Teilmenge von V . Sei $\text{Lin}(M)$ die Menge aller endlichen Linearkombinationen

$$\sum_{i=1}^r \lambda_i a_i$$

mit $r \in \mathbb{N}$, $\lambda_i \in K$ und $a_i \in M$. Wir nennen $\text{Lin}(M)$ das *lineare Erzeugnis* von M . Es gilt:

- $\text{Lin}(M)$ ist ein Untervektorraum von V .
- M heißt Erzeugendensystem von V , falls $\text{Lin}(M) = V$.

- M heißt linear unabhängig, falls für jede endliche Summe $s = \sum_{i=1}^r \lambda_i a_i$ mit $r \in \mathbb{N}$, $\lambda_i \in K$ und $a_i \in M$ gilt:

$$s = 0 \Rightarrow \lambda_i = 0 \text{ für alle } i = 1, \dots, r.$$

- Man nennt M eine *Basis* von V , falls M linear unabhängig und ein Erzeugendensystem ist.

Man kann mit Hilfe des Lemmas von Zorn zeigen, dass jeder K -Vektorraum V eine Basis besitzt. Eine Basis von V besteht aus endlich vielen Elementen, genau dann wenn V endlich erzeugt ist.

2.5.1 Übungen

Aufgabe 2.7. 1. Entscheiden Sie, ob die folgende Aussage wahr oder falsch ist: Wenn U_1 und U_2 Untervektorräume eines Vektorraums V sind, ist auch $U_1 \cup (U_1 \cap U_2)$ ein Untervektorraum.

2. Entscheiden Sie, ob die folgende Aussage wahr oder falsch ist: Zwei Untervektorräume desselben Vektorraums können disjunkt sein.

3. Im Vektorraum $\mathbb{R}^2$ gelte für paarweise verschiedene Vektoren a, b, c, d , dass $\text{Lin}(a, b) = \text{Lin}(c, d)$. Welche der folgenden Behauptungen sind unter dieser Voraussetzung richtig?

- (a) Das Tupel (a, b, c, d) muss linear abhängig sein.
- (b) a und b müssen linear unabhängig sein.
- (c) a muss weder von c noch von d linear abhängig sein.
- (d) Es gilt stets $\text{Lin}(a, c) = \text{Lin}(b, d)$.

4. Welche der folgenden Behauptungen ist stets in jedem Vektorraum richtig?

- (a) $\text{Lin}(v_1, v_2) = \text{Lin}(v_2, v_1)$.
- (b) $\text{Lin}(v_1, v_2, v_1 + v_2) = \text{Lin}(v_1, v_2)$.
- (c) $\text{Lin}(v_1, v_2, v_1 + v_2) = \text{Lin}(v_1 + v_2)$.

Aufgabe 2.8. Sei V ein endlich erzeugter K -Vektorraum. Zeigen Sie, dass jedes endliche Erzeugendensystem $\{v_1, \dots, v_n\}$ von V eine Teilmenge besitzt, die eine Basis von V ist.

Aufgabe 2.9. Sei V ein Vektorraum und seien $W_1, \dots, W_k \subset V$ Untervektorräume. Wir schreiben $V = W_1 \oplus \dots \oplus W_k$, falls $V = W_1 + \dots + W_k$ und $W_i \cap \sum_{j \neq i} W_j = \{0\}$ für $i \in \{1, \dots, k\}$. In diesem Fall nennen wir V die *direkte Summe* der Unterräume W_i . Zeigen Sie, dass folgende Bedingungen äquivalent sind:

1. $V = W_1 \oplus \dots \oplus W_k$

2. Jedes $v \in V$ ist eindeutig darstellbar als

$$v = w_1 + \dots + w_k$$

mit $w_i \in W_i$.

3 Lineare Abbildungen

Von nun an bezeichnen wir mit $K^n := K^{n \times 1}$ den K -Vektorraum der *Spaltenvektoren* der Länge n .

3.1 Der Begriff der linearen Abbildung

Beispiel 3.1. Sei K ein Körper und $A \in K^{m \times n}$. Ist $x \in K^n$, so ist $y = Ax$ ein Spaltenvektor in K^m . Wir erhalten eine Abbildung

$$\begin{aligned} f_A : K^n &\rightarrow K^m, \\ x &\mapsto Ax. \end{aligned}$$

Für $i = 1, \dots, m$ ist die i -te Komponente von $y = Ax$ durch

$$y_i = \sum_{j=1}^n a_{ij} x_j$$

gegeben. Da Matrixmultiplikation dem Distributivgesetz genügt, gilt

$$f_A(x + y) = A(x + y) = Ax + Ay = f_A(x) + f_A(y)$$

und

$$f_A(\lambda x) = A(\lambda x) = \lambda(Ax) = \lambda f_A(x)$$

für alle $x, y \in K^n$ und $\lambda \in K$.

Die Abbildung f_A ist ein Beispiel für eine lineare Abbildung. Allgemeiner:

Definition 3.2. Seien V und W K -Vektorräume. Eine Abbildung

$$f : V \rightarrow W$$

heißt *linear*, falls

- i) $f(x + y) = f(x) + f(y)$ für alle $x, y \in V$,
- ii) $f(\lambda x) = \lambda f(x)$ für alle $\lambda \in K$ und $x \in V$.

Bemerkung 3.3.

- Lineare Abbildungen heißen auch Vektorraumhomomorphismen.
- Durch Induktion nach n folgt aus i) und ii), dass

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) = \sum_{i=1}^n \lambda_i f(x_i)$$

für alle $\lambda_i \in K$ und $x_i \in V$.

- Es gilt $f(0_V) = 0_W$.

Beispiel 3.4.

1. Für ein $A \in K^{m \times n}$ ist die Abbildung $f_A : K^n \rightarrow K^m$ linear.
2. Ist $U \subset V$ ein Untervektorraum, so ist die Inklusion

$$i : U \rightarrow V, \quad i(u) = u,$$

linear.

3. Sei V ein K -Vektorraum und $a_1, \dots, a_n \in V$. Die Abbildung

$$\varphi : K^n \rightarrow V, \quad \varphi(x) = \sum_{i=1}^n x_i \cdot a_i$$

ist linear.

4. Eine lineare Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}$ hat die Form $x \mapsto ax$, wobei $a = f(1)$.

Bemerkung 3.5. Sei φ die Abbildung aus Beispiel 3.4 (3). Es gilt

- i) φ ist genau dann surjektiv, wenn $a_1, \dots, a_n$ ein Erzeugendensystem von V ist.
- ii) φ ist genau dann injektiv, wenn die Vektoren $a_1, \dots, a_n$ linear unabhängig sind.

Beweis. Übung. □

Bemerkung 3.6. Sind V, W und X drei K -Vektorräume und

$$f : V \rightarrow W, \quad g : W \rightarrow X$$

lineare Abbildungen, so ist $g \circ f$ linear.

Beweis. Übung. □

Bemerkung 3.7. Sei V ein endlich erzeugter K -Vektorraum und $b_1, \dots, b_n$ eine Basis von V . Dann lässt sich jedes $v \in V$ in *eindeutiger Weise* als

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

mit $\lambda_i = \lambda_i(v) \in K$ darstellen.

Beweis. Dass es eine Darstellung gibt, folgt aus der Tatsache, dass eine Basis ein Erzeugendensystem ist. Sei

$$v = \sum_{i=1}^n \mu_i b_i,$$

mit $\mu_i \in K$, ein zweite Darstellung. Es folgt

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

Da die Basisvektoren $b_1, \dots, b_n$ linear unabhängig sind gilt

$$\lambda_i = \mu_i$$

für alle $i = 1, \dots, n$. Damit ist die Darstellung eindeutig. □

Daraus folgern wir:

Satz 3.8. *Seien V und W K -Vektorräume. Sei $b_1, \dots, b_n$ eine Basis von V und $w_1, \dots, w_n$ beliebige Vektoren in W . Dann gibt es genau eine lineare Abbildung*

$$f : V \rightarrow W$$

mit

$$f(b_i) = w_i$$

für alle $i = 1, \dots, n$.

Beweis. Wir definieren f durch lineare Fortsetzung der Bedingung $f(b_i) = w_i$ für $i = 1, \dots, n$. Ist $v = \sum_{i=1}^n \lambda_i \cdot b_i \in V$, so setzen wir

$$\begin{aligned} f(v) &= f\left(\sum_{i=1}^n \lambda_i b_i\right) \\ &= \sum_{i=1}^n \lambda_i f(b_i) \\ &= \sum_{i=1}^n \lambda_i w_i. \end{aligned}$$

Wegen Bemerkung 3.7 ist die Abbildung f wohldefiniert. Wie man leicht nachprüft, ist f linear. Aufgrund von $f(b_i) = w_i$ für alle $i = 1, \dots, n$ und der Linearität ist f eindeutig bestimmt. □

3.2 Kern und Bild

Definition 3.9. Sei $f : V \rightarrow W$ eine lineare Abbildung. Dann definiert man das *Bild* von f als

$$\text{Bild}(f) = f(V) = \{f(x) \mid x \in V\},$$

und den *Kern* von f durch

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

Bemerkung 3.10. Für eine lineare Abbildung $f : V \rightarrow W$ ist $\text{Bild}(f)$ ein Unterraum von W und $\text{Kern}(f)$ ein Unterraum von V .

Beweis. $\text{Bild}(f)$ ist ein Unterraum:

Es gilt $0 = f(0) \in \text{Bild}(f)$.

Seien $y_1 = f(x_1)$ und $y_2 = f(x_2)$ in $\text{Bild}(f)$, wobei $x_1, x_2 \in V$. Dann ist

$$y_1 + y_2 = f(x_1) + f(x_2) \stackrel{f \text{ linear}}{=} f(x_1 + x_2) \in \text{Bild}(f).$$

Sei $\lambda \in K$ und $y = f(x) \in \text{Bild}(f)$. Dann ist

$$\lambda y = \lambda f(x) \stackrel{f \text{ linear}}{=} f(\lambda x) \in \text{Bild}(f).$$

Wir haben gezeigt, dass $\text{Bild}(f)$ die 0 enthält und abgeschlossen unter Addition und Skalarmultiplikation ist. Damit ist $\text{Bild}(f)$ ein Untervektorraum. Die Behauptung für $\text{Kern}(f)$ wird in der Übung gezeigt. $\square$

Bemerkung 3.11. Sei $f : V \rightarrow W$ eine lineare Abbildung. Es gilt:

$$f \text{ ist injektiv} \Leftrightarrow \text{Kern}(f) = \{0\}.$$

Beweis. „ $\Rightarrow$ “: Klar.

„ $\Leftarrow$ “: Sei $\text{Kern}(f) = \{0\}$. Seien $x_1, x_2 \in V$ mit $f(x_1) = f(x_2)$. Wir müssen zeigen, dass $x_1 = x_2$ gilt. Es gilt

$$0 = f(x_1) - f(x_2) \stackrel{f \text{ linear}}{=} f(x_1 - x_2).$$

Daraus folgt

$$x_1 - x_2 \in \text{Kern}(f) = \{0\},$$

Es gilt daher

$$x_1 - x_2 = 0,$$

was gleichbedeutend mit

$$x_1 = x_2$$

ist. $\square$

Satz 3.12 (Dimensionsformel). *Seien V und W endlichdimensionale K -Vektorräume und $f : V \rightarrow W$ eine lineare Abbildung. Es gilt*

$$\dim(V) = \dim(\text{Kern}(f)) + \dim(\text{Bild}(f)).$$

Beweis. Sei $m = \dim(\text{Kern}(f))$ und $e_1, \dots, e_m$ eine Basis von $\text{Kern}(f)$. Wir ergänzen diese zu einer Basis $e_1, \dots, e_m, e_{m+1}, \dots, e_n$ von V , wobei $n = \dim(V) \geq m$.

Es gilt

$$\text{Lin}(f(e_1), \dots, f(e_n)) = \text{Bild}(f).$$

Wegen $f(e_1) = \dots = f(e_m) = 0$ gilt sogar

$$\text{Lin}(f(e_{m+1}), \dots, f(e_n)) = \text{Bild}(f).$$

Behauptung: Die Vektoren $f(e_{m+1}), \dots, f(e_n)$ sind linear unabhängig.

Sei dazu

$$\lambda_{m+1}f(e_{m+1}) + \dots + \lambda_n f(e_n) = 0,$$

wobei $\lambda_{m+1}, \dots, \lambda_n \in K$. Da f linear ist, folgt

$$f(\lambda_{m+1}e_{m+1} + \dots + \lambda_n e_n) = 0.$$

Dies bedeutet, dass

$$v = \lambda_{m+1}e_{m+1} + \dots + \lambda_n e_n \in \text{Kern}(f).$$

Da $e_1, \dots, e_m$ eine Basis von $\text{Kern}(f)$ ist, gibt es $\lambda_1, \dots, \lambda_m \in K$ mit

$$-v = \lambda_1 e_1 + \dots + \lambda_m e_m.$$

Es folgt

$$\lambda_1 e_1 + \dots + \lambda_n e_n = 0.$$

Da die Vektoren $e_1, \dots, e_n$ linear unabhängig sind, gilt

$$\lambda_1 = \dots = \lambda_n = 0.$$

Die Vektoren $f(e_{m+1}), \dots, f(e_n)$ sind daher linear unabhängig. Da sie $\text{Bild}(f)$ erzeugen, bilden sie eine Basis von $\text{Bild}(f)$. Es gilt daher

$$\dim(\text{Bild}(f)) = n - m,$$

woraus

$$\dim(\text{Kern}(f)) + \dim(\text{Bild}(f)) = m + (n - m) = n = \dim(V)$$

folgt. □

Korollar 3.13. *Seien V und W K -Vektorräume der gleichen Dimension n . Sei $f : V \rightarrow W$ eine lineare Abbildung. Dann sind äquivalent:*

i) f ist injektiv.

ii) f ist surjektiv.

iii) f ist bijektiv.

Beweis. Es genügt die Äquivalenz von i) und ii) zu zeigen. Da V und W die gleiche Dimension haben gilt nach Satz 3.12, dass

$$\dim(W) = \dim(\text{Kern}(f)) + \dim(\text{Bild}(f)).$$

Es ist nun

$$\begin{aligned} f \text{ ist injektiv} &\Leftrightarrow \dim(\text{Kern}(f)) = 0 \\ &\Leftrightarrow \dim(\text{Bild}(f)) = \dim(W) \\ &\Leftrightarrow f \text{ ist surjektiv.} \end{aligned}$$

□

Man vergleiche dies mit der analogen Aussage für beliebige Abbildungen $f : M \rightarrow M'$ zwischen endlichen Mengen mit $\#M = \#M' = n$.

3.2.1 Übungen

Aufgabe 3.1. Welche der folgenden Abbildungen von $\mathbb{R}^2$ nach $\mathbb{R}$ bzw. $\mathbb{R}^2$ sind lineare Abbildungen? Begründen Sie kurz Ihre Antworten.

$$\phi_1(x_1, x_2) = x_1 + x_2$$

$$\phi_2(x_1, x_2) = x_1 \cdot x_2$$

$$\phi_3(x_1, x_2) = -x_2$$

$$\phi_4(x_1, x_2) = (x_1 - 2, 3x_2)$$

$$\phi_5(x_1, x_2) = (2x_1 + x_2, x_2 - x_1)$$

$$\phi_6(x_1, x_2) = (2x_1 - 3x_2, 0)$$

Aufgabe 3.2. Es sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K^{2 \times 2}$ eine fest gewählte Matrix. Wir betrachten die Abbildung

$$\varphi : K^{2 \times 2} \rightarrow K^{2 \times 2}, \quad \varphi(B) = AB.$$

1. Zeigen Sie, dass φ linear ist.
2. Bestimmen Sie $\text{Kern}(\varphi) = \{B \in K^{2 \times 2} : AB = 0\}$ in Abhängigkeit von a, b, c, d .
3. Zeigen Sie, dass φ genau dann bijektiv ist, wenn $ad - bc \neq 0$.

Aufgabe 3.3. Es sei K ein endlicher Körper mit q Elementen. Bestimmen Sie jeweils die Anzahl der Elemente in den folgenden Mengen:

1. K^3 für $q = 5$.
2. Die Menge der K -linearen Abbildungen von K^2 nach K für $q = 17$.
3. Die Menge der 1-dimensionalen Untervektorräume von K^4 für $q = 3$.

3.3 Isomorphismen

Definition 3.14. Eine lineare Abbildung $f : V \rightarrow W$ heißt *Isomorphismus*, falls sie bijektiv ist.

Bemerkung 3.15. Für einen Isomorphismus $f : V \rightarrow W$ ist auch die Umkehrabbildung $f^{-1} : W \rightarrow V$ linear und damit ein Isomorphismus.

Beweis. Übung. □

Bemerkung 3.16. Sei $f : V \rightarrow W$ ein Isomorphismus und $v_1, \dots, v_n \in V$. Es gilt:

- i) $v_1, \dots, v_n$ sind genau dann linear unabhängig, wenn $f(v_1), \dots, f(v_n)$ linear unabhängig sind.
- ii) $v_1, \dots, v_n$ erzeugen V genau dann, wenn W von $f(v_1), \dots, f(v_n)$ erzeugt wird.
- iii) $v_1, \dots, v_n$ sind genau dann eine Basis von V , wenn $f(v_1), \dots, f(v_n)$ eine Basis von W sind.

Beweis. Übung. □

Bemerkung 3.17. Zwei K -Vektorräume V und W heißen *isomorph*, falls es einen Isomorphismus $f : V \rightarrow W$ gibt. Schreibe $V \simeq W$. Durch „ $\simeq$ “ wird eine Äquivalenzrelation auf der Menge der K -Vektorräume definiert.

Satz 3.18. Sei V ein K -Vektorraum und $B = (b_1, \dots, b_n)$ ein n -Tupel von Vektoren in V . Die Abbildung

$$\sigma_B : K^n \rightarrow V, \quad x \mapsto \sigma_B(x) = \sum_{i=1}^n x_i \cdot b_i,$$

ist linear. Sie ist

- i) injektiv genau dann, wenn $b_1, \dots, b_n$ linear unabhängig sind.
- ii) surjektiv genau dann, wenn $b_1, \dots, b_n$ ein Erzeugendensystem bilden.
- iii) bijektiv genau dann, wenn $b_1, \dots, b_n$ eine Basis von V bilden.

Beweis. Klar. □

Bemerkung 3.19. Ist $B = (b_1, \dots, b_n)$ eine Basis von V , so können wir in Satz 3.18 (iii) zu $v \in V$ das Urbild $\sigma_B^{-1}(v) \in K^n$ betrachten. Man nennt es den *Koordinatenvektor* von v bezüglich der Basis B .

Bemerkung 3.20. Satz 3.18 (iii) besagt, dass es das Gleiche ist, eine Basis eines n -dimensionalen Vektorraums V anzugeben, wie einen Isomorphismus $K^n \rightarrow V$ anzugeben.

Satz 3.21. Sei V ein K -Vektorraum der Dimension $m < \infty$. Es ist

$$V \simeq K^n \quad \Leftrightarrow \quad n = m.$$

Beweis. „ $\Leftarrow$ “: Sei $B = (b_1, \dots, b_n)$ eine Basis von V und $\sigma_B : K^n \rightarrow V$ wie in Satz 3.18. Es folgt mit Satz 3.18, dass σ_B ein Isomorphismus ist.

„ $\Rightarrow$ “: Sei $\dim(V) = m$ und $g : K^n \rightarrow V$ ein Isomorphismus. Es gilt dann $\dim(\text{Bild}(g)) = \dim(V) = m$ und $\dim(\text{Kern}(g)) = \dim(\{0\}) = 0$. Aus der Dimensionsformel folgt $m = n$. $\square$

Folgerung 3.22. Die Dimension legt einen endlichdimensionalen K -Vektorraum bis auf Isomorphie fest.

3.4 Lineare Abbildungen und Matrizen

Seien V, W K -Vektorräume und $f : V \rightarrow W$ eine lineare Abbildung.

Ziel. Beschreibe f durch eine Matrix.

Dazu müssen wir Basen

$$\begin{aligned} B &= (b_1, \dots, b_n) \quad \text{von } V, \\ C &= (c_1, \dots, c_m) \quad \text{von } W \end{aligned}$$

wählen. Für $j = 1, \dots, n$ schreibe

$$f(b_j) = \sum_{i=1}^m a_{ij} c_i$$

mit $a_{ij} \in K$. Dies ist nach Bemerkung 3.7 in eindeutiger Weise möglich. Dann ist

$$A = (a_{ij}) \in K^{m \times n}$$

eine $m \times n$ -Matrix. Die Abbildung f ist durch A eindeutig bestimmt.

Definition 3.23. Die Matrix A heißt *Koordinatenmatrix* (auch *darstellende Matrix*) der linearen Abbildung f bezüglich der Basen B und C . Schreibe

$$A = \mathcal{M}_C^B(f).$$

Wichtig. Die Matrix A hängt von der Wahl der Basen B und C ab (Wahl der Koordinatensysteme).

Bemerkung 3.24.

- $\text{Hom}(V, W) = \{K\text{-lineare Abbildungen von } V \text{ nach } W\}$ ist ein K -Vektorraum.
- $K^{m \times n}$ ist ein K -Vektorraum der Dimension mn .

Bemerkung 3.25. Die Abbildung

$$\text{Hom}(V, W) \rightarrow K^{m \times n}, \quad f \mapsto A = \mathcal{M}_C^B(f)$$

ist bijektiv (und K -linear, also ein K -Vektorraumisomorphismus).

Beweis. Die Abbildung f ist durch Angabe von $f(b_j)$ (also der a_{ij} mit $i = 1, \dots, m$) für $j = 1, \dots, n$ eindeutig bestimmt. Daraus folgt, dass die Abbildung $f \mapsto \mathcal{M}_C^B(f)$ injektiv ist.

Zu beliebig vorgegebenen Vektoren $w_j = a_{1j}c_1 + \dots + a_{mj}c_m \in W$ existiert nach Satz 3.8 eine lineare Abbildung mit $f(b_j) = w_j$. Es folgt, dass die Abbildung $f \mapsto \mathcal{M}_C^B(f)$ surjektiv ist. $\square$

Beispiel 3.26. Sei $V = K^n$, $W = K^m$ und $A \in K^{m \times n}$. Sei

$$f_A : K^n \rightarrow K^m, \quad x \mapsto f_A(x) = Ax,$$

die A zugeordnete lineare Abbildung. Sei

$$\begin{aligned} B &= (e_1, \dots, e_n) && \text{die Standardbasis von } K^n, \\ C &= (e'_1, \dots, e'_m) && \text{die Standardbasis von } K^m. \end{aligned}$$

Was ist $\mathcal{M}_C^B(f_A)$?

Was ist $f_A(e_j)$ für $j = 1, \dots, n$? Es gilt

$$\begin{aligned} f_A(e_j) = Ae_j &= \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j\text{-te Stelle} \\ &= \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = a_{1j}e'_1 + \dots + a_{mj}e'_m, \end{aligned}$$

woraus

$$\mathcal{M}_C^B(f_A) = A$$

folgt.

Bemerkung 3.27. Die f_A bezüglich der Standardbasen zugeordnete Koordinatenmatrix (darstellende Matrix) ist A selbst.

Dies lässt sich leicht verallgemeinern:

Bemerkung 3.28. Seien V und W K -Vektorräume und $f : V \rightarrow W$ eine lineare Abbildung. Seien $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_m)$ Basen von V , beziehungsweise W . Sei σ_B der Isomorphismus

$$K^n \rightarrow V, \quad x \mapsto \sigma_B(x) = \sum_{i=1}^n x_i \cdot b_i$$

und σ_C der Isomorphismus

$$K^m \rightarrow W, \quad y \mapsto \sigma_C(y) = \sum_{j=1}^m y_j \cdot c_j$$

wie in Satz 3.18. Sei $A = \mathcal{M}_C^B(f)$. Dann kommutiert das folgende Diagramm:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \sigma_B \uparrow & & \uparrow \sigma_C \\ K^n & \xrightarrow{f_A} & K^m \end{array}$$

Das heißt $f \circ \sigma_B = \sigma_C \circ f_A$.

Korollar 3.29. In der j -ten Spalte von A steht der Koordinatenvektor $\sigma_C^{-1}(f(b_j))$ des Bildes des j -ten Basisvektors b_j .

Wir zeigen nun, dass die Verkettung linearer Abbildungen dem Produkt von Matrizen entspricht.

Satz 3.30. Seien V, W und X K -Vektorräume mit Basen $B = (b_1, \dots, b_n)$, $C = (c_1, \dots, c_m)$ und $D = (d_1, \dots, d_l)$. Seien $f : V \rightarrow W$ und $g : W \rightarrow X$ lineare Abbildungen. Seien A, A' und A'' die Koordinatenmatrizen von f, g und $g \circ f$ bezüglich der gewählten Basen. Dann gilt

$$A'' = A' A.$$

Beweis. Wir wissen, dass das folgende Diagramm kommutiert:

$$\begin{array}{ccccc}
 & & g \circ f & & \\
 & \nearrow f & & \searrow g & \\
 V & \xrightarrow{\quad} & W & \xrightarrow{\quad} & X \\
 \uparrow \sigma_B & & \uparrow \sigma_C & & \uparrow \sigma_D \\
 K^n & \xrightarrow{x \mapsto Ax} & K^m & \xrightarrow{x \mapsto A'x} & K^l \\
 & \searrow x \mapsto A''x & & \nearrow &
 \end{array}$$

Hieraus folgt die Behauptung.

Alternativer Beweis:

Wir zeigen, dass die Spalten von A'' und $A'A$ übereinstimmen. Sei $A = (a_{ij})$, $A' = (a'_{ij})$ und $A'' = (a''_{ij})$. Die j -te Spalte von A'' ist durch

$$(g \circ f)(b_j) = \sum_{i=1}^l a''_{ij} d_i$$

gegeben. Andererseits gilt

$$\begin{aligned}
 (g \circ f)(b_j) &= g(f(b_j)) \\
 &= g\left(\sum_{k=1}^m a_{kj} c_k\right) \\
 &= \sum_{k=1}^m a_{kj} g(c_k) \\
 &= \sum_{k=1}^m a_{kj} \sum_{i=1}^l a'_{ik} d_i \\
 &= \sum_{i=1}^l \left(\sum_{k=1}^m a'_{ik} a_{kj}\right) d_i.
 \end{aligned}$$

Durch Koeffizientenvergleich folgt die Behauptung. $\square$

Definition 3.31. Sei $f : V \rightarrow W$ eine lineare Abbildung. Der Rang von f ist definiert als

$$\text{Rang}(f) = \dim(\text{Bild}(f)).$$

Bemerkung 3.32. Seien B und C Basen von V und W . Sei A die zugehörige Koordinatenmatrix von f . Es gilt

$$\text{Rang}(f) = \text{Rang}(A).$$

Beweis. Wegen Bemerkung 3.28 gilt

$$\begin{aligned} \text{Rang}(f) &= \text{Rang}(f_A) \\ &= \dim\{Ax \mid x \in K^n\} \\ &= \text{Dimension des von den Spalten von } A \text{ aufgespannten Teilraums von } K^n \\ &= \text{Rang}(A). \end{aligned}$$

□

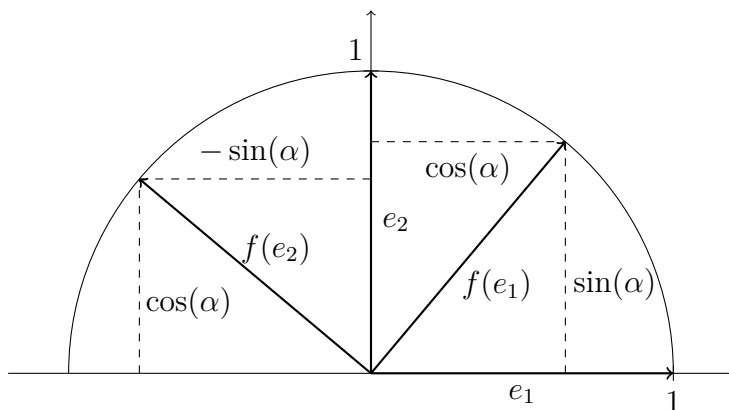
Beispiel 3.33 (Drehungen im $\mathbb{R}^2$). Sei $B = (e_1, e_2) = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die Drehung um $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ mit Winkel α .

Es gilt

$$f(e_1) = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}, \quad f(e_2) = \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix}$$

und

$$\mathcal{M}_B^B(f) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$



3.4.1 Übungen

Aufgabe 3.4. Es seien A und B reelle $n \times n$ -Matrizen. Zeigen Sie, dass

$$\text{Rang}(AB) \leq \text{Rang}(B)$$

gilt.

Hinweis: Betrachten Sie die zugehörigen linearen Abbildungen $f_A(x) = Ax$, $f_B(x) = Bx$ und $f_{AB}(x) = ABx$.

Aufgabe 3.5.

1. Wir betrachten die lineare Abbildung

$$\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}.$$

Bestimmen Sie eine Basis B von $\mathbb{R}^2$ und eine Basis C von $\mathbb{R}^2$ so dass die Darstellungsmatrix $\mathcal{M}_C^B(\varphi)$ die Einheitsmatrix ist.

2. Sei ψ die orthogonale Spiegelung in $\mathbb{R}^2$ entlang einer Ursprungsgeraden. Bestimmen Sie die Darstellungsmatrix von ψ bzgl. einer ‘geeigneten’ Basis.

Aufgabe 3.6. Es sei $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ein Endomorphismus mit $\varphi^3 = 0$ und $\varphi^2 \neq 0$. Zeigen Sie, dass es eine Basis B von $\mathbb{R}^3$ gibt, so dass die Darstellungsmatrix von φ bzgl. B die Form

$$\mathcal{M}_B^B(\varphi) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

hat. Gehen Sie dazu wie folgt vor.

1. Zeigen Sie, dass es einen Vektor $v \in \mathbb{R}^3$ gibt, so dass $v, \varphi(v)$ und $\varphi^2(v)$ ungleich Null sind.
2. Zeigen Sie, dass für dieses v die Vektoren $v, \varphi(v), \varphi^2(v)$ linear unabhängig sind und eine Basis von $\mathbb{R}^3$ bilden.
3. Stellen Sie die Abbildung φ bezüglich einer Basis B dar, welche durch Vertauschen der drei Vektoren entsteht.

3.5 Isomorphismen von Vektorräumen und Basiswechsel

Definition 3.34. Eine $m \times n$ Matrix $A \in K^{m \times n}$ heißt *invertierbar*, falls $m = n$ und falls es ein $A' \in K^{n \times n}$ gibt, so dass

$$A' \cdot A = 1_n.$$

Notation. Schreibe A^{-1} für die inverse Matrix A' .

Bemerkung 3.35. Die Menge

$$\mathrm{GL}_n(K) = \{A \in K^{n \times n} \mid A \text{ ist invertierbar}\}$$

bildet mit der Matrixmultiplikation eine Gruppe. Sie heißt *lineare Gruppe* (n -ten Grades) über K .

Beweis. Übung. □

Bemerkung 3.36. Sei V ein K -Vektorraum der Dimension n , W ein K -Vektorraum der Dimension m und $f : V \rightarrow W$ eine lineare Abbildung. Sei A die Koordinatenmatrix von f bezüglich zweier Basen B und C von V beziehungsweise W . Es sind äquivalent:

- i) f ist ein Isomorphismus,
- ii) Es ist $m = n$ und $\mathrm{Rang}(f) = n$,

iii) A ist invertierbar.

In diesem Fall ist die Koordinatenmatrix von f^{-1} bezüglich C und B gleich A^{-1} .

Beweis. (i) $\Leftrightarrow$ (ii): Satz 3.18 und Satz 3.21.

(i) $\Rightarrow$ (iii): Sei f ein Isomorphismus und A' die Koordinatenmatrix von f^{-1} bezüglich C und B . Es gilt

$$f^{-1} \circ f = \text{id}.$$

Mit Satz 3.30 folgt

$$1_n = \mathcal{M}_B^B(\text{id}) = \mathcal{M}_B^B(f^{-1} \circ f) = \mathcal{M}_B^C(f^{-1}) \cdot \mathcal{M}_C^B(f),$$

das heißt

$$1_n = A' \cdot A.$$

A ist somit invertierbar und $A^{-1} = \mathcal{M}_B^C(f^{-1})$.

(iii) $\Rightarrow$ (i): Sei A invertierbar, $A^{-1} \cdot A = 1_n$. Sei

$$\begin{aligned} f_{A^{-1}} : K^n &\rightarrow K^n \\ f_{A^{-1}}(x) &= A^{-1} \cdot x \end{aligned}$$

die zur inversen Matrix gehörende lineare Abbildung. Es ist offenbar f_A ein Isomorphismus mit $(f_A)^{-1} = f_{A^{-1}}$. Mit Bemerkung 3.28 folgt, dass f ein Isomorphismus ist. $\square$

Definition 3.37. Sei V ein K -Vektorraum und $B = (b_1, \dots, b_n)$ und $B' = (b'_1, \dots, b'_n)$ zwei Basen von V . Schreibe

$$b_j = \sum_{i=1}^n s_{ij} b'_i$$

für $j = 1, \dots, n$. Die Matrix $S = (s_{ij}) \in K^{n \times n}$ heißt *Basiswechselmatrix* von B nach B' .

Bemerkung 3.38.

i) Es gilt $S = \mathcal{M}_{B'}^B(\text{id}_V)$.

ii) Sei $f : V \rightarrow V$ die lineare Abbildung mit $f(b'_j) = b_j$. Dann gilt $S = \mathcal{M}_{B'}^{B'}(f)$.

iii) S ist invertierbar und $S^{-1} = \mathcal{M}_B^{B'}(\text{id}_V)$.

Beweis. (i) folgt direkt aus der Definition.

(ii): Übung.

(iii) folgt aus Bemerkung 3.36. $\square$

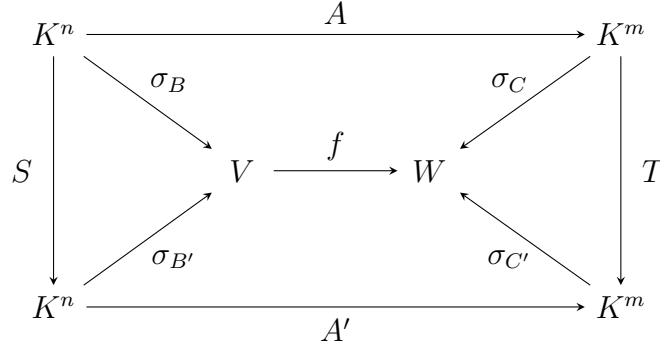
Satz 3.39 (Transformationsformel). Seien V und W K -Vektorräume und $f : V \rightarrow W$ eine lineare Abbildung. Seien B und B' Basen von V , seien C und C' Basen von W . Sei

$$\begin{aligned} A &= \mathcal{M}_C^B(f), & A' &= \mathcal{M}_{C'}^{B'}(f), \\ S &= \mathcal{M}_{B'}^B(\text{id}_V), & T &= \mathcal{M}_{C'}^C(\text{id}_W). \end{aligned}$$

Es gilt

$$A' = T \cdot A \cdot S^{-1}.$$

Beweis. Dies folgt aus Satz 3.30 mit dem kommutativen Diagramm:



□

Korollar 3.40. Sei V ein K -Vektorraum und $f : V \rightarrow V$ eine lineare Abbildung. Seien B und B' Basen von V und $A = \mathcal{M}_B^B(f)$, $A' = \mathcal{M}_{B'}^{B'}(f)$. Sei $S = \mathcal{M}_{B'}^B(\text{id}_V)$ die Basiswechselmatrix von B nach B' . Es gilt

$$A' = SAS^{-1}.$$

Definition 3.41.

- i) Zwei Matrizen $A, A' \in K^{m \times n}$ heißen *äquivalent*, falls es Matrizen $S \in \text{GL}_n(K)$ und $T \in \text{GL}_m(K)$ gibt, so dass

$$A' = TAS^{-1}.$$

- ii) Zwei Matrizen $A, A' \in K^{n \times n}$ heißen *ähnlich*, falls es eine Matrix $S \in \text{GL}_n(K)$ gibt, so dass

$$A' = SAS^{-1}.$$

Beispiel 3.42.

- i) Sei $V = W = \mathbb{R}^2$. Sei $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gegeben durch

$$\begin{pmatrix} a \\ b \end{pmatrix} \mapsto \frac{1}{5} \begin{pmatrix} -3a + 4b \\ 4a + 3b \end{pmatrix} = \frac{1}{5} \begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix},$$

das heißt $\varphi = f_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit $A = \frac{1}{5} \begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}$. Sei $B = (e_1, e_2)$ mit $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Sei $B' = (v_1, v_2)$ mit $v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und $v_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$. Dann sind B und B' Basen von V (und von W). Die Basiswechselmatrizen sind durch

$$S^{-1} = T^{-1} = \mathcal{M}_{B'}^B(\text{id}_{\mathbb{R}^2}) = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

und

$$T = S = \mathcal{M}_{B'}^B(\text{id}_{\mathbb{R}^2}) = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$$

gegeben. Mit der Transformationsformel folgt

$$\begin{aligned}
\mathcal{M}_{B'}^{B'}(\varphi) &= \mathcal{M}_{B'}^B(\text{id}_W) \mathcal{M}_B^B(\varphi) \mathcal{M}_B^{B'}(\text{id}_V) \\
&= \mathcal{M}_{B'}^B(\text{id}_{\mathbb{R}^2}) \mathcal{M}_B^B(f_A) \mathcal{M}_B^{B'}(\text{id}_{\mathbb{R}^2}) \\
&= T A S^{-1} \\
&= \frac{1}{25} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \\
&= \frac{1}{25} \begin{pmatrix} 5 & 10 \\ 10 & -5 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \\
&= \frac{1}{25} \begin{pmatrix} 25 & 0 \\ 0 & -25 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\end{aligned}$$

ii) Sei $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ die lineare Fortsetzung von

$$\psi \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \quad \psi \left(\begin{pmatrix} 1 \\ 2 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Diese ist nach Satz 3.8 eindeutig, da $B = ((\frac{1}{1}), (\frac{1}{2}))$ eine Basis von $\mathbb{R}^2$ ist. Für $\mathbb{R}^3$ wählen wir die Basis $C = ((\frac{1}{0}), (\frac{0}{1}), (\frac{0}{1}))$. Die Koordinatenmatrix von ψ ist dann durch

$$A = \mathcal{M}_C^B(\psi) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -2 & 1 \end{pmatrix}$$

gegeben.

Gesucht ist die Matrix $\mathcal{M}_{C'}^{B'}(\psi)$ für die Basen

$$B' = ((\frac{1}{0}), (\frac{0}{1})), \quad C' = ((\frac{1}{0}), (\frac{0}{1}), (\frac{0}{1})).$$

Dazu berechnen wir

$$S = \mathcal{M}_{B'}^B(\text{id}_{\mathbb{R}^2}) = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

und

$$T = \mathcal{M}_{C'}^C(\text{id}_{\mathbb{R}^3}) = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Mit der Transformationsformel folgt

$$\begin{aligned}
 \mathcal{M}_{C'}^{B'}(\psi) &= T A S^{-1} \\
 &= \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} 2 & -1 \\ 2 & -1 \\ -5 & 3 \end{pmatrix}.
 \end{aligned}$$

3.6 Die allgemeine lineare Gruppe

Sei K ein Körper. Die Menge

$$\mathrm{GL}_n(K) = \{A \in K^{n \times n} \mid A \text{ ist invertierbar}\}$$

ist eine Gruppe mit neutralem Element

$$1_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}.$$

Satz 3.43. *Ein Matrix $A \in K^{n \times n}$ ist genau dann invertierbar, wenn $\mathrm{Rang}(A) = n$.*

Beweis.

$$\begin{aligned}
 &A \text{ ist invertierbar} \\
 \Leftrightarrow &f_A : K^n \rightarrow K^n, \quad x \mapsto Ax \text{ ist ein Isomorphismus} \\
 \Leftrightarrow &f_A \text{ ist surjektiv} \\
 \Leftrightarrow &\mathrm{Bild}(f_A) = K^n \\
 \Leftrightarrow &\mathrm{Rang}(A) = n.
 \end{aligned}$$

Damit ist die Behauptung bewiesen. □

Elementarmatrizen

Permutationsmatrizen:

Für $\sigma \in S_n$ sei

$$P_\sigma = (e_{\sigma(1)}, \dots, e_{\sigma(n)}) \in \mathrm{GL}_n(K),$$

wobei $(e_1, \dots, e_n)$ die Standardbasis von K^n ist. Dann gilt

$$f_{P_\sigma}(e_j) = e_{\sigma(j)}$$

und

$$(f_{P_{\sigma'}} \circ f_{P_\sigma})(e_j) = f_{P_{\sigma'}}(e_{\sigma(j)}) = e_{\sigma'(\sigma(j))} = e_{\sigma' \circ \sigma(j)}.$$

Es folgt, dass

$$S_n \rightarrow \mathrm{GL}_n(K), \quad \sigma \mapsto P_\sigma$$

ein Gruppenhomomorphismus ist.

Beispiel 3.44. Es gilt

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = P_{(12)}.$$

Scherungsmatrizen:

Für $i \neq j$ sei

$$Q_{ij} = 1_n + E_{ij} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \ddots \\ & & & & 1 \end{pmatrix}. \quad \begin{array}{l} \leftarrow i\text{-te Zeile} \\ \uparrow \\ j\text{-te Spalte} \end{array}$$

Für $\lambda \in K$ sei

$$Q_{ij}(\lambda) = 1_n + \lambda E_{ij}.$$

Spezielle Diagonalmatrizen:

Für $\lambda \in K^\times$ und $1 \leq i \leq n$, sei

$$S_i(\lambda) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \lambda & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} \quad \leftarrow i\text{-te Zeile}$$

Bemerkung 3.45. Jede Elementarmatrix ist invertierbar. Es gilt

$$\begin{aligned} P_\sigma^{-1} &= P_{\sigma^{-1}}, \\ S_i(\lambda)^{-1} &= S_i(\lambda^{-1}), \\ Q_{ij}(\lambda)^{-1} &= Q_{ij}(-\lambda). \end{aligned}$$

Zusätzlich gilt

$$Q_{ij}(\lambda) = S_j(\lambda^{-1})Q_{ij}S_j(\lambda).$$

Beweis. Nachrechnen. □

Bemerkung 3.46. Das Multiplizieren einer Elementarmatrix aus $\text{GL}_m(K)$ von links an eine Matrix $A \in K^{m \times n}$ bewirkt eine entsprechende elementare Zeilenumformung bei A :

$S_i(\lambda) \cdot A$: Multiplizieren der i -ten Zeile mit λ .

$Q_{ij} \cdot A$: Addieren der j -ten Zeile zur i -ten Zeile.

$P_{(ij)} \cdot A$: Vertauschen der i -ten und der j -ten Zeile.

Das Multiplizieren einer Elementarmatrix aus $\text{GL}_n(K)$ von rechts bewirkt eine entsprechende elementare Spaltenumformung.

Satz 3.47. Jede Matrix aus $\text{GL}_n(K)$ lässt sich als Produkt endlich vieler Elementarmatrizen schreiben.

Beweis. Nach Lemma 1.48 lässt sich jede Matrix $A \in K^{n \times n}$ durch endlich viele elementare Zeilen- und Spaltenumformungen in die Form

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

bringen. Laut Bemerkung 3.46 existieren dann Elementarmatrizen $X_1, \dots, X_k, Y_1, \dots, Y_l$, so dass

$$X_1 \cdot \dots \cdot X_k \cdot A \cdot Y_1 \cdot \dots \cdot Y_l = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Falls $A \in \text{GL}_n(K)$, so ist $r = n$ und es folgt

$$A = X_k^{-1} \cdot \dots \cdot X_1^{-1} \cdot Y_l^{-1} \cdot \dots \cdot Y_1^{-1}.$$

□

Satz 3.48. Jede Matrix $A \in \text{GL}_n(K)$ lässt sich durch ausschließlich elementare Zeilenumformungen in die Einheitsmatrix überführen.

Bemerkung 3.49. Ist $A \in \text{GL}_n(K)$ und sind $X_1, \dots, X_k$ Elementarmatrizen mit

$$X_1 \cdot \dots \cdot X_k \cdot A = 1_n,$$

so gilt

$$A^{-1} = X_1 \cdot \dots \cdot X_k.$$

Dies liefert ein Verfahren zur Berechnung von A^{-1} .

3.6.1 Übungen

Aufgabe 3.7. Wir leiten einen Algorithmus zum Invertieren von Matrizen her.

Sei K ein Körper und $A \in K^{n \times n}$. Zeigen Sie, dass A invertierbar ist genau dann, wenn sich $(A \mid E_n) \in K^{n \times 2n}$ durch elementare Zeilenumformungen in $(E_n \mid B)$ überführen lässt. In diesem Fall ist $B = A^{-1}$.

Aufgabe 3.8. Berechnen Sie die inverse Matrix von $A := \begin{pmatrix} 1 & 1 & 2 \\ 2 & 3 & 3 \\ 2 & 6 & -1 \end{pmatrix} \in \mathbb{R}^{3 \times 3}$.

Aufgabe 3.9. Zeigen Sie, dass die Menge

$$\mathrm{GL}_n(K) = \{A \in K^{n \times n} : A \text{ invertierbar}\}$$

mit der Matrixmultiplikation eine Gruppe ist.

4 Die Determinante

4.1 Die Leibniz-Formel

Ziel. Sei $A \in K^{n \times n}$. Wir suchen ein Kriterium, um zu entscheiden ob A invertierbar ist (das heißt $A \in \text{GL}_n(K)$). Wenn A invertierbar ist, was ist A^{-1} ?

Die Determinante liefert eine Abbildung $K^{n \times n} \rightarrow K$, mit der man dies entscheiden kann.

Beispiel 4.1.

$$1) \ n = 1: K^{1 \times 1} \simeq K, \quad \det((a)) = a.$$

$$2) \ n = 2: \text{Für } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K^{2 \times 2} \text{ definiert man}$$

$$\det(A) = ad - bc \in K.$$

Bemerkung 4.2. Eine Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K^{2 \times 2}$ ist genau dann invertierbar, wenn $\det(A) \neq 0$. In diesem Fall ist

$$A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Beweis. „ $\Leftarrow$ “: Sei $\det(A) \neq 0$. Dann gilt

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} = \det(A) \cdot 1_2.$$

Es folgen die Invertierbarkeit und die Formel.

„ $\Rightarrow$ “: Sei A invertierbar. Es ist zu zeigen, dass $\det(A) \neq 0$. Dies folgt aus dem nächsten Lemma. $\square$

Lemma 4.3. Für $A, B \in K^{2 \times 2}$ gilt $\det(A \cdot B) = \det(A) \cdot \det(B)$. Insbesondere definiert $\det$ einen Gruppenhomomorphismus

$$\det : \text{GL}_2(K) \rightarrow K^\times.$$

Beweis. Nachrechnen. $\square$

Bemerkung 4.4 (Geometrische Bedeutung der Determinante von reellen Matrizen). Für $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ ist $\det(A)$ ein Maß für den Flächeninhalt des von den Spaltenvektoren aufgespannten Parallelogramms in $\mathbb{R}^2$. Eine analoge Interpretation gibt es auch für $n > 2$.

Beispiel 4.5 (Fortsetzung von Beispiel 4.1).

3) $n = 3$: Wir schreiben

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in K^{3 \times 3}.$$

In diesem Fall leistet die Definition

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

das Gewünschte.

4) $n \geq 4$: Was ist die Struktur der Formel für $n = 1, 2, 3$?

Die Summanden sind von der Form

$$a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)}$$

mit $\sigma \in S_n$. Das Vorzeichen ist $\text{sgn}(\sigma)$. Dies führt zu:

Definition 4.6 (Leibniz-Formel). Die Determinante einer $n \times n$ -Matrix $A = (a_{ij})$ ist definiert durch

$$\det(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdot \dots \cdot a_{n,\sigma(n)}.$$

Bemerkung 4.7.

- Für $n = 1, 2, 3$ ist dies unsere vorherige Definition.
- Die Summe hat $n!$ Summanden.

$$\begin{aligned} 10! &= 3628800, \\ 20! &\approx 2,4 \cdot 10^{18}, \\ 30! &\approx 2,7 \cdot 10^{32}. \end{aligned}$$

- Falls n groß ist, so ist die Berechnung praktisch sehr aufwändig.

$\rightsquigarrow$ Um $\det(A)$ zu verstehen, muss man die Struktur genauer untersuchen.

Wir bezeichnen nun die Spalten von $A = (a_{ij}) \in K^{n \times n}$ mit $a_1, \dots, a_n$, das heißt

$$a_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \quad a_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \quad \dots, \quad a_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix}.$$

Definition 4.8. Eine Funktion $f : K^{n \times n} \rightarrow K$ heißt linear in der k -ten Spalte, falls für beliebige feste Spalten $a_1, \dots, a_{k-1}, a_{k+1}, \dots, a_n$ die Funktion

$$g : K^n \rightarrow K, \quad g(x) = f(a_1, \dots, a_{k-1}, x, a_{k+1}, \dots, a_n)$$

linear ist. Das heißt $g(x + y) = g(x) + g(y)$ und $g(\lambda x) = \lambda g(x)$ für $x, y \in K^n$ und $\lambda \in K$.

Definition 4.9. Eine Funktion $f : K^{n \times n} \rightarrow K$ heißt alternierende Multilinearform (der Spalten), falls gilt:

- i) f ist linear in jeder Spalte.
- ii) f ist alternierend: Falls A zwei gleiche Spalten hat, so gilt $f(A) = 0$.

Bemerkung 4.10. Aus Bedingung ii) folgt wegen i) immer

- ii') f ist schiefsymmetrisch: Vertauscht man zwei Spalten, so ändert sich das Vorzeichen, das heißt

$$\begin{aligned} & f(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_{j-1}, a_j, a_{j+1}, \dots, a_n) \\ &= -f(a_1, \dots, a_{i-1}, a_j, a_{i+1}, \dots, a_{j-1}, a_i, a_{j+1}, \dots, a_n). \end{aligned}$$

Falls $\text{char}(K) \neq 2$ (das heißt $1 + 1 \neq 0 \in K$), so sind ii) und ii') äquivalent.

Beispiel 4.11. Für $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K^{2 \times 2}$ ist $A \mapsto \det(A)$ eine alternierende Multilinearform, denn:
Es gilt

$$\begin{aligned} \det \begin{pmatrix} a+a' & b \\ c+c' & d \end{pmatrix} &= (a+a')d - b(c+c') \\ &= ad - bc + a'd - bc' \\ &= \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \det \begin{pmatrix} a' & b \\ c' & d \end{pmatrix}, \end{aligned}$$

sowie

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

Es folgt, dass $\det$ linear in der ersten Spalte ist. Die zweite Spalte behandelt man analog. Weiterhin gilt

$$\det \begin{pmatrix} a & a \\ c & c \end{pmatrix} = ac - ac = 0,$$

die Abbildung ist somit alternierend.

Um dies für beliebiges n zu beweisen benötigen wir das folgende Lemma.

Lemma 4.12. Sei $A = (a_{ij}) \in K^{n \times n}$. Dann gilt

$$\det(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdot \dots \cdot a_{\sigma(n),n}.$$

Das heißt, dass $\det(A) = \det(A^t)$, wobei

$$A^t = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}$$

die zu A transponierte Matrix ist.

Beweis. Für $\tau, \sigma \in S_n$ gilt

$$a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdot \dots \cdot a_{n,\sigma(n)} = a_{\tau(1),\sigma(\tau(1))} \cdot a_{\tau(2),\sigma(\tau(2))} \cdot \dots \cdot a_{\tau(n),\sigma(\tau(n))},$$

da lediglich die Reihenfolge der Faktoren vertauscht wird. Es folgt

$$\begin{aligned} \det(A) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdot \dots \cdot a_{n,\sigma(n)} \\ &= \sum_{\sigma \in S_n} \underbrace{\operatorname{sgn}(\sigma)}_{=\operatorname{sgn}(\sigma^{-1})} \cdot a_{\sigma^{-1}(1),\sigma(\sigma^{-1}(1))} \cdot a_{\sigma^{-1}(2),\sigma(\sigma^{-1}(2))} \cdot \dots \cdot a_{\sigma^{-1}(n),\sigma(\sigma^{-1}(n))} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) \cdot a_{\sigma^{-1}(1),1} \cdot a_{\sigma^{-1}(2),2} \cdot \dots \cdot a_{\sigma^{-1}(n),n} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdot \dots \cdot a_{\sigma(n),n} \\ &= \det(A^t). \end{aligned}$$

□

Satz 4.13. Die Determinante $\det : K^{n \times n} \rightarrow K$ (definiert durch die Leibniz-Formel) ist eine alternierende Multilinearform.

Beweis.

- $\det$ ist linear in der k -ten Spalte:

Nach Lemma 4.12 gilt

$$\begin{aligned} &\det(a_1, \dots, a_{k-1}, a'_k + a''_k, a_{k+1}, \dots, a_n) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{\sigma(1),1} \cdot \dots \cdot a_{\sigma(k-1),k-1} \cdot (a'_{\sigma(k),k} + a''_{\sigma(k),k}) \cdot a_{\sigma(k+1),k+1} \cdot \dots \cdot a_{\sigma(n),n} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{\sigma(1),1} \cdot \dots \cdot a_{\sigma(k-1),k-1} \cdot a'_{\sigma(k),k} \cdot a_{\sigma(k+1),k+1} \cdot \dots \cdot a_{\sigma(n),n} \\ &\quad + \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{\sigma(1),1} \cdot \dots \cdot a_{\sigma(k-1),k-1} \cdot a''_{\sigma(k),k} \cdot a_{\sigma(k+1),k+1} \cdot \dots \cdot a_{\sigma(n),n} \\ &= \det(a_1, \dots, a_{k-1}, a'_k, a_{k+1}, \dots, a_n) + \det(a_1, \dots, a_{k-1}, a''_k, a_{k+1}, \dots, a_n). \end{aligned}$$

Weiterhin gilt

$$\begin{aligned} & \det(a_1, \dots, a_{k-1}, \lambda a_k, a_{k+1}, \dots, a_n) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{\sigma(1),1} \cdot \dots \cdot a_{\sigma(k-1),k-1} \cdot \lambda a_{\sigma(k),k} \cdot a_{\sigma(k+1),k+1} \cdot \dots \cdot a_{\sigma(n),n} \\ &= \lambda \det(a_1, \dots, a_{k-1}, a_k, a_{k+1}, \dots, a_n). \end{aligned}$$

- $\det$ ist alternierend:

Angenommen die i -te und j -te Spalte von A sind gleich mit $i \neq j$. Es ist zu zeigen, dass

$$\det(a_1, \dots, a_n) = 0.$$

Sei $\tau \in S_n$ die Transposition, die i und j vertauscht. Sei $A_n = \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\}$ die alternierende Gruppe. Dann gilt $S_n = A_n \cup \tau A_n$ und die Vereinigung ist disjunkt. Für $\sigma \in A_n$ gilt $\operatorname{sgn}(\sigma) = +1$ und $\operatorname{sgn}(\tau\sigma) = -1$. Nun ist

$$\begin{aligned} \det(A) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{1,\sigma(1)} \cdot \dots \cdot a_{n,\sigma(n)} \\ &= \sum_{\sigma \in A_n} a_{1,\sigma(1)} \cdot \dots \cdot a_{n,\sigma(n)} - \sum_{\sigma \in A_n} a_{1,\tau(\sigma(1))} \cdot \dots \cdot a_{n,\tau(\sigma(n))} \\ &= \sum_{\sigma \in A_n} a_{1,\sigma(1)} \cdot \dots \cdot a_{n,\sigma(n)} - \sum_{\sigma \in A_n} a_{1,\sigma(1)} \cdot \dots \cdot a_{n,\sigma(n)} \\ &= 0. \end{aligned}$$

□

Satz 4.14. Sei $f : K^{n \times n} \rightarrow K$ eine beliebige alternierende Multilinearform der Spalten. Dann gilt

$$f(A) = \det(A) \cdot f(1_n)$$

für alle $A \in K^{n \times n}$.

Folgerung 4.15. Die Determinante ist die einzige alternierende Multilinearform f , die normiert ist, das heißt, für die

$$f(1_n) = 1$$

gilt. In diesem Sinne ist die Determinante eindeutig bestimmt.

Beweis von Satz 4.14. Die j -te Spalte a_j von A ist eine Linearkombination der Einheitsvektoren,

$$a_j = a_{1j}e_1 + a_{2j}e_2 + \dots + a_{nj}e_n.$$

Aufgrund der Multilinearität von f gilt daher:

$$\begin{aligned} f(A) &= \sum_{i=1}^n a_{i1} \cdot f(e_i, a_2, \dots, a_n) \\ &= \sum_{i_1, \dots, i_n=1}^n a_{i_1,1} \cdot a_{i_2,2} \cdot \dots \cdot a_{i_n,n} \cdot f(e_{i_1}, e_{i_2}, \dots, e_{i_n}). \end{aligned}$$

Da f alternierend ist, gilt

$$f(e_{i_1}, e_{i_2}, \dots, e_{i_n}) = 0,$$

falls $i_k = i_l$ für ein Paar (k, l) mit $k \neq l$. Der zugehörige Summand ist daher gleich 0, es sei denn die Abbildung

$$k \mapsto i_k$$

ist eine Permutation der Menge $\{1, \dots, n\}$. Folglich gilt

$$f(A) = \sum_{\sigma \in S_n} a_{\sigma(1),1} \cdot \dots \cdot a_{\sigma(n),n} \cdot f(e_{\sigma(1)}, \dots, e_{\sigma(n)}).$$

Da f alternierend ist, gilt

$$\begin{aligned} f(e_{\sigma(1)}, \dots, e_{\sigma(n)}) &= \operatorname{sgn}(\sigma) f(e_1, \dots, e_n) \\ &= \operatorname{sgn}(\sigma) f(1_n). \end{aligned}$$

Es folgt

$$f(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdot \dots \cdot a_{\sigma(n),n} \cdot f(1_n) = \det(A) \cdot f(1_n).$$

□

4.2 Weitere Eigenschaften der Determinante

Satz 4.16 (Determinantenmultiplikationssatz). *Für $A, B \in K^{n \times n}$ gilt*

$$\det(A \cdot B) = \det(A) \cdot \det(B).$$

Beweis. Sei A fest. Betrachte die Funktion $f : K^{n \times n} \rightarrow K$, gegeben durch

$$\begin{aligned} f(B) &= \det(A \cdot B) \\ &= \det(A \cdot b_1, A \cdot b_2, \dots, A \cdot b_n), \end{aligned}$$

wobei $B = (b_1, \dots, b_n)$ mit $b_i \in K^n$. Wegen Satz 4.13 ist f eine alternierende Multilinearform in den Spalten von B . Mit Satz 4.14 folgt

$$\begin{aligned} f(B) &= \det(B) \cdot f(1_n) \\ &= \det(B) \cdot \det(A). \end{aligned}$$

□

Satz 4.17 (Entwicklung nach der i -ten Zeile). *Sei $A = (a_{ij}) \in K^{n \times n}$ und $i \in \{1, \dots, n\}$. Sei A_{ij} die $(n-1) \times (n-1)$ -Matrix, die man erhält, indem man die i -te Zeile und die j -te Spalte von A streicht, also*

$$A_{ij} = \begin{pmatrix} a_{1,1} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{n,n} \end{pmatrix}.$$

Dann gilt

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

Beispiel 4.18. Sei

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in K^{3 \times 3}.$$

Entwickelt man die Determinante von A nach der 1. Zeile, so folgt

$$\det(A) = a_{11} \cdot \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \cdot \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \cdot \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}.$$

Um Satz 4.17 zu beweisen, benötigen wir einige Lemmata.

Lemma 4.19. Sei $M \in K^{n \times n}$ eine Block-Dreiecksmatrix

$$M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

mit $A \in K^{r \times r}$, $B \in K^{r \times (n-r)}$ und $D \in K^{(n-r) \times (n-r)}$. Dann gilt

$$\det(M) = \det(A) \cdot \det(D).$$

Beweis.

1) Sei zunächst $A = 1_r$. Die Funktion $f : K^{(n-r) \times (n-r)} \rightarrow K$ gegeben durch

$$f(D) = \det \begin{pmatrix} 1_r & B \\ 0 & D \end{pmatrix}$$

ist eine alternierende Multilinearform der Zeilen von D . Mit Satz 4.14 folgt

$$f(D) = \det(D) \cdot \det \begin{pmatrix} 1_r & B \\ 0 & 1_{n-r} \end{pmatrix}.$$

Aus der Leibniz-Formel folgt leicht, dass

$$\det \begin{pmatrix} 1_r & B \\ 0 & 1_{n-r} \end{pmatrix} = 1.$$

Es folgt

$$f(D) = \det(D).$$

2) Betrachte nun für festes D die Funktion $g : K^{r \times r} \rightarrow K$ gegeben durch

$$g(A) = \det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}.$$

Diese ist eine alternierende Multilinearform in den Spalten von A . Mit Satz 4.14 und Teil 1) folgt nun

$$\begin{aligned} g(A) &= \det(A) \cdot \det \begin{pmatrix} 1_r & B \\ 0 & D \end{pmatrix} \\ &= \det(A) \cdot \det(D). \end{aligned}$$

□

Lemma 4.20. Sei $A = (a_{ij}) \in K^{n \times n}$. Mit den Bezeichnungen von Satz 4.17 gilt

$$\det(A_{ij}) = (-1)^{i+j} \det(a_1, \dots, a_{j-1}, e_i, a_{j+1}, \dots, a_n),$$

wobei $e_i \in K^n$ der i -te Einheitsvektor ist, und a_k die k -te Spalte von A .

Beweis. Nach Lemma 4.19 ist

$$\det(A_{ij}) = \det \left(\begin{array}{c|c} 1 & * \\ \hline 0 & A_{ij} \end{array} \right).$$

Durch $(j-1)$ Vertauschungen benachbarter Spalten und $(i-1)$ Vertauschungen benachbarter Zeilen bringt man die Matrix in die Form

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ * & \dots & * & 1 & * & \dots & * \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{-te Zeile.}$$

$\uparrow$
 $j\text{-te Spalte}$

Dabei sind die Einträge außerhalb der i -ten Zeile und der j -ten Spalte gleich den Einträgen von A . Wählt man $*$ geeignet, so erhält man die Matrix $(a_1, \dots, a_{j-1}, e_i, a_{j+1}, \dots, a_n)$. Folglich gilt

$$\det(A_{ij}) = (-1)^{(i-1)+(j-1)} \det(a_1, \dots, a_{j-1}, e_i, a_{j+1}, \dots, a_n).$$

□

Lemma 4.21. Sei $A = (a_{ij}) \in K^{n \times n}$. Sei $\tilde{A} = (\tilde{a}_{ij}) \in K^{n \times n}$ die zu A komplementäre Matrix, definiert durch

$$\tilde{a}_{ij} = (-1)^{i+j} \det(A_{ji}).$$

(Man beachte die Umkehrung der Reihenfolge der Indizes auf der rechten Seite.) Dann gilt

$$\tilde{A} \cdot A = A \cdot \tilde{A} = \det(A) \cdot 1_n.$$

Beweis. Mit Hilfe von Lemma 4.20 bestimmen wir den Eintrag von $\tilde{A} \cdot A$ an der Stelle (i, k) . Dieser ist gleich

$$\begin{aligned} \sum_{j=1}^n \tilde{a}_{ij} a_{jk} &= \sum_{j=1}^n a_{jk} \cdot (-1)^{i+j} \cdot \det(A_{ji}) \\ &= \sum_{j=1}^n a_{jk} \cdot \det(a_1, \dots, a_{i-1}, e_j, a_{i+1}, \dots, a_n) \\ &= \det(a_1, \dots, a_{i-1}, \sum_{j=1}^n a_{jk} e_j, a_{i+1}, \dots, a_n) \\ &= \det(a_1, \dots, a_{i-1}, a_k, a_{i+1}, \dots, a_n) \\ &= \delta_{ik} \cdot \det(A) \end{aligned}$$

mit

$$\delta_{ik} = \begin{cases} 1, & i = k, \\ 0, & i \neq k. \end{cases}$$

Die Matrix $A \cdot \tilde{A}$ berechnet man analog. □

Korollar 4.22. Eine Matrix $A \in K^{n \times n}$ ist genau dann invertierbar, wenn $\det(A) \neq 0$. In diesem Fall gilt

$$A^{-1} = \frac{1}{\det(A)} \cdot \tilde{A}.$$

Beispiel 4.23. Für $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K^{2 \times 2}$ ist $\tilde{A} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Beweis von Satz 4.17. In der Formel $A \cdot \tilde{A} = \det(A) \cdot 1_n$ betrachten wir die Matrixeinträge an der Stelle (i, k) :

$$\sum_{j=1}^n a_{ij} \tilde{a}_{jk} = \det(A) \cdot \delta_{ik}.$$

Für $i = k$ folgt

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

□

Bemerkung 4.24. Ebenso kann man für $A \in K^{n \times n}$ die Determinante nach der j -ten Spalte entwickeln:

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

Bemerkung 4.25 (Elementare Spaltenumformungen). Sei $A \in K^{n \times n}$.

i) B entstehe aus A durch Addition der j -ten Spalte zur i -ten Spalte ($i \neq j$). Dann gilt

$$\det(B) = \det(A).$$

ii) B entstehe aus A durch Multiplikation der i -ten Spalte mit $\lambda \in K$. Dann gilt

$$\det(B) = \lambda \cdot \det(A).$$

iii) B entstehe aus A durch Vertauschen zweier Spalten. Dann gilt

$$\det(B) = -\det(A).$$

iv) B entstehe aus A durch Addition des λ -fachen der j -ten Spalte zur i -ten Spalte ($i \neq j$). Dann gilt

$$\det(B) = \det(A).$$

Beweis.

iii) Die Determinante ist alternierend.

ii) Die Determinante ist linear in der i -ten Spalte.

iv) Folgt aus Multilinearität und daher, dass die Determinante alternierend ist:

$$\begin{aligned} \det(A) &= \det(a_1, \dots, a_n) + \underbrace{\lambda \cdot \det(a_1, \dots, a_{i-1}, a_j, a_{i+1}, \dots, a_j, \dots, a_n)}_{=0} \\ &= \det(a_1, \dots, a_n) + \det(a_1, \dots, a_{i-1}, \lambda a_j, a_{i+1}, \dots, a_j, \dots, a_n) \\ &= \det(a_1, \dots, a_{i-1}, a_i + \lambda a_j, a_{i+1}, \dots, a_j, \dots, a_n) \\ &= \det(B). \end{aligned}$$

i) Folgt aus iv).

□

Bemerkung 4.26. Bemerkung 4.25 gilt analog für elementare Zeilenumformungen.

Praktische Berechnung der Determinante:

- Man bringe $A \in K^{n \times n}$ durch elementare Zeilenumformungen und Spaltenvertauschungen auf Normalform

$$B = \left(\begin{array}{ccc|c} b_{1,1} & & * & \\ & \ddots & & * \\ 0 & & b_{r,r} & \\ \hline & 0 & & 0 \end{array} \right) \quad (4.1)$$

- Man führe Buch über die Faktoren λ bei Umformungen vom Typ (ii) und die Anzahl der Vertauschungen vom Typ (iii).
- Benutze $\det(B) = b_{11} \cdot \dots \cdot b_{nn}$.

4.2.1 Übungen

Aufgabe 4.1. Zeigen Sie: Für $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ ist $|\det(A)|$ der Flächeninhalt des von den Spaltenvektoren von A aufgespannten Parallelogramms.

Aufgabe 4.2. Gegeben Sei eine beliebige 2×2 -Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{C}^{2 \times 2}$$

mit komplexen Einträgen a, b, c und d . Die Determinante einer solchen Matrix ist definiert durch $\det(A) := ad - bc$, die Spur als $\text{Spur}(A) := a + d$.

1. Zeigen Sie, dass

$$\det(A - \lambda E_2) = \lambda^2 - \lambda \text{Spur}(A) + \det(A)$$

für alle $\lambda \in \mathbb{C}$ gilt.

2. Zeigen Sie, dass

$$\det(A + B) = \det(A) + \det(B) + \text{Spur}(BA^\#)$$

für alle Matrizen $A, B \in \mathbb{C}^{2 \times 2}$ gilt, wobei $A^\# = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ für $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist.

Aufgabe 4.3. Sei K ein Körper der Charakteristik ungleich 2 (das heißt $1 + 1 \neq 0$ in K , also $\frac{1}{2} \in K$).

1. Sei $M \in K^{n \times n}$ eine obere oder untere Dreiecksmatrix. Zeigen Sie, dass die Determinante von M das Produkt ihrer Diagonaleinträge ist.

2. Sei $f : K^{n \times n} \rightarrow K$ eine Multilinearform (der Spalten). Zeigen Sie, dass die folgenden beiden Eigenschaften äquivalent sind:

- (a) $f(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = -f(a_1, \dots, a_j, \dots, a_i, \dots, a_n)$ für alle $i \neq j$.
 (b) $f(A) = 0$ für jede Matrix A , die zwei gleiche Spalten besitzt.

Aufgabe 4.4. Wir betrachten $\mathbb{Z}/5\mathbb{Z}$ als Körper mit 5 Elementen und die Matrix

$$A_\lambda := \begin{pmatrix} 1 & 2 & \lambda & 0 \\ 0 & \lambda + 4 & 0 & 0 \\ 4\lambda & 3 & 1 & 0 \\ 2 & 4 & 2\lambda & \lambda + 1 \end{pmatrix}$$

mit Einträgen aus $\mathbb{Z}/5\mathbb{Z}$.

1. Berechnen Sie $\det(A_\lambda) \in \mathbb{Z}/5\mathbb{Z}$.
2. Bestimmen Sie alle Werte $\lambda \in \mathbb{Z}/5\mathbb{Z}$ für welche A_λ invertierbar ist.

Aufgabe 4.5. Zeigen Sie mit Hilfe einer Induktion nach $n \in \mathbb{N}$, dass

$$V_n(x_1, \dots, x_n) := \det \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix} = \prod_{1 \leq i < j \leq n} (x_j - x_i)$$

gilt. Dabei seien $x_1, \dots, x_n \in \mathbb{R}$ beliebig und $n \geq 1$. Eine Matrix dieser Form heißt auch *Vandermonde-Matrix*.

Hinweis: Bringen Sie die Matrix zunächst durch Umformungen, die die Determinante der Matrix nicht ändern, in eine Form, die es Ihnen ermöglicht, nach einer Zeile oder Spalte zu entwickeln.

4.3 Die Determinante eines Endomorphismus

Sei V ein endlichdimensionaler K -Vektorraum. Ein *Endomorphismus* von V ist eine K -lineare Abbildung $f : V \rightarrow V$. Ist f bijektiv, so heißt f auch *Automorphismus*.

Notation.

$$\begin{aligned} \text{End}(V) &= \text{Hom}(V, V), \\ \text{Aut}(V) &= \text{Iso}(V, V). \end{aligned}$$

Sei $f \in \text{End}(V)$.

Ziel. Definiere $\det(f)$.

Dazu wählt man eine Basis von V und betrachtet die darstellende Matrix $A = \mathcal{M}_B^B(f)$. Definiere $\det(f) = \det(A)$.

Problem. Was passiert, wenn man eine andere Basis B' von V wählt? In diesem Fall erhält man

$$A' = \mathcal{M}_{B'}^{B'}(f) \quad \text{und} \quad \det(f) = \det(A').$$

Mit der Basiswechselmatrix $S = \mathcal{M}_{B'}^B(\text{id}_V)$ gilt aber

$$A' = S \cdot A \cdot S^{-1}.$$

Es folgt

$$\begin{aligned} \det(A') &= \det(S) \cdot \det(A) \cdot \det(S^{-1}) \\ &= \det(A) \cdot \det(S \cdot S^{-1}) \\ &= \det(A), \end{aligned}$$

das heißt $\det(f)$ ist unabhängig von der Wahl der Basis von V .

Bemerkung 4.27. Es gibt eine eindeutige Abbildung $\det : \text{End}(V) \rightarrow K$ mit der Eigenschaft: Wird f bezüglich irgendeiner Basis B von V durch $A \in K^{n \times n}$ dargestellt, so gilt

$$\det(f) = \det(A).$$

Das heißt, das Diagramm

$$\begin{array}{ccc} \text{End}(V) & \xrightarrow{\det} & K \\ f \downarrow & \searrow \swarrow & \uparrow \\ \mathcal{M}_B^B(f) & \xrightarrow{A} & \det(A) \\ & \downarrow & \\ & K^{n \times n} & \end{array}$$

kommutiert.

Bemerkung 4.28. Für $f, g \in \text{End}(V)$ gilt

$$\det(f \circ g) = \det(f) \cdot \det(g).$$

Beweis. Dies folgt aus dem determinantenmultiplikationssatz, Satz 4.16. □

Notation. Die *lineare Gruppe* oder *Automorphismengruppe von V* bezeichnet man mit

$$\text{GL}(V) = \text{Aut}(V) = \{f \in \text{End}(V) \mid f \text{ ist invertierbar}\}.$$

Bemerkung 4.29. Sei B eine Basis von V . Dann definiert die Abbildung $f \mapsto \mathcal{M}_B^B(f)$ einen Gruppenisomorphismus

$$\mathrm{GL}(V) \rightarrow \mathrm{GL}_n(K).$$

Bemerkung 4.30. Ein Endomorphismus $f \in \mathrm{End}(V)$ ist genau dann invertierbar, wenn $\det(f) \neq 0$. Die Zuordnung $f \mapsto \det(f)$ definiert einen Gruppenhomomorphismus

$$\mathrm{GL}(V) \rightarrow K^\times.$$

5 Eigenwerte und Eigenvektoren

5.1 Diagonalisierbarkeit

Sei K ein Körper, V ein endlichdimensionaler K -Vektorraum. Sei $f : V \rightarrow V$ ein *Endomorphismus*, das heißt eine K -lineare Abbildung von V nach V . Ist $B = (b_1, \dots, b_n)$ eine Basis von V , so können wir f durch die Koordinatenmatrix $A = \mathcal{M}_B^B(f)$ darstellen. Dabei ist $A = (a_{ij}) \in K^{n \times n}$ mit

$$f(b_j) = \sum_{i=1}^n a_{ij} b_i.$$

Frage. Kann man B so wählen, dass A besonders „einfach“ wird?

Anders ausgedrückt:

Gibt es ein Koordinatensystem in V , in dem sich f „einfach“ darstellen lässt?

Was heißt hier einfach?

Wir wollen, dass A eine Diagonalmatrix,

$$A = \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix},$$

ist. Dies bedeutet

$$f(b_j) = a_{jj} b_j$$

für alle $j = 1, \dots, n$, das heißt f bewirkt eine Streckung um den Faktor a_{jj} in Richtung b_j .

Genauere Frage. Sei $f : V \rightarrow V$ ein Endomorphismus von V . Gibt es eine Basis $B = (b_1, \dots, b_n)$ von V , so dass die darstellende Matrix

$$A = \mathcal{M}_B^B(f)$$

eine Diagonalmatrix ist?

Äquivalente Frage. Sei $A \in K^{n \times n}$. Gibt es eine Matrix $S \in \text{GL}_n(K)$, so dass

$$SAS^{-1}$$

eine Diagonalmatrix ist?

In diesem Fall sagt man, dass f bzw. A *diagonalisierbar* ist. Die Äquivalenz der beiden Fragen erkennt man, indem man S als Basiswechselmatrix interpretiert.

Achtung.

- Es gibt Matrizen, die nicht diagonalisierbar sind.

- Es gibt Matrizen $A \in K^{n \times n}$ die nicht über dem Körper K , aber über einem größeren Körper $L \supset K$ diagonalisierbar sind.

Beispiel 5.1. Die Matrix

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

ist nicht diagonalisierbar.

Beweis. Angenommen es gibt eine Matrix $S = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_2(\mathbb{C})$ mit

$$SAS^{-1} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Es folgt

$$\begin{pmatrix} 0 & a \\ 0 & c \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = SA = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} S = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda_1 a & \lambda_1 b \\ \lambda_2 c & \lambda_2 d \end{pmatrix}.$$

Wäre $\lambda_1 \neq 0$, so folgt durch den Vergleich der linken und rechten Seite $a = 0$, woraus wiederum $b = 0$ folgt. Dies steht jedoch im Widerspruch zur Invertierbarkeit von S . Es muss daher $\lambda_1 = 0$ gelten. Analog folgt $\lambda_2 = 0$. Wegen $\lambda_1 = \lambda_2 = 0$ ist die rechte Seite (und damit die linke Seite) die Nullmatrix, woraus $a = c = 0$ folgt. Dies widerspricht der Invertierbarkeit von S , es kann folglich keine Matrix S mit den gewünschten Eigenschaften geben. $\square$

Beispiel 5.2. Die Matrix

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

ist nicht diagonalisierbar über $\mathbb{R}$. Sie ist jedoch diagonalisierbar über $\mathbb{C}$.

Beweis. Über $\mathbb{R}$:

Angenommen es gibt eine Matrix $S = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_2(\mathbb{R})$ mit

$$SAS^{-1} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Berechnet man die Determinanten beider Seiten, so folgt

$$1 = \det(A) = \det(SAS^{-1}) = \det \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \lambda_1 \lambda_2.$$

Dis bedeutet insbesondere, dass $\lambda_1 \neq 0$. Weiterhin folgt

$$\begin{pmatrix} b & -a \\ d & -c \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = SA = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} S = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda_1 a & \lambda_1 b \\ \lambda_2 c & \lambda_2 d \end{pmatrix}.$$

Es muss daher

$$b = \lambda_1 a, \quad -a = \lambda_1 b$$

gelten. Wäre $b = 0$, so wäre auch $a = 0$ und S nicht invertierbar. Es gilt daher $b \neq 0$. Aus den obigen Gleichungen folgt

$$b = -\lambda_1^2 b$$

und wegen $b \neq 0$

$$\lambda_1^2 = -1.$$

Diese Gleichung hat keine Lösung in $\mathbb{R}$. Es kann daher kein $S \in \text{GL}_2(\mathbb{R})$ mit den geforderten Eigenschaften geben.

Über $\mathbb{C}$:

In $\mathbb{C}$ hat die Gleichung

$$\lambda_1^2 = -1$$

die Lösung $\lambda_1 = i$. In der Tat ist

$$SAS^{-1} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

für

$$S = \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}.$$

□

Beispiel 5.3. Die Matrix

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

ist diagonalisierbar über $\mathbb{R}$. Es gilt

$$SAS^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

für

$$S = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \in \text{GL}_2(\mathbb{R}).$$

Wir wollen die Frage der Diagonalisierbarkeit nun genauer untersuchen.

Ist die darstellende Matrix eines Endomorphismus $f : V \rightarrow V$ bezüglich der Basis $B = (b_1, \dots, b_n)$ diagonal, so gilt

$$f(b_i) = \lambda_i b_i$$

mit $\lambda_i \in K$ für $i = 1, \dots, n$. Dies führt zu

Definition 5.4. Sei $f : V \rightarrow V$ ein Endomorphismus von V . Ein Vektor $v \in V$ heißt *Eigenvektor* von f , falls

i) $v \neq 0$,

ii) es gibt ein $\lambda \in K$ mit

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

Der Skalar $\lambda \in K$ heißt dann *Eigenwert* von f (zum Eigenvektor v).

Der Endomorphismus f ist genau dann diagonalisierbar, wenn V eine Basis aus Eigenvektoren von f besitzt.

Satz 5.5. *Es sind äquivalent:*

i) $\lambda \in K$ ist Eigenwert von f (zu einem geeigneten Eigenvektor).

ii) $\det(f - \lambda \cdot \text{id}_V) = 0$.

iii) Ist A die darstellende Matrix von f bezüglich einer gewählten Basis, so gilt

$$\det(A - \lambda \cdot E) = 0.$$

Beweis. (ii) $\Leftrightarrow$ (iii): klar.

(i) $\Rightarrow$ (ii): Sei $\lambda \in K$ ein Eigenwert von f zum Eigenvektor $v \in V$, das heißt $f(v) = \lambda v$. Es folgt

$$0 = f(v) - \lambda \cdot v = (f - \lambda \cdot \text{id}_V)(v).$$

Da $v \neq 0$ im Kern von $f - \lambda \cdot \text{id}_V$ liegt, gilt

$$\text{Kern}(f - \lambda \cdot \text{id}_V) \neq \{0\}.$$

$f - \lambda \cdot \text{id}_V : V \rightarrow V$ ist daher nicht invertierbar und $\det(f - \lambda \cdot \text{id}_V) = 0$.

(ii) $\Rightarrow$ (i): Sei $\det(f - \lambda \cdot \text{id}_V) = 0$. $f - \lambda \cdot \text{id}_V : V \rightarrow V$ ist dann nicht invertierbar und es gilt

$$\text{Kern}(f - \lambda \cdot \text{id}_V) \neq \{0\}.$$

Es gibt daher einen Vektor $v \neq 0$ im Kern von $f - \lambda \cdot \text{id}_V$. Für diesen Vektor gilt

$$f(v) - \lambda \cdot v = (f - \lambda \cdot \text{id}_V)(v) = 0.$$

Das heißt

$$f(v) = \lambda v,$$

λ ist ein Eigenwert von f zum Eigenvektor v . □

Folgerung 5.6. *Um die Eigenwerte von f zu finden, muss man die Funktion*

$$P_f : K \rightarrow K$$

$$\lambda \mapsto P_f(\lambda) = \det(\lambda \cdot \text{id}_V - f) = \det(\lambda \cdot E - A)$$

betrachten.

Korollar 5.7. $\lambda \in K$ ist genau dann ein Eigenwert von f , wenn $P_f(\lambda) = 0$.

Beispiel 5.8. Für die Matrix $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ gilt

$$P_A(\lambda) = P_{f_A}(\lambda) = \det(\lambda \cdot E - A) = \det \begin{pmatrix} \lambda & -1 \\ 0 & \lambda \end{pmatrix} = \lambda^2.$$

Hieraus folgt, dass A genau einen Eigenwert, nämlich 0, hat.

Für die Matrix $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ gilt

$$P_B(\lambda) = \det(\lambda \cdot E - B) = \det \begin{pmatrix} \lambda & 1 \\ -1 & \lambda \end{pmatrix} = \lambda^2 + 1 = (\lambda - i)(\lambda + i).$$

B hat keine reellen Eigenwerte, aber die komplexen Eigenwerte $\pm i$.

Für die Matrix $C = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ gilt

$$P_C(\lambda) = \det \begin{pmatrix} \lambda - 2 & -1 \\ -1 & \lambda - 2 \end{pmatrix} = (\lambda - 2)^2 - 1 = \lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1).$$

Daher hat C die Eigenwerte 1 und 3.

Die Funktion P_f ist eine polynomiale Funktion in λ . Bevor wir sie weiter untersuchen, stellen wir einige Grundlagen über Polynome zusammen.

5.2 Ringe und Polynome

Definition 5.9. Ein *Ring* ist eine Menge R mit zwei inneren Verknüpfungen

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

so dass gilt:

(R1) $(R, +)$ ist eine abelsche Gruppe.

(R2) Die Verknüpfung „ $\cdot$ “ ist assoziativ und es existiert dafür ein neutrales Element.

(R3) Es gelten die Distributivgesetze. Für $a, b, c \in R$ gilt also:

$$\begin{aligned}(a + b) \cdot c &= a \cdot c + b \cdot c \\ c \cdot (a + b) &= c \cdot a + c \cdot b.\end{aligned}$$

Der Ring R heißt *kommutativ*, falls die Verknüpfung $\cdot$ kommutativ ist. Das bezüglich $+$ neutrale Element bezeichnet man mit 0 , das bezüglich $\cdot$ neutrale Element mit 1 .

Beispiel 5.10. Die ganzen Zahlen $(\mathbb{Z}, +, \cdot)$ sind ein kommutativer Ring. Jeder Körper ist ein kommutativer Ring. Ist K ein Körper, so bildet $K^{n \times n}$ mit der Addition und Multiplikation von Matrizen einen nicht-kommutativen Ring. Ist V ein K -Vektorraum, so bildet $\text{End}(V)$ mit der Addition von Endomorphismen und der Verkettung von Abbildungen einen nicht-kommutativen Ring, den *Endomorphismenring* von V .

Bemerkung 5.11. In einem Ring R gelten für alle $a, b \in R$:

- (i) $0 \cdot a = a \cdot 0 = 0$
- (ii) $(-a) \cdot b = -(a \cdot b) = a \cdot (-b)$

Beweis. (i) Mit $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$ folgt die Behauptung durch Kürzen. (ii) und (iii): Übung. $\square$

Bemerkung 5.12. Für einen Ring R gilt $1 = 0$ genau dann, wenn $R = \{0\}$.

Beweis. Ist $0 = 1$, so gilt für alle $a \in R$ wegen Bemerkung 5.11, dass $a = a \cdot 1 = a \cdot 0 = 0$. Die Umkehrung ist klar. $\square$

Ein weiteres wichtiges Beispiel für einen Ring ist der Polynomring über einem Körper. Sei K ein Körper.

Definition 5.13. Der *Polynomring* über K ist die Menge

$$\begin{aligned}K[X] &= K^{(\mathbb{N}_0)} \\ &= \{a : \mathbb{N}_0 \rightarrow K, i \mapsto a_i \mid a_i = 0 \text{ für fast alle } i\}\end{aligned}$$

zusammen mit den inneren Verknüpfungen Addition und Multiplikation von Polynomen.

Addition: Wie in $K^{(\mathbb{N}_0)}$ (komponentenweise). Für $(a_i), (b_i) \in K[X]$ ist

$$(a_i) + (b_i) = (a_i + b_i).$$

Multiplikation: *Nicht* komponentenweise, sondern wie folgt:

Für $(a_i), (b_i) \in K[X]$ definiert man $(a_i) \cdot (b_i) = (c_i)$, wobei

$$c_i = \sum_{j=0}^i a_j \cdot b_{i-j}.$$

Mit diesen Verknüpfungen bildet $K[X]$ einen kommutativen Ring. Die Elemente von $K[X]$ heißen *Polynome*.

Andere Schreibweise:

Ein $(a_i) \in K[X]$ schreiben wir auch als formalen Ausdruck

$$\sum_{i \in \mathbb{N}_0} a_i X^i.$$

Die a_i nennt man auch *Koeffizienten* des Polynoms, X nennt man *Variable*. Damit haben wir für $(a_i), (b_i) \in K[X]$:

$$\begin{aligned} \sum_{i \in \mathbb{N}_0} a_i X^i + \sum_{i \in \mathbb{N}_0} b_i X^i &= \sum_{i \in \mathbb{N}_0} (a_i + b_i) X^i, \\ \left(\sum_{i \in \mathbb{N}_0} a_i X^i \right) \cdot \left(\sum_{i \in \mathbb{N}_0} b_i X^i \right) &= \sum_{i \in \mathbb{N}_0} \left(\sum_{j=0}^i a_j \cdot b_{i-j} \right) X^i. \end{aligned}$$

Bemerkung 5.14. Die Abbildung $K \rightarrow K[X]$, $\lambda \mapsto \lambda \cdot X^0$ ist ein injektiver Gruppenhomomorphismus. Wir identifizieren K mit seinem Bild in $K[X]$.

$K[X]$ ist damit ein K -Vektorraum. Es gilt $\dim_K(K[X]) = \infty$.

Definition 5.15. Der *Grad* eines Polynoms $P = \sum_{i \in \mathbb{N}_0} a_i X^i \in K[X]$ ist definiert als

$$\text{Grad}(P) = \begin{cases} \max\{i \in \mathbb{N}_0 \mid a_i \neq 0\}, & P \neq 0, \\ -\infty, & P = 0. \end{cases}$$

Bemerkung 5.16. Für $P, Q \in K[X]$ gilt

- i) $\text{Grad}(P + Q) \leq \max(\text{Grad}(P), \text{Grad}(Q))$,
- ii) $\text{Grad}(P \cdot Q) = \text{Grad}(P) + \text{Grad}(Q)$.

Beweis. Übung. □

5.2.1 Polynomfunktionen

Sei $P = \sum_{i \in \mathbb{N}_0} a_i X^i = a_0 + a_1 X + \dots + a_n X^n \in K[X]$ gegeben. Durch Einsetzen von Elementen aus K für die Variable X erhält man eine Abbildung

$$\tilde{P} : K \rightarrow K, \quad \lambda \mapsto P(\lambda) = a_0 + a_1 \lambda + \dots + a_n \lambda^n.$$

Wir erhalten eine (K -lineare) Abbildung

$$K[X] \rightarrow \text{Abb}(K, K), \quad P \mapsto \tilde{P}. \quad (5.1)$$

Definition 5.17. Ein $f \in \text{Abb}(K, K)$ heißt Polynomfunktion, falls es ein $P \in K[X]$ mit

$$f = \tilde{P}$$

gibt. $\tilde{P}$ heißt zu P gehörige Polynomfunktion.

Bemerkung 5.18. Die Abbildung (5.1) ist im Allgemeinen nicht injektiv.

Beispiel 5.19. Sei $K = \mathbb{F}_2$, $P = X + X^2 \in \mathbb{F}_2[X]$. Dann ist

$$\begin{aligned}\tilde{P} : \mathbb{F}_2 &\rightarrow \mathbb{F}_2, \\ \tilde{P}(0) &= 0^2 + 0 = 0, \\ \tilde{P}(1) &= 1^2 + 1 = 0,\end{aligned}$$

die Nullfunktion.

Beispiel 5.20. Allgemeiner gilt für $K = \mathbb{F}_p$ und $P = X^p - X \in \mathbb{F}_p[X]$, dass $\tilde{P}$ die Nullfunktion ist.

Beweis. Übung, kleiner Satz von Fermat. □

Bemerkung 5.21. Das Einsetzen von $\lambda \in K$ in Polynome ist mit Summe und Produkt (von Polynomen) verträglich, das heißt

$$(P + Q)(\lambda) = P(\lambda) + Q(\lambda), \quad (P \cdot Q)(\lambda) = P(\lambda) \cdot Q(\lambda).$$

Wir zeigen nun, dass die Abbildung (5.1) injektiv ist, falls $\#K = \infty$.

5.2.2 Division mit Rest

Satz 5.22. Sei $0 \neq g \in K[X]$. Für alle $f \in K[X]$ gibt es eindeutig bestimmte Polynome $q, r \in K[X]$ mit

$$i) \quad f = q \cdot g + r,$$

$$ii) \quad \text{Grad}(r) < \text{Grad}(g). \text{ Dabei ist } r = 0 \text{ möglich.}$$

Beweis. Sei $d = \text{Grad}(g) \in \mathbb{N}_0$.

Eindeutigkeit:

Seien $f = q \cdot g + r = q' \cdot g + r'$ mit $\text{Grad}(r) < d$ und $\text{Grad}(r') < d$ zwei Darstellungen. Es gilt

$$(q - q') \cdot g = r' - r$$

woraus mit Bemerkung 5.16

$$\text{Grad}(q - q') + d = \text{Grad}(r' - r) < d$$

folgt. Dies ist nur möglich, falls

$$\text{Grad}(q - q') < 0.$$

Es folgt $\text{Grad}(q - q') = -\infty$, das heißt $q = q'$. Es muss dann auch $r = r'$ gelten, das heißt beide Darstellungen sind gleich.

Existenz:

Induktion nach $n = \text{Grad}(f)$.

$n < d$: Setze $q = 0$ und $r = f$. Dann gilt $f = 0 \cdot g + f$ und $\text{Grad}(f) < d$.

$n \geq d$: Schreibe $f = c_n X^n + \dots + c_1 X + c_0$ mit $c_n \neq 0$ und $g = a_d X^d + \dots + a_1 X + a_0$. Es ist

$$f_1 = f - c_n \cdot a_d^{-1} \cdot X^{n-d} \cdot g \in K[X]$$

ein Polynom vom Grad kleiner als n . Nach der Induktionsannahme gibt es $q_1, r_1 \in K[X]$ mit

$$f_1 = q_1 \cdot g + r_1 \quad \text{und} \quad \text{Grad}(r_1) < d.$$

Es folgt

$$\begin{aligned} f &= f_1 + c_n \cdot a_d^{-1} \cdot X^{n-d} \cdot g \\ &= q_1 \cdot g + r_1 + c_n \cdot a_d^{-1} \cdot X^{n-d} \cdot g \\ &= (q_1 + c_n \cdot a_d^{-1} \cdot X^{n-d}) \cdot g + r_1. \end{aligned}$$

Die Polynome

$$q = q_1 + c_n \cdot a_d^{-1} \cdot X^{n-d}, \quad r = r_1$$

haben die gewünschten Eigenschaften i) und ii). □

Bemerkung 5.23. Das im Beweis benutzte Verfahren entspricht dem bekannten Algorithmus für Polynomdivision.

Beispiel 5.24. Durch die Polynomdivision

$$\begin{array}{r} (X^3 + X^2 + 1) : (X + 2) = X^2 - X + 2 \\ \underline{X^3 + 2X^2} \\ -X^2 + 1 \\ \underline{-X^2 - 2X} \\ 2X + 1 \\ \underline{2X + 4} \\ -3 \end{array}$$

erhält man

$$(X^3 + X^2 + 1) = \underbrace{(X^2 - X + 2)}_q \cdot \underbrace{(X + 2)}_g \underbrace{- 3}_r.$$

Korollar 5.25. Sei $0 \neq f \in K[X]$ und sei $\lambda \in K$ eine Nullstelle von f , das heißt $f(\lambda) = 0$. Dann gibt es ein $q \in K[X]$ mit $\text{Grad}(q) = \text{Grad}(f) - 1$ und

$$f = q \cdot (X - \lambda).$$

Beweis. Division mit Rest impliziert, dass es $q, r \in K[X]$ mit

$$f = q \cdot (X - \lambda) + r$$

und

$$\text{Grad}(r) \leq 0,$$

also r konstant, gibt. Durch Einsetzen von λ erhält man

$$0 = f(\lambda) = q \cdot (\lambda - \lambda) + r = r.$$

Es folgt die Behauptung. □

Korollar 5.26. *Ein Polynom ungleich Null vom Grad n hat höchstens n Nullstellen in K .*

Beweis. Induktion nach n , Übung. □

Korollar 5.27. *Falls $\#K = \infty$, so ist die Abbildung (5.1),*

$$K[X] \rightarrow \text{Abb}(K, K), \quad P \mapsto \tilde{P},$$

injektiv.

Beweis. Falls $\tilde{P}$ die Nullabbildung ist, so gilt $P(\lambda) = 0$ für alle $\lambda \in K$. Dies bedeutet für $\#K = \infty$, dass P unendlich viele Nullstellen hat. Nach Korollar 5.26 gilt dann $P = 0$. □

Für Körper mit unendlich vielen Elementen (wie zum Beispiel $\mathbb{Q}, \mathbb{R}, \mathbb{C}$) können wir daher Polynome mit Polynomfunktionen identifizieren.

Definition 5.28. Ist $0 \neq P \in K[X]$ und $\lambda \in K$, so heißt

$$\mu(P, \lambda) = \max\{r \in \mathbb{N}_0 \mid P = (X - \lambda)^r \cdot Q \text{ für ein } Q \in K[X]\}$$

die *algebraische Vielfachheit* von λ in P . Ist λ eine Nullstelle von P , so nennt man $\mu(P, \lambda)$ auch die *Vielfachheit* der Nullstelle λ von P .

Nach Korollar 5.25 gilt

$$P(\lambda) \neq 0 \quad \Leftrightarrow \quad \mu(P, \lambda) = 0.$$

Ist $P = (X - \lambda)^r \cdot Q$ mit $r = \mu(P, \lambda)$, so folgt $Q(\lambda) \neq 0$.

5.3 Das charakteristische Polynom

Wir beginnen mit der Bemerkung, dass die Definition der Determinante und deren grundlegende Eigenschaften sich direkt auf Matrizen $B \in R^{n \times n}$ mit Einträgen in einem kommutativen Ring R verallgemeinern lassen. Sei im folgenden K ein Körper.

Definition 5.29. Das *charakteristische Polynom* einer Matrix $A \in K^{n \times n}$ ist definiert als

$$P_A(X) = \det(X \cdot 1_n - A) \in K[X].$$

(Beachte, dass die Matrix $B = X \cdot 1_n - A$ Koeffizienten im Polynomring $K[X]$ hat.)
Das charakteristische Polynom eines Endomorphismus $f : V \rightarrow V$ eines endlich dimensional K -Vektorraums V ist definiert als

$$P_f(X) = P_A(X),$$

wobei A die Koordinatenmatrix von f bezüglich einer Basis B von V ist. (Diese Definition ist unabhängig von der Wahl der Basis B .)

Bemerkung 5.30. Das charakteristische Polynom $P_A(X)$ einer Matrix $A \in K^{n \times n}$ ist ein normiertes Polynom vom Grad n . (Normiert heißt, dass der führende Koeffizient gleich 1 ist, das heißt $P_f(X) = X^n + \text{Terme niedrigerer Ordnung}$.)

Beweis. Wende die Leibnizformel auf die Definition an. □

Definition 5.31. Sei $A = (a_{ij}) \in K^{n \times n}$. Die *Spur* von A ist die Summe der Einträge auf der Diagonalen,

$$\text{Spur}(A) = \sum_{i=1}^n a_{ii}.$$

Die Spur eines Endomorphismus $f : V \rightarrow V$ eines endlich dimensional K -Vektorraums V ist definiert als die Spur einer Koordinatenmatrix von f .

Satz 5.32. Sei $A \in K^{n \times n}$ und

$$P_A(X) = c_0 + c_1 X + \dots + c_n X^n$$

das charakteristische Polynom von A . Es gilt

$$\begin{aligned} c_n &= 1, \\ c_{n-1} &= -\text{Spur}(A), \\ c_0 &= (-1)^n \cdot \det(A). \end{aligned}$$

Beweis. c_n : Bemerkung 5.30.

c_0 :

$$c_0 = P_A(0) = \det(0 \cdot 1_n - A) = \det(-A) = (-1)^n \det(A).$$

c_{n-1} : Beweis mit der Leibniz-Formel (Übung). □

Als Polynom vom Grad n hat das charakteristische Polynom P_A höchstens n Nullstellen in K (Korollar 5.26). Nach Korollar 5.7 sind die Nullstellen von P_A genau die Eigenwerte von A . Es folgt

Bemerkung 5.33. Eine Matrix $A \in K^{n \times n}$ hat höchstens n Eigenwerte. Analog gilt dies für Endomorphismen $f : V \rightarrow V$.

Lemma 5.34. Seien $v_1, \dots, v_m \in V$ Eigenvektoren eines Endomorphismus $f : V \rightarrow V$ zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_m \in K$. Dann sind die Vektoren $v_1, \dots, v_m$ linear unabhängig.

Beweis. Da Eigenvektoren ungleich 0 sind brauchen wir nur den Fall $m \geq 2$ zu betrachten. Dann existiert ein maximales $r \in \{2, \dots, m\}$, so dass $v_1, \dots, v_r$ linear unabhängig sind. Zu zeigen ist $r = m$.

Angenommen es gilt $r < m$. Dann ist also $r + 1 \leq m$ und $v_1, \dots, v_{r+1}$ sind linear abhängig. Seien $\alpha_1, \dots, \alpha_{r+1} \in K$ mit

$$\alpha_1 v_1 + \dots + \alpha_{r+1} v_{r+1} = 0. \quad (*)$$

Wenden wir f auf diese Linearkombination an, so folgt

$$\alpha_1 \lambda_1 v_1 + \dots + \alpha_{r+1} \lambda_{r+1} v_{r+1} = 0.$$

Andererseits können wir $(*)$ mit λ_{r+1} multiplizieren und erhalten

$$\alpha_1 \lambda_{r+1} v_1 + \dots + \alpha_{r+1} \lambda_{r+1} v_{r+1} = 0.$$

Subtrahieren wir diese Gleichung von der oberen, so folgt

$$\alpha_1 (\lambda_1 - \lambda_{r+1}) v_1 + \dots + \alpha_{r+1} (\lambda_{r+1} - \lambda_{r+1}) v_{r+1} = 0.$$

Da die Vektoren $v_1, \dots, v_r$ linear unabhängig und die Eigenwerte $\lambda_1, \dots, \lambda_r$ verschieden von λ_{r+1} sind, folgt

$$\alpha_1, \dots, \alpha_r = 0.$$

Die Linearkombination $(*)$ muss daher trivial sein und somit sind $v_1, \dots, v_{r+1}$ linear unabhängig, im Widerspruch zur Annahme $r < m$. $\square$

5.4 Kriterien für Diagonalisierbarkeit

Sei K ein unendlicher Körper (zum Beispiel $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$) und V ein K -Vektorraum der Dimension $n \in \mathbb{N}$. Sei $f : V \rightarrow V$ ein Endomorphismus und $P_f(X) = \det(X \cdot \text{id}_V - f) \in K[X]$ dessen charakteristisches Polynom (normiertes Polynom vom Grad n).

Frage. Wann ist f diagonalisierbar?

Äquivalente Frage. Wann besitzt V eine Basis aus Eigenvektoren von f ?

Bemerkung 5.35. Wenn f diagonalisierbar ist, so zerfällt P_f in Linearfaktoren in $K[X]$. P_f hat also n Nullstellen in K (mit Vielfachheiten gezählt).

Dieses Kriterium ist notwendig, aber nicht hinreichend.

Beweis. Sei $B = (b_1, \dots, b_n)$ eine Basis von Eigenvektoren. Seien $\lambda_1, \dots, \lambda_n$ die zugehörigen Eigenwerte. Dann ist

$$\mathcal{M}_B^B(f) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

woraus

$$P_f(X) = (X - \lambda_1) \cdot \dots \cdot (X - \lambda_n)$$

folgt. □

Satz 5.36 (Ein hinreichendes Kriterium). *Sei V ein K -Vektorraum der Dimension n und $f : V \rightarrow V$ ein Endomorphismus. Das charakteristische Polynom $P_f(X)$ besitze n paarweise verschiedene Nullstellen in K . Dann gibt es eine Basis von V aus Eigenvektoren von f .*

Beweis. Seien $\lambda_1, \dots, \lambda_n$ die Nullstellen von P_f (also die Eigenwerte von f). Seien $v_1, \dots, v_n \in V$ Eigenvektoren zu $\lambda_1, \dots, \lambda_n$. Nach Annahme sind die λ_i paarweise verschieden. Nach Lemma 5.34 sind dann die Eigenvektoren $v_1, \dots, v_n$ linear unabhängig. Da $\dim(V) = n$ bilden sie eine Basis von V . □

Beispiel 5.37. Für die Matrix $C = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ gilt

$$P_C(X) = \det \begin{pmatrix} X-2 & -1 \\ -1 & X-2 \end{pmatrix} = X^2 - 4X + 3 = (X-3)(X-1).$$

Nach Satz 5.36 ist C diagonalisierbar.

Beispiel 5.38. Für die Matrix $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ gilt

$$P_A(X) = (X-1)^2.$$

Das Polynom P_A hat demnach eine Nullstelle mit Vielfachheit 2. Mit Satz 5.36 ist keine Aussage möglich.

Beispiel 5.39. Für die Matrix $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ gilt

$$P_A(X) = X^2 + 1.$$

Das Polynom P_A hat keine Nullstellen in $\mathbb{R}$. Mit Satz 5.36 ist (für $K = \mathbb{R}$) keine Aussage möglich. In $\mathbb{C}[X]$ gilt $P_A(X) = (X+i)(X-i)$. Nach Satz 5.36 ist A daher über $\mathbb{C}$ diagonalisierbar.

Frage. Wann hat das charakteristische Polynom P_f genau n verschiedene Nullstellen in K ?

Antwort. Dies hängt von P_f und vom Grundkörper K ab. Im allgemeinen muss ein Polynom gar keine Nullstellen in K haben.

Beispiel 5.40. Das Polynom $X^2 + 1 \in \mathbb{R}[X]$ hat keine Nullstellen in $\mathbb{R}$ und das Polynom $X^2 - 2 \in \mathbb{Q}[X]$ hat keine Nullstellen in $\mathbb{Q}$.

Definition 5.41. Ein Körper K heißt *algebraisch abgeschlossen*, falls jedes nicht konstante Polynom in $K[X]$ eine Nullstelle in K hat. Wegen Korollar 5.25 impliziert dies, dass sich jedes Polynom $P \in K[X]$ vom Grad n in der Form

$$P = C \cdot (X - \alpha_1) \cdot (X - \alpha_2) \cdot \dots \cdot (X - \alpha_n)$$

schreiben lässt, wobei $\alpha_1, \dots, \alpha_n \in K$ und $C \in K^\times$.

Bemerkung 5.42.

i) Dann sind C und $\alpha_1, \dots, \alpha_n$ eindeutig durch P bestimmt (Übung).

Beispiel: $C =$ führender Koeffizient.

ii) Die Nullstellen sind genau die α_i .

iii) Die α_i müssen nicht verschieden sein.

Bemerkung 5.43. Ein Polynom $P \in K[X]$ vom Grad $n \geq 0$ über einem algebraisch abgeschlossenen Körper hat genau n Nullstellen (mit Vielfachheiten gezählt).

Satz 5.44 (Fundamentalsatz der Algebra). *Der Körper $\mathbb{C}$ ist algebraisch abgeschlossen.*

Beweis. Der Beweis wird zum Beispiel in den Vorlesungen Funktionentheorie und Algebra geführt. $\square$

Folgerung 5.45. *Über $\mathbb{C}$ lassen sich „mehr“ Matrizen diagonalisieren als über $\mathbb{R}$ oder $\mathbb{Q}$.*

Dieses Kriterium ist hinreichend, aber nicht notwendig.

Ziel. Ein hinreichendes und notwendiges Kriterium.

Definition 5.46. Der *Eigenraum* von f zu λ ist definiert als

$$\text{Eig}(f, \lambda) = \{v \in V \mid f(v) = \lambda \cdot v\} \subset V.$$

Dies ist ein Untervektorraum von V . Die Berechnung von $\text{Eig}(f, \lambda)$ besteht in der Lösung eines linearen Gleichungssystems. Es gilt:

$$\begin{aligned} & \lambda \text{ ist Eigenwert von } f \\ \Leftrightarrow & \text{Eig}(f, \lambda) \neq \{0\} \\ \Leftrightarrow & \dim(\text{Eig}(f, \lambda)) > 0. \end{aligned}$$

Die *geometrische Vielfachheit* von λ (bezüglich f) ist definiert als $\dim(\text{Eig}(f, \lambda))$.

Bemerkung 5.47. Es gilt stets

$$\mu(P_f, \lambda) \geq \dim(\text{Eig}(f, \lambda)).$$

Beweis. Sei $v_1, \dots, v_r$ eine Basis des Eigenraums $\text{Eig}(f, \lambda)$. Ergänze diese zu einer Basis $B = (v_1, \dots, v_r, v_{r+1}, \dots, v_n)$ von V . Die Koordinatenmatrix von f bezüglich B hat die Form

$$A := \mathcal{M}_B^B(f) = \left(\begin{array}{ccc|ccc} \lambda & & 0 & & & \\ & \ddots & & & & \\ 0 & & \lambda & & & \\ \hline & & & & & \\ & 0 & & & A' & \end{array} \right) \left. \begin{array}{l} \left. \begin{array}{ccc} & & \\ & & \\ & & \end{array} \right\} r \\ \left. \begin{array}{ccc} & & \\ & & \\ & & \end{array} \right\} n-r. \end{array} \right.$$

Es gilt daher

$$P_f(X) = (X - \lambda)^r \cdot \det(X \cdot E_{n-r} - A'),$$

woraus

$$\mu(P_f, \lambda) \geq r$$

folgt. □

Beispiel 5.48. Für die Matrix $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ und den Endomorphismus $f = f_A$ gilt

$$P_A(X) = P_f(X) = X^2.$$

A hat nur den Eigenwert 0 mit algebraischer Vielfachheit $\mu(P_A, 0) = 2$. Der zugehörige Eigenraum $\text{Eig}(A, 0)$ ist die Lösung des linearen Gleichungssystems

$$A \cdot \begin{pmatrix} a \\ b \end{pmatrix} = 0.$$

Für eine Lösung $\begin{pmatrix} a \\ b \end{pmatrix}$ gilt

$$\begin{pmatrix} 0 \\ b \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = A \cdot \begin{pmatrix} a \\ b \end{pmatrix} = 0,$$

das heißt $b = 0$. Es folgt

$$\text{Eig}(A, 0) = \mathbb{R} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Daher gilt

$$\dim(\text{Eig}(A, 0)) = 1.$$

Satz 5.49. Sei $f : V \rightarrow V$ eine Endomorphismus. Es sind äquivalent:

- 1) f ist diagonalisierbar.
- 2) i) P_f zerfällt in $K[X]$ vollständig in Linearfaktoren (hat also mit Vielfachheiten genau $n = \dim(V)$ Nullstellen in K).

ii) Für jede Nullstelle λ von P_f gilt

$$\mu(P_f, \lambda) = \dim(\text{Eig}(f, \lambda)).$$

3) Sind $\lambda_1, \dots, \lambda_r$ die paarweise verschiedenen Eigenwerte von f , so gilt

$$V = \text{Eig}(f, \lambda_1) \oplus \dots \oplus \text{Eig}(f, \lambda_r).$$

Beweis. „(1) $\Rightarrow$ (2)“: Klar (wie Bemerkung 5.35).

„(2) $\Rightarrow$ (1)“: Wir können nach i) P_f als Produkt von Linearfaktoren

$$P_f(X) = (X - \lambda_1)^{n_1} \cdot \dots \cdot (X - \lambda_r)^{n_r} \in K[X]$$

schreiben, wobei $\lambda_1, \dots, \lambda_r$ die paarweise verschiedenen Nullstellen von P_f mit zugehörigen algebraischen Vielfachheiten und $n_i = \mu(P_f, \lambda_i)$ sind. Es gilt

$$n = \text{Grad}(P_f) = n_1 + \dots + n_r.$$

Nach ii) ist

$$n_i = \dim(\text{Eig}(f, \lambda_i)).$$

Sei $v_1^{(i)}, \dots, v_{n_i}^{(i)}$ eine Basis von $\text{Eig}(f, \lambda_i)$ für $i = 1, \dots, r$. Wir betrachten das n -Tupel

$$B = (v_1^{(1)}, \dots, v_{n_1}^{(1)}, v_1^{(2)}, \dots, v_{n_2}^{(2)}, \dots, v_1^{(r)}, \dots, v_{n_r}^{(r)})$$

von Vektoren aus V .

Behauptung: B ist linear unabhängig (und damit eine Basis aus Eigenvektoren).

Beweis: Seien $\alpha_j^{(i)} \in K$ für $i = 1, \dots, r$ und $j = 1, \dots, n_i$ mit

$$\begin{aligned} & \alpha_1^{(1)} \cdot v_1^{(1)} + \dots + \alpha_{n_1}^{(1)} \cdot v_{n_1}^{(1)} \\ & + \alpha_1^{(2)} \cdot v_1^{(2)} + \dots + \alpha_{n_2}^{(2)} \cdot v_{n_2}^{(2)} \\ & + \dots + \alpha_1^{(r)} \cdot v_1^{(r)} + \dots + \alpha_{n_r}^{(r)} \cdot v_{n_r}^{(r)} = 0. \end{aligned}$$

Schreibe

$$w_i = \alpha_1^{(i)} \cdot v_1^{(i)} + \dots + \alpha_{n_i}^{(i)} \cdot v_{n_i}^{(i)} \in \text{Eig}(f, \lambda_i).$$

Dann gilt

$$w_1 + \dots + w_r = 0.$$

Mit Lemma 5.34 folgt

$$w_1 = \dots = w_r = 0,$$

das heißt

$$\alpha_1^{(i)} \cdot v_1^{(i)} + \dots + \alpha_{n_i}^{(i)} \cdot v_{n_i}^{(i)} = 0$$

für alle $i = 1, \dots, r$. Da die Vektoren $v_j^{(i)}$ für jedes i linear unabhängig sind folgt

$$\alpha_j^{(i)} = 0$$

für alle $i = 1, \dots, r$ und $j = 1, \dots, n_i$. Die Behauptung ist damit bewiesen.

„(1) $\Rightarrow$ (3)“: klar.

„(3) $\Rightarrow$ (1)“: Übung. □

5.5 Trigonalisierung

Definition 5.50. Sei V ein endlichdimensionaler K -Vektorraum. Ein Endomorphismus $f : V \rightarrow V$ heißt *trigonalisierbar*, falls V eine Basis B besitzt, so dass

$$\mathcal{M}_B^B(f) = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$$

eine obere Dreiecksmatrix ist.

Satz 5.51. Ein Endomorphismus $f : V \rightarrow V$ ist genau dann trigonalisierbar, wenn $P_f \in K[X]$ vollständig in Linearfaktoren zerfällt.

Beweis. „ $\Rightarrow$ “: Klar.

„ $\Leftarrow$ “: Es sei $P_f(X) = (X - \lambda_1) \cdot \dots \cdot (X - \lambda_n)$ (wobei die λ_i nicht verschieden sein müssen). Sei $0 \neq v_1 \in V$ ein Eigenvektor von f zum Eigenwert λ_1 . Diesen können wir zu einer Basis $v_1, v_2, \dots, v_n$ von V ergänzen. Die darstellende Matrix von f bezüglich dieser Basis hat die Form

$$\left(\begin{array}{c|c} \lambda_1 & * \\ \hline 0 & \tilde{A} \end{array} \right).$$

Dabei ist $\tilde{A}$ die darstellende Matrix des folgenden Endomorphismus:

Sei $W = \text{Lin}(v_2, \dots, v_n) \subset V$, $\dim(W) = n - 1$. Für $w \in W$ gilt

$$f(w) = \underbrace{\tilde{w}}_{\in W} + \underbrace{t(w)}_{\in K} \cdot v_1$$

mit eindeutig bestimmten $\tilde{w} \in W$ und $t(w) \in K$. Die Zuordnung

$$w \mapsto \tilde{w} =: g(w)$$

definiert einen Endomorphismus $g : W \rightarrow W$. Die darstellende Matrix von g bezüglich der Basis $v_2, \dots, v_n$ von W ist $\tilde{A}$ und es gilt $P_{\tilde{A}}(X) = (X - \lambda_2) \cdot \dots \cdot (X - \lambda_n)$. Per Induktion nach $n = \dim(V)$ können wir annehmen, dass $\tilde{A}$ eine obere Dreiecksmatrix ist. $\square$

Korollar 5.52. Über einem algebraisch abgeschlossenen Körper K ist jeder Endomorphismus eines endlichdimensionalen K -Vektorraums trigonalisierbar.

Bemerkung 5.53. Trigonalisierung ist nicht eindeutig.

Beispiel 5.54. Die obere Dreiecksmatrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ist ähnlich zur oberen Dreiecksmatrix

$$\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a^{-1} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$$

für $a \in K^\times$.

Frage. Kann man durch zusätzliche Bedingungen erreichen, dass die Transformation auf obere Dreiecksgestalt eindeutig wird?

Diese Frage behandeln wir später (im Kapitel zur Jordanschen Normalform).

6 Die Jordansche Normalform

Sei K ein Körper und V ein K -Vektorraum der Dimension $n < \infty$.

Definition 6.1.

- i) Ein Polynom $Q \in K[X] \setminus \{0\}$ *teilt* $P \in K[X]$, falls es ein $A \in K[X]$ gibt mit

$$P(X) = Q(X) \cdot A(X).$$

Schreibe $Q|P$.

- ii) $P, Q \in K[X]$ heißen *teilerfremd*, falls die konstanten Polynome ungleich Null die einzigen gemeinsamen Teiler von P und Q sind.

Satz 6.2. Seien $P, Q \in K[X]$ teilerfremd. Dann existieren $A, B \in K[X]$, so dass

$$AP + BQ = 1$$

gilt.

Beweis. Übung. □

Bemerkung 6.3. Sei K algebraisch abgeschlossen. Dann sind $P, Q \in K[X]$ genau dann teilerfremd, wenn sie keine gemeinsamen Nullstellen in K haben.

Definition 6.4. Sei $f : V \rightarrow V$ ein Endomorphismus. Ein Unterraum $W \subset V$ heißt *invariant*, falls $f(W) \subset W$. In diesem Fall kann man f auf W einschränken und erhält eine lineare Abbildung

$$f|_W : W \rightarrow W.$$

Bemerkung 6.5. Sei W ein invarianter Unterraum von V und $g = f|_W$. Dann ist P_g ein Teiler von P_f in $K[X]$.

Beweis. Übung (H6). □

6.1 Der Satz von Cayley-Hamilton

Sei $P = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0 \in K[X]$ ein Polynom und $f : V \rightarrow V$ ein Endomorphismus. Wir können f in P einsetzen, das heißt

$$P(f) = a_n f^n + a_{n-1} f^{n-1} + \dots + a_0 \operatorname{id}_V \in \operatorname{End}(V)$$

betrachten.

Hierbei schreiben wir auch $f^0 = \operatorname{id}_V$.

Bemerkung 6.6. Das Einsetzen ist mit den Ringoperationen verträglich, d.h. es gilt

$$(P + Q)(f) = P(f) + Q(f) \text{ und } (P \cdot Q)(f) = P(f) \circ Q(f)$$

für alle $P, Q \in K[X]$. D.h. einfach, dass die Abbildung $K[X] \rightarrow \text{End}(V)$ gegeben durch $P \mapsto P(f)$ ein Homomorphismus von Ringen ist. Ihr Bild $K[f] = \{P(f) \mid P \in K[X]\}$ ist ein kommutativer Unterring von $\text{End}(V)$.

Satz 6.7 (Cayley-Hamilton). *Sei $f : V \rightarrow V$ ein Endomorphismus und $P_f \in K[X]$ das charakteristische Polynom von f . Dann gilt*

$$P_f(f) = 0 \in \text{End}(V),$$

das heißt $P_f(f)(v) = 0$ für alle $v \in V$.

Beweis. Sei $v \in V$. Es ist zu zeigen, dass $P_f(f)(v) = 0$. Für $v = 0$ ist dies klar. Sei $v \neq 0$.

Schritt 1:

Sei $r \in \mathbb{N}$ minimal, so dass $v, f(v), f^2(v), \dots, f^r(v)$ linear abhängig sind. Dann gibt es $c_0, \dots, c_{r-1} \in K$, so dass $f^r(v) = c_0 v + c_1 f(v) + \dots + c_{r-1} f^{r-1}(v)$. Offenbar ist $r \leq n = \dim(V)$ und $v, f(v), f^2(v), \dots, f^{r-1}(v)$ sind linear unabhängig. Sei

$$U = \text{Lin}(v, f(v), f^2(v), \dots, f^{r-1}(v)) \subset V.$$

Offenbar ist $B = (v, f(v), f^2(v), \dots, f^{r-1}(v))$ eine Basis des Unterraums U und U ist f -invariant, das heißt

$$f|_U \in \text{End}(U).$$

Bezüglich der Basis B hat $f|_U$ die darstellende Matrix

$$A = \mathcal{M}_B^B(f|_U) = \begin{pmatrix} 0 & & c_0 \\ 1 & \ddots & c_1 \\ & \ddots & 0 & \vdots \\ & & 1 & c_{r-1} \end{pmatrix}.$$

Folglich ist das charakteristische Polynom von $f|_U$

$$\begin{aligned} P_{f|_U}(X) &= \det \begin{pmatrix} X & & -c_0 \\ -1 & \ddots & -c_1 \\ & \ddots & X & \vdots \\ & & -1 & X - c_{r-1} \end{pmatrix} \\ &= (X - c_{r-1}) \cdot X^{r-1} - c_{r-2} X^{r-2} - c_{r-3} X^{r-3} - \dots - c_0 \\ &= X^r - c_{r-1} X^{r-1} - c_{r-2} X^{r-2} - c_{r-3} X^{r-3} - \dots - c_0. \end{aligned}$$

Die Determinante wurde hierbei nach der letzten Spalte entwickelt. Es folgt

$$P_{f|_U}(f)(v) = f^r(v) - c_{r-1}f^{r-1}(v) - \dots - c_1f(v) - c_0v = 0.$$

Schritt 2:

Nach Bemerkung 6.5 gilt $P_{f|_U}|P_f$, das heißt es gibt ein Polynom $Q \in K[X]$ mit

$$P_f(X) = P_{f|_U}(X) \cdot Q(X).$$

Es folgt

$$P_f(f)(v) = \underbrace{P_{f|_U}(f)(v)}_{=0} \cdot Q(f)(v) = 0.$$

□

6.2 Zerlegung in Haupträume

Satz 6.8. *Sei $f : V \rightarrow V$ ein Endomorphismus und $P \in K[X]$ ein Polynom mit $P(f) = 0$ (zum Beispiel das charakteristische Polynom von f). Es sei $P = P_1 \cdot P_2 \in K[X]$ mit teilerfremden Faktoren $P_1, P_2 \in K[X]$. Dann gilt*

$$V = V_1 \oplus V_2$$

mit

$$V_i = \text{Kern}(P_i(f)).$$

Die Räume V_i sind f -invariant.

Beweis. Nach Satz 6.2 existieren $A_1, A_2 \in K[X]$ mit

$$A_1 \cdot P_1 + A_2 \cdot P_2 = 1.$$

Es folgt

$$A_1(f) \circ P_1(f) + A_2(f) \circ P_2(f) = \text{id}_V,$$

das heißt

$$\underbrace{A_1(f) \circ P_1(f)(v)}_{=:v_2} + \underbrace{A_2(f) \circ P_2(f)(v)}_{=:v_1} = v$$

für alle $v \in V$. Nach Voraussetzung gilt

$$P_2(f)(v_2) = P_2(f) \circ A_1(f) \circ P_1(f)(v) = A_1(f) \circ \underbrace{P(f)(v)}_{=0} = 0,$$

das heißt

$$v_2 \in \text{Kern}(P_2(f)).$$

Analog zeigt man

$$v_1 \in \text{Kern}(P_1(f)),$$

es folgt

$$V = V_1 + V_2.$$

Es bleibt noch $V_1 \cap V_2 = \{0\}$ zu zeigen. Für $v \in V_1 \cap V_2$ gilt

$$v = A_1(f) \circ \underbrace{P_1(f)(v)}_{=0} + A_2(f) \circ \underbrace{P_2(f)(v)}_{=0} = 0.$$

□

Bemerkung 6.9. Sei $f : V \rightarrow V$ ein nilpotenter Endomorphismus, das heißt $f^m = 0$ für ein $m \in \mathbb{N}$. Dann ist $P_f(X) = X^n$ mit $n = \dim(V)$.

Beweis.

$$\begin{aligned} f^m &= 0 \\ \Rightarrow \det(f)^m &= 0 \\ \Rightarrow \det(f) &= 0 \end{aligned}$$

Da f linear ist, gibt es daher ein nicht-triviales Element im Kern von f , das heißt es gibt einen Eigenvektor zum Eigenwert 0. Ergänze v zu einer Basis $B = (v, b_2, \dots, b_n)$ von V . Die darstellende Matrix ist

$$\mathcal{M}_B^B(f) = \left(\begin{array}{c|c} 0 & * \\ \hline 0 & A' \end{array} \right),$$

wobei $A' \in K^{(n-1) \times (n-1)}$ nilpotent ist. Es gilt

$$P_f(X) = X \cdot P_{A'}(X),$$

per Induktion nach $\dim(V)$ folgt die Behauptung. □

Definition 6.10. Sei $f : V \rightarrow V$ ein Endomorphismus, λ ein Eigenwert von f und $r \in \mathbb{N}$. Der r -te verallgemeinerte Eigenraum von f zum Eigenwert λ ist definiert als

$$\text{Eig}_r(f, \lambda) = \text{Kern}((f - \lambda \text{id}_V)^r).$$

Bemerkung 6.11.

i) Für $r = 1$ erhält man den gewöhnlichen Eigenraum,

$$\text{Eig}_1(f, \lambda) = \text{Eig}(f, \lambda).$$

ii) $\text{Eig}_r(f, \lambda)$ ist f -invariant.

iii) Es gilt $\text{Eig}_1(f, \lambda) \subset \text{Eig}_2(f, \lambda) \subset \text{Eig}_3(f, \lambda) \subset \dots$

iv) Da V endlichdimensional ist, existiert ein $N \in \mathbb{N}$ mit

$$\text{Eig}_N(f, \lambda) = \text{Eig}_{N+1}(f, \lambda) = \dots,$$

das heißt, dass die Inklusionskette stationär wird. Dann heißt

$$\text{Hau}(f, \lambda) := \text{Eig}_N(f, \lambda) = \bigcup_{r=1}^{\infty} \text{Eig}_r(f, \lambda)$$

Hauptraum von f zum Eigenwert λ .

Bemerkung 6.12. Sei λ ein Eigenwert von f mit algebraischer Vielfachheit m .

- i) Dann gilt $\dim(\text{Eig}_r(f, \lambda)) \leq m$.
- ii) Die Einschränkung g von f auf $\text{Eig}_r(f, \lambda)$ hat das charakteristische Polynom $P_g(X) = (X - \lambda)^\mu$ mit $\mu = \dim(\text{Eig}_r(f, \lambda))$.

Beweis. ii): Sei $W = \text{Eig}_r(f, \lambda)$, $g : W \rightarrow W$ die Einschränkung von f auf W . Nach der Definition von W gilt

$$(g - \lambda \text{id}_W)^r = 0,$$

das heißt $g - \lambda \text{id}_W \in \text{End}(W)$ ist nilpotent. Nach Bemerkung 6.9 gilt $P_{g-\lambda \text{id}_W}(X) = X^\mu$ mit $\mu = \dim(W)$. Es folgt

$$\begin{aligned} P_g(X) &= \det(X \text{id}_W - g) \\ &= \det((X - \lambda) \text{id}_W - (g - \lambda \text{id}_W)) \\ &= P_{g-\lambda \text{id}_W}(X - \lambda) \\ &= (X - \lambda)^\mu. \end{aligned}$$

i): $W = \text{Eig}_r(f, \lambda)$ ist nach Bemerkung 6.11 f -invariant. Nach Bemerkung 6.5 ist P_g ein Teiler von P_f . Daraus folgt $\mu \leq m$. $\square$

Satz 6.13. Das charakteristische Polynom P_f von f zerfalle in Linearfaktoren,

$$P_f = (X - \lambda_1)^{m_1} \cdot \dots \cdot (X - \lambda_k)^{m_k}$$

mit $\lambda_i \neq \lambda_j$ für $i \neq j$. Dann gilt

$$V = \text{Eig}_{m_1}(f, \lambda_1) \oplus \dots \oplus \text{Eig}_{m_k}(f, \lambda_k).$$

Durch Einschränken von f erhält man Endomorphismen

$$f_i : \text{Eig}_{m_i}(f, \lambda_i) \rightarrow \text{Eig}_{m_i}(f, \lambda_i)$$

und es gilt

$$P_{f_i}(X) = (X - \lambda_i)^{m_i} \in K[X].$$

Beweis. Induktion nach k :

$k = 1$: Hier ist $P_f(X) = (X - \lambda_1)^{m_1} = (X - \lambda_1)^n$. Nach dem Satz von Cayley-Hamilton (Satz 6.7) gilt $P_f(f) = 0 \in \text{End}(V)$. Dies bedeutet gerade, dass $(f - \lambda_1)^n = 0$, woraus $\text{Eig}_n(f, \lambda_1) = V$ folgt.

Sei nun $k > 1$: Wir faktorisieren $P_f(X) = P_1(X) \cdot P_2(X)$ mit

$$\begin{aligned} P_1(X) &= (X - \lambda_1)^{m_1}, \\ P_2(X) &= (X - \lambda_2)^{m_2} \cdot \dots \cdot (X - \lambda_k)^{m_k}. \end{aligned}$$

Die Polynome P_1 und P_2 sind teilerfremd und nach Satz 6.8 gilt

$$V = V_1 \oplus V_2$$

mit

$$\begin{aligned} V_1 &= \text{Kern}(P_1(f)) = \text{Kern}((f - \lambda_1 \text{id}_V)^{m_1}) = \text{Eig}_{m_1}(f, \lambda_1), \\ V_2 &= \text{Kern}(P_2(f)). \end{aligned}$$

Sei $f_1 = f|_{V_1} : V_1 \rightarrow V_1$. Nach Bemerkung 6.12 gilt

$$P_{f_1}(X) = (X - \lambda_1)^\mu$$

mit $\mu \leq m_1$. Da V_1 und V_2 f -invariant sind mit $V = V_1 \oplus V_2$, gilt $P_f(X) = P_{f_1} \cdot P_{f|_{V_2}}$.

Behauptung: $\mu = m_1$.

Angenommen die Behauptung gilt nicht, das heißt $\mu < m_1$. Dann gilt $(X - \lambda_1)|_{P_{f|_{V_2}}}$, das heißt λ_1 ist eine Nullstelle von $P_{f|_{V_2}}$. Daraus folgt, dass λ_1 ein Eigenwert von $f|_{V_2} : V_2 \rightarrow V_2$ ist. Es existiert daher ein Vektor $0 \neq v_2 \in V_2$ mit $f(v_2) = \lambda_1 v_2$. Weiterhin gilt

$$v_2 \in \text{Kern}(f - \lambda_1 \text{id}_V) \subset \text{Kern}((f - \lambda_1 \text{id}_V)^{m_1}) = V_1.$$

Es folgt

$$0 \neq v \in V_1 \cap V_2,$$

was im Widerspruch zu $V = V_1 \oplus V_2$ steht. Es folgt die Behauptung.

Da $\mu = m_1$ folgt

$$\begin{aligned} P_{f_1}(X) &= (X - \lambda_1)^\mu = (X - \lambda_1)^{m_1} = P_1, \\ P_{f|_{V_2}}(X) &= P_2(X). \end{aligned}$$

Per Induktion können wir annehmen, dass die Behauptung für V_2 und $f|_{V_2}$ bereits bewiesen ist. Es folgt die Behauptung für V und f . $\square$

Bemerkung 6.14. P_f zerfalle in Linearfaktoren. Sei λ ein Eigenwert von f mit algebraischer Vielfachheit m . Dann gilt für alle $r \geq m$

$$\dim(\text{Eig}_r(f, \lambda)) = m.$$

Insbesondere gilt, dass $\text{Hau}(f, \lambda) = \text{Eig}_m(f, \lambda)$.

Beweis. Folgt aus Bemerkung 6.12 (i) und Satz 6.13. $\square$

Satz 6.15. Sei $f : V \rightarrow V$ ein Endomorphismus und P_f zerfalle in Linearfaktoren. Es sind äquivalent:

i) f ist diagonalisierbar.

ii) Für alle Eigenwerte λ von f gilt $\text{Eig}(f, \lambda) = \text{Hau}(f, \lambda)$.

Beweis. (ii) $\Rightarrow$ (i): Nach Satz 6.13 und der Voraussetzung gilt

$$V = \text{Hau}(f, \lambda_1) \oplus \dots \oplus \text{Hau}(f, \lambda_k) = \text{Eig}(f, \lambda_1) \oplus \dots \oplus \text{Eig}(f, \lambda_k).$$

Dann ist f bereits diagonalisierbar.

(i) $\Rightarrow$ (ii): Da f diagonalisierbar ist, gilt nach Satz 5.49 und Bemerkung 6.14

$$\dim(\text{Eig}(f, \lambda)) = \mu(P_f, \lambda) = \dim(\text{Hau}(f, \lambda)).$$

Da $\text{Eig}(f, \lambda) \subset \text{Hau}(f, \lambda)$ folgt daraus $\text{Eig}(f, \lambda) = \text{Hau}(f, \lambda)$. $\square$

Korollar 6.16. Sei $f : V \rightarrow V$ ein diagonalisierbarer Endomorphismus und $W \subset V$ ein invarianter Unterraum. Dann ist auch $f|_W : W \rightarrow W$ diagonalisierbar.

Beweis. Sei $g = f|_W \in \text{End}(W)$. Da f diagonalisierbar ist zerfällt P_f über K in Linearfaktoren. Da dann P_g ein Teiler von P_f ist, muss auch P_g über K in Linearfaktoren zerfallen.

Wir zeigen nun, dass ii) in 6.13 für g erfüllt ist. Wir haben die Inklusionen

$$\text{Eig}(g, \lambda) \subset \text{Hau}(g, \lambda) \subset \text{Hau}(f, \lambda) = \text{Eig}(f, \lambda).$$

Sei $w \in \text{Hau}(g, \lambda)$. Da $\text{Hau}(f, \lambda) = \text{Eig}(f, \lambda)$ gilt, ist $g(w) = f(w) = \lambda w$ für alle $w \in \text{Hau}(g, \lambda)$ und damit bereits $\text{Hau}(g, \lambda) \subset \text{Eig}(g, \lambda)$ womit die Gleichheit $\text{Hau}(g, \lambda) = \text{Eig}(g, \lambda)$ gezeigt ist. $\square$

Satz 6.17. Seien $f_1, \dots, f_m \in \text{End}(V)$ diagonalisierbar und es gelte $f_i \circ f_j = f_j \circ f_i$ für alle i, j . Dann können $f_1, \dots, f_m$ simultan diagonalisiert werden, das heißt es existiert eine Basis $v_1, \dots, v_n$ von V , so dass jedes v_i ein Eigenvektor aller f_j (für $i = 1, \dots, n$, und $j = 1, \dots, m$) ist.

Beweis. Hier beweisen wir nur den Fall $m = 2$. Sei $f := f_1$, $g := f_2$. Seien $\lambda_1, \dots, \lambda_k$ die Eigenwerte von f und

$$V_i := \text{Eig}(f, \lambda_i).$$

Es ist $f|_{V_i} = \lambda_i \text{id}_{V_i}$.

Behauptung: V_i ist g -invariant.

Sei $v \in V_i$ (das heißt $f(v) = \lambda_i v$). Es gilt

$$f(g(v)) = g(f(v)) = g(\lambda_i v) = \lambda_i g(v).$$

Dies bedeutet gerade, dass $g(v) \in \text{Eig}(f, \lambda_i) = V_i$, der Raum V_i ist also g -invariant.

Mit Korollar 6.16 folgt, dass $g|_{V_i}$ für alle $i = 1, \dots, k$ diagonalisierbar ist. Man erhält eine Basis von Eigenvektoren von g , die automatisch auch Eigenvektoren von f sind. $\square$

Korollar 6.18. *Sind $f, g \in \text{End}(V)$ diagonalisierbar und vertauschbar, so sind auch $f \circ g$ und $f \pm g$ diagonalisierbar.*

Beweis. Übung. $\square$

6.3 Die Jordan-Zerlegung

Sei $f : V \rightarrow V$ ein Endomorphismus des K -Vektorraums V . Das charakteristische Polynom $P_f \in K[X]$ zerfalle in Linearfaktoren

$$P_f(X) = (X - \lambda_1)^{m_1} \cdot \dots \cdot (X - \lambda_k)^{m_k}$$

mit $\lambda_i \neq \lambda_k$ für $i \neq j$. Nach Satz 6.13 gilt

$$V = V_1 \oplus \dots \oplus V_k$$

mit

$$V_i = \text{Hau}(f, \lambda_i) = \text{Eig}_{m_i}(f, \lambda_i).$$

Es sei

$$p_i : V \rightarrow V$$

der *Projektionsoperator* auf V_i , das heißt die eindeutig bestimmte lineare Abbildung mit

$$p_i(v) = \begin{cases} v, & \text{falls } v \in V_i, \\ 0, & \text{falls } v \in V_j, j \neq i. \end{cases}$$

Für $v = v_1 + \dots + v_k$ mit $v_i \in V_i$ ist also

$$p_i(v) = v_i.$$

Die p_i sind offensichtlich diagonalisierbar und kommutieren miteinander.

Bemerkung 6.19. Der Projektionsoperator p_i lässt sich als Polynom in f mit Koeffizienten in K darstellen.

Beweis. Sei ohne Beschränkung der Allgemeinheit $i = 1$. Wir faktorisieren $P_f = P_1 \cdot P_2$ mit

$$\begin{aligned} P_1(X) &= (X - \lambda_1)^{m_1}, \\ P_2(X) &= (X - \lambda_2)^{m_2} \cdot \dots \cdot (X - \lambda_k)^{m_k}. \end{aligned}$$

Die Polynome P_1 und P_2 sind teilerfremd. Nach Satz 6.2 existieren $Q_1, Q_2 \in K[X]$ mit

$$Q_1 \cdot P_1 + Q_2 \cdot P_2 = 1 \in K[X] \quad (*)$$

Nach Satz 6.13 gilt $V_1 = \text{Kern}(P_1(f))$. Es folgt $p_1 = (Q_2 \cdot P_2)(f)$, denn für $v \in V_1$ gilt

$$\begin{aligned} (Q_2 \cdot P_2)(f)(v) &\stackrel{(*)}{=} v - (Q_1 \cdot P_1)(f)(v) \\ &= v - Q_1(f) \circ \underbrace{P_1(f)(v)}_{=0} \\ &= v \end{aligned}$$

und für $v \in V_i$ mit $i = 2, \dots, k$ gilt nach Satz 6.13

$$v \in V_i = \text{Kern}((f - \lambda_i \text{id}_V)^{m_i}) \subset \text{Kern}(P_2(f)),$$

das heißt $P_2(f)(v) = 0$. □

Satz 6.20 (Jordan-Zerlegung). *Sei $f \in \text{End}(V)$ und $P_f \in K[X]$ zerfalle in Linearfaktoren in $K[X]$. Dann existiert eine Zerlegung*

$$f = g + h$$

mit $g, h \in \text{End}(V)$ wobei g diagonalisierbar, h nilpotent und $g \circ h = h \circ g$. Eine solche Zerlegung ist eindeutig und g, h sind Polynome in f mit Koeffizienten in K .

Beweis.

Existenz: Mit den Eigenwerten $\lambda_1, \dots, \lambda_k$ und den Projektionsoperatoren $p_1, \dots, p_k$ definieren wir

$$g = \lambda_1 p_1 + \dots + \lambda_k p_k \in \text{End}(V).$$

Per Konstruktion ist g diagonalisierbar und nach Bemerkung 6.19 ist g ein Polynom in f mit Koeffizienten in K . Sei

$$h = f - g.$$

Dann ist auch h ein Polynom in f mit Koeffizienten in K .

Es ist noch zu zeigen, dass h nilpotent ist. Da V_i für alle $i = 1, \dots, k$ ein h -invarianter Raum ist, genügt es zu zeigen, dass

$$h|_{V_i} \in \text{End}(V_i)$$

nilpotent ist. Es gilt

$$h|_{V_i} = f|_{V_i} - \lambda_i \text{id}_{V_i} = (f - \lambda_i \text{id}_V)|_{V_i}.$$

Wegen $V_i = \text{Kern}((f - \lambda_i \text{id}_V)^{m_i})$ gilt $(h|_{V_i})^{m_i} = 0$, das heißt $h|_{V_i}$ ist nilpotent.

Eindeutigkeit: Sei $f = g + h$ eine Zerlegung von f mit g diagonalisierbar, h nilpotent und g, h seien Polynome in f mit Koeffizienten in K . Sei $f = g' + h'$ eine zweite Zerlegung mit g' diagonalisierbar, h' nilpotent und g', h' vertauschbar (nicht notwendigerweise Polynome in f). Da g' und h' vertauschbar sind, vertauschen sie auch mit $f = g' + h'$. Da g und h Polynome in f sind, vertauschen g' und h' dann auch mit g und h . Aus $g + h = f = g' + h'$ folgt

$$g - g' = h' - h.$$

Die Endomorphismen g, g' sind diagonalisierbar und vertauschen. Nach Korollar 6.18 ist daher auch $g - g'$ diagonalisierbar. Die Endomorphismen h, h' sind nilpotent und vertauschbar. Dann ist auch $h' - h$ nilpotent, denn in der Summe

$$(h' - h)^N = \sum_{i=0}^N \binom{N}{i} (-1)^{N-i} (h')^i \circ h^{N-i}$$

ist, für N groß genug, in jedem Summanden $(h')^i$ oder h^{N-i} gleich 0. Damit ist $g - g'$ diagonalisierbar und nilpotent, das heißt alle Eigenwerte von $g - g'$ sind 0. Es folgt $g = g'$ und $h = h'$. $\square$

Korollar 6.21. Sei $A \in K^{n \times n}$. Das charakteristische Polynom $P_A \in K[X]$ zerfalle in Linearfaktoren

$$P_A(X) = (X - \lambda_1)^{m_1} \cdot \dots \cdot (X - \lambda_k)^{m_k},$$

wobei $\lambda_i \neq \lambda_j$ für $i \neq j$. Dann existiert eine Matrix $S \in \text{GL}_n(K)$, so dass

$$SAS^{-1} = \begin{pmatrix} C_1 & & \\ & \ddots & \\ & & C_k \end{pmatrix},$$

wobei die Blöcke $C_i \in K^{m_i \times m_i}$ obere Dreiecksmatrizen der folgenden Gestalt sind:

$$\begin{pmatrix} \lambda_i & * & * \\ & \ddots & * \\ & & \lambda_i \end{pmatrix}.$$

Beweis. Übung. $\square$

6.4 Zyklische Vektoren

Zyklische Vektoren sind das „Gegenteil“ von Eigenvektoren.

Sei V ein endlichdimensionaler K -Vektorraum und $f : V \rightarrow V$ ein Endomorphismus.

Definition 6.22. Ein Vektor $a \in V$ heißt *zyklisch* bezüglich f , falls

$$V = \text{Lin}(a, f(a), f^2(a), \dots).$$

Falls V einen zyklischen Vektor besitzt, so nennt man (V, f) *zyklisch*.

Beispiel 6.23. Für $a \in V$ ist

$$V(a) := \{P(f)(a) \mid P \in K[X]\}$$

der von a erzeugte *zyklische Untervektorraum*. Für $a_1, \dots, a_m \in V$ ist

$$V(a_1, \dots, a_m) := V(a_1) + \dots + V(a_m)$$

der kleinste Untervektorraum von V der $V(a_1), \dots, V(a_m)$ enthält. Für geeignete Vektoren $a_1, \dots, a_m \in V$ (zum Beispiel eine Basis) gilt $V = V(a_1, \dots, a_m)$.

Satz 6.24. *Es existiert eine Zerlegung*

$$V = V_1 \oplus \dots \oplus V_m$$

so dass für alle $i = 1, \dots, m$ gilt:

i) V_i ist f -invariant, so dass $f_i := f|_{V_i} \in \text{End}(V_i)$,

ii) (V_i, f_i) ist *zyklisch*.

Beweis. Durch Induktion nach $\dim(V)$. Falls $\dim(V) = 1$, ist die Aussage klar.

Sei nun $\dim(V) > 1$ und die Behauptung bewiesen für Vektorräume kleinerer Dimension.

Schritt 1: Es sei $m \in \mathbb{N}$ die minimale Anzahl von Erzeugern von V als $K[f]$ -Modul. Das heißt, m sei minimal, so dass es $a_1, \dots, a_m \in V$ mit

$$V = V(a_1, \dots, a_m)$$

gibt. Im Allgemeinen gibt es dann viele solche m -Tupel und die Summenzerlegung

$$V = V(a_1) + \dots + V(a_m)$$

ist im Allgemeinen nicht direkt. Wir zeigen nun, dass die Summe für eine geeignete Wahl von $a_1, \dots, a_m \in V$ direkt ist.

Schritt 2: Eine *Relation* zu $(a_1, \dots, a_m)$ ist ein m -Tupel $(R_1, \dots, R_m)$ von Polynomen in $K[X]$ so dass

$$R_1(f)(a_1) + \dots + R_m(f)(a_m) = 0$$

gilt. Man kann zum Beispiel $R_1 = \dots = R_m = P_f$ wählen. Ist $(R_1, \dots, R_m)$ eine nicht-triviale Relation, so heißt

$$\min\{\text{Grad}(R_i) \mid i = 1, \dots, m, R_i \neq 0\} \in \mathbb{N}_0$$

der *Grad der Relation*. Es sei $d \in \mathbb{N}_0$ minimal mit der Eigenschaft:

Es gibt $a_1, \dots, a_m \in V$ mit $V = V(a_1, \dots, a_m)$ und
eine nicht-triviale Relation $(R_1, \dots, R_m)$ zu $(a_1, \dots, a_m)$ vom Grad d .

Sei $(R_1, \dots, R_m)$ eine solche Relation mit minimalem Grad d und ohne Beschränkung der Allgemeinheit $R_1 \neq 0$ und $\text{Grad}(R_1) = d$.

Behauptung: Bei geeigneter Wahl von $(a_1, \dots, a_m)$ können wir sogar erreichen, dass

$$R_2 = \dots = R_m = 0$$

(und $R_1 \neq 0$, $\text{Grad}(R_1) = d$).

Denn: Falls $R_2 \neq 0$, so wähle $A, B \in K[X]$ mit

$$R_2 = AR_1 + B$$

und $\text{Grad}(B) < \text{Grad}(R_1)$. Setze

$$\begin{aligned} b_1 &= a_1 + A(f)(a_2), \\ b_i &= a_i \end{aligned}$$

für $i = 2, \dots, m$ und

$$\begin{aligned} S_1 &= R_1, \\ S_2 &= R_2 - AR_1 = B, \\ S_i &= R_i \end{aligned}$$

für $i = 3, \dots, m$. Dann gilt $V = V(b_1, \dots, b_m)$ und

$$S_1(f)(b_1) + \dots + S_m(f)(b_m) = 0,$$

das heißt $(S_1, \dots, S_m)$ ist eine nicht-triviale Relation zu $(b_1, \dots, b_m)$. Weiterhin gilt dann $B = 0$, denn sonst wäre $(S_1, \dots, S_m)$ eine Relation mit Grad kleiner als d , was der Minimalität von d widerspräche. Auf diese Weise erreicht man $R_2 = 0$ und analog $R_3 = \dots = R_m = 0$.

Schritt 3: Sei $(R_1, \dots, R_m)$ eine Relation mit der Form wie in der obigen Behauptung. Es ist also insbesondere $R_1(f)(a_1) = 0$.

Behauptung: Es gilt $V(a_1) \cap V(a_2, \dots, a_m) = \{0\}$.

Sei dazu $P_1 \in K[X]$ mit

$$P_1(f)(a_1) \in V(a_2, \dots, a_m).$$

Dann existieren $P_2, \dots, P_m \in K[X]$ mit

$$P_1(f)(a_1) + \dots + P_m(f)(a_m) = 0.$$

Wie in Schritt 2 ziehen wir ein Vielfaches der Relation $R_1(f)(a_1) = 0$ ab, nutzen Division mit Rest und die Minimalität der Relation $R_1(f)(a_1) = 0$, um zu zeigen, dass P_1 durch R_1 teilbar ist. Daraus folgt $P_1(f)(a_1) = 0$ und damit die Behauptung.

Schritt 4: Wir haben nun

$$V = V(a_1) \oplus V(a_2, \dots, a_m).$$

Da $\dim V(a_2, \dots, a_m) < \dim(V)$, folgt die Behauptung per Induktion. $\square$

Wir betrachten nun die die Zerlegung in zyklische Untervektorräume speziell für nilpotente Endomorphismen.

Sei V ein K -Vektorraum mit $n = \dim(V) < \infty$.

Bemerkung 6.25. Ein Endomorphismus ist genau dann nilpotent, wenn $P_f(X) = X^n$.

Beweis. „ $\Rightarrow$ “: Bemerkung 6.9.

„ $\Leftarrow$ “: Sei $P_f(X) = X^n$. Nach dem Satz von Cayley-Hamilton (Satz 6.7) gilt

$$f^n = P_f(f) = 0,$$

das heißt f ist nilpotent. $\square$

Satz 6.26. Sei $f : V \rightarrow V$ ein nilpotenter Endomorphismus und (V, f) sei zyklisch. Dann existiert eine Basis B von V , so dass die darstellende Matrix $\mathcal{M}_B^B(f)$ die Form

$$\begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$

hat.

Beweis. Sei $a \in V$ ein zyklischer Vektor. Es ist $P_f(X) = X^n$ mit $n = \dim(V)$ (Bemerkung 6.25). Nach dem Satz von Cayley-Hamilton gilt $f^n = 0$ und es folgt $f^r(a) = 0$ für alle $r \geq n$. Damit ist $f^{n-1}(a), f^{n-2}(a), \dots, f(a), a$ ein Erzeugendensystem des Vektorraums V . Da dies genau $n = \dim(V)$ Vektoren sind, bilden diese eine Basis. Bezüglich dieser hat die darstellende Matrix die geforderte Form. $\square$

Satz 6.27. Sei $f : V \rightarrow V$ ein nilpotenter Endomorphismus und sei $V = V_1 \oplus \dots \oplus V_m$ eine Zerlegung von V in f -invariante zyklische Unterräume, so dass

$$\dim(V_1) \geq \dim(V_2) \geq \dots \geq \dim(V_m).$$

Dann sind m und das Tupel $(\dim(V_1), \dots, \dim(V_m))$ eindeutig durch f bestimmt.

Beweis. Für den Beweis bezeichnen wir für $a \in V$ die Größe

$$h(a) := \dim(V(a)) \in \mathbb{N}_0$$

als die Höhe von a . Nach Satz 6.26 ist $h(a) = \min\{r \in \mathbb{N}_0 \mid f^r(a) = 0\}$. Ist $a = a_1 + \dots + a_m$ mit $a_i \in V_i$, so gilt

$$h(a_i) \leq h(a)$$

für alle $i = 1, \dots, m$ und es existiert ein i mit $h(a_i) = h(a)$.

Sei $a \in V$ mit $V(a) = V_1$. Wegen $\dim(V_1) \geq \dim(V_i)$ ist a ein Element maximaler Höhe in V . Sei $V = W_1 \oplus \dots \oplus W_k$ eine zweite Zerlegung in f -invariante zyklische Unterräume mit

$$\dim(W_1) \geq \dim(W_2) \geq \dots \geq \dim(W_k).$$

Wir schreiben

$$a = w_1 + \dots + w_k$$

mit $w_i \in W_i$. Es existiert ein i mit $h(w_i) = h(a)$. Weil a maximale Höhe hat gilt dann

$$h(a) = h(w_i) \leq \dim(W_i) \leq \dim(W_{i-1}) \leq \dots \leq \dim(W_1) \leq h(a).$$

Es folgt die Gleichheit

$$\dim(W_i) = \dim(W_{i-1}) = \dots = \dim(W_1) = h(a) = h(w_i).$$

Sei ohne Beschränkung der Allgemeinheit $i = 1$, das heißt

$$h(w_1) = h(a) = \dim(W_1) = \dim(V_1).$$

Es gilt

$$w_1 = a - w_2 - \dots - w_k \in V_1 + (W_2 \oplus \dots \oplus W_k).$$

Für alle $r \in \mathbb{N}_0$ gilt dann

$$f^r(w_1) \in V_1 + (W_2 \oplus \dots \oplus W_k).$$

Dies bedeutet

$$W_1 \subset V_1 + (W_2 \oplus \dots \oplus W_k),$$

woraus

$$V = V_1 + (W_2 \oplus \dots \oplus W_k)$$

folgt. Aus Dimensionsgründen ist die Summe direkt, also

$$V_1 \cap (W_2 \oplus \dots \oplus W_k) = \{0\}. \quad (*)$$

Es sei $p : V \rightarrow V_2 \oplus \dots \oplus V_m$ die Projektion auf den zweiten Summanden der Zerlegung $V = V_1 \oplus (V_2 \oplus \dots \oplus V_m)$. Insbesondere ist $\text{Kern}(p) = V_1$. Die Einschränkung $\sigma = p|_{W_2 \oplus \dots \oplus W_k}$ definiert eine lineare Abbildung

$$\sigma : W_2 \oplus \dots \oplus W_k \rightarrow V_2 \oplus \dots \oplus V_m.$$

Wegen (*) ist $\text{Kern}(\sigma) = \{0\}$, das heißt σ ist injektiv. Aus Dimensionsgründen ist σ dann auch surjektiv, also ein Isomorphismus. Wir erhalten die zwei Zerlegungen

$$V_2 \oplus \dots \oplus V_m = \sigma(W_2) \oplus \dots \oplus \sigma(W_k)$$

in $f|_{V_2 \oplus \dots \oplus V_m}$ -invariante zyklische Unterräume. Per Induktion nach $\dim(V)$ folgt $k = m$ und $\dim(V_2) = \dim(W_2), \dots, \dim(V_m) = \dim(W_m)$. $\square$

6.5 Die Jordansche Normalform

Erinnerung: Eine $r \times r$ Matrix J heißt Jordanblock zum Eigenwert $\lambda \in K$, falls sie von der Form

$$J = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

ist. Eine Matrix aus $K^{n \times n}$ ist in Jordanscher Normalform, falls sie eine Block-Diagonalmatrix der Form

$$\begin{pmatrix} J_1 & & 0 \\ & J_2 & \\ & & \ddots \\ 0 & & & J_m \end{pmatrix}$$

ist, wobei die J_i Jordanblöcke (möglicherweise unterschiedlicher Größe) sind.

Beispiel 6.28 (Jordanblöcke).

$$\begin{aligned} r = 1 : & \quad (\lambda) \\ r = 2 : & \quad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \\ r = 3 : & \quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \\ \vdots & \quad \vdots \end{aligned}$$

Satz 6.29 (Jordansche Normalform). *Sei $f : V \rightarrow V$ ein Endomorphismus. Wenn $P_f \in K[X]$ vollständig in Linearfaktoren zerfällt, so gibt es eine Basis B von V , so dass $\mathcal{M}_B^B(f)$ in Jordanscher Normalform ist. Die Matrix ist dann durch f (bis auf die Reihenfolge der Jordanblöcke) eindeutig bestimmt.*

Beweis. Existenz: Nach Satz 6.13 können wir V zunächst in die Haupträume von f zerlegen:

$$V = \text{Hau}(f, \lambda_1) \oplus \dots \oplus \text{Hau}(f, \lambda_k).$$

Es genügt daher, die einzelnen Haupträume zu betrachten. Ohne Beschränkung der Allgemeinheit hat f dann nur einen Eigenwert λ und $f - \lambda \text{id}$ ist nilpotent. Es genügt, $f - \lambda \text{id}$ zu betrachten. Ohne Beschränkung der Allgemeinheit ist f dann nilpotent. Nach Satz 6.24 hat V eine Zerlegung in f -invariante zyklische Teilräume

$$V = V_1 \oplus \dots \oplus V_m,$$

ohne Beschränkung der Allgemeinheit sei daher f nilpotent und (V, f) zyklisch. In diesem Fall folgt die Behauptung aus Satz 6.26.

Eindeutigkeit: Es sei eine Basis B gegeben, so dass $\mathcal{M}_B^B(f)$ Jordansche Normalform hat. Fasst man alle Jordanblöcke zu einem festen Eigenwert zusammen, so erhält man die darstellende Matrix von f eingeschränkt auf den zugehörigen Hauptraum $\text{Hau}(f, \lambda)$. Ohne Beschränkung der Allgemeinheit habe f nur einen Eigenwert λ . Wieder betrachten wir

$$\mathcal{M}_B^B(f - \lambda \text{id}) = \mathcal{M}_B^B(f) - \lambda E,$$

ohne Beschränkung der Allgemeinheit ist f somit nilpotent. Die Jordanblöcke entsprechen f -invarianten zyklischen Unterräumen. Die Eindeutigkeit folgt nun aus Satz 6.27. $\square$

Korollar 6.30. *Sei $A \in K^{n \times n}$ und $P_A \in K[X]$ zerfalle in Linearfaktoren. Dann existiert $S \in \text{GL}(n, K)$, so dass*

$$SAS^{-1}$$

in Jordanscher Normalform ist. Die Normalformmatrix ist bis auf die Reihenfolge der Jordanblöcke eindeutig bestimmt.

Bestimmung der Jordanschen Normalform eines Endomorphismus $f : V \rightarrow V$:

1. Berechne das charakteristische Polynom und faktorisiere es in Linearfaktoren

$$P_f(X) = (X - \lambda_1)^{n_1} \cdot \dots \cdot (X - \lambda_k)^{n_k}.$$

2. Bestimme die Haupträume

$$\text{Hau}(f, \lambda_i) = \text{Eig}_{n_i}(f, \lambda_i) = \text{Kern}((f - \lambda_i \text{id}_V)^{n_i}).$$

3. Es genügt die einzelnen Haupträume zu betrachten. Ohne Beschränkung der Allgemeinheit hat f nur einen Eigenwert λ und

$$P_f(X) = (X - \lambda)^n.$$

4. Es genügt $f - \lambda \text{id}$ zu betrachten. Ohne Beschränkung der Allgemeinheit ist f nilpotent und

$$P_f(X) = X^n.$$

Es existiert eine Basis B von V , so dass

$$A = \mathcal{M}_B^B(f) = \begin{pmatrix} J(d_1) & & \\ & \ddots & \\ & & J(d_m) \end{pmatrix} \in K^{n \times n},$$

$$J(d) = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \in K^{d \times d}.$$

Dabei ist m Anzahl der Jordanblöcke und

$$\text{Rang}(A) = n - m.$$

Es gilt

$$\text{Rang}(J(d)^k) = \begin{cases} d - k, & k \leq d, \\ 0, & k > d. \end{cases}$$

Sei $\beta_d = \#\{i \mid d_i = d\}$ die Anzahl der Jordanblöcke der Länge d . Es folgt

$$\text{Rang}(A^k) = \text{Rang}(f^k) = \sum_{i=1}^m \text{Rang}(J(d_i)^k) = \sum_{d > k} \beta_d \cdot (d - k).$$

Damit kann man die Anzahl β_d rekursiv berechnen.

i) Angenommen $A^k = 0$ (zum Beispiel $k = n$). Dann gilt

$$0 = \text{Rang}(A^k) = \sum_{d > k} \underbrace{\beta_d}_{\geq 0} \cdot \underbrace{(d - k)}_{> 0},$$

woraus $\beta_d = 0$ für alle $d > k$ folgt.

ii) Es gilt

$$\text{Rang}(A^{k-1}) = \sum_{d > k-1} \beta_d \cdot (d - k + 1) = \beta_k.$$

iii) Es gilt

$$\text{Rang}(A^{k-2}) = \sum_{d > k-2} \beta_d \cdot (d - k + 2) = \beta_k \cdot 2 + \beta_{k-1},$$

woraus

$$\beta_{k-1} = \text{Rang}(A^{k-2}) - 2\beta_k$$

folgt.

...

Allgemein gilt

$$\beta_j = \text{Rang}(f^{j-1}) - 2\beta_{j+1} - 3\beta_{j+2} - \dots$$

Bestimmung einer Jordanbasis:

Ohne Beschränkung der Allgemeinheit ist f ein nilpotenter Endomorphismus, das heißt $P_f(X) = X^n$. Es sei $d \in \mathbb{N}$ mit $f^d = 0 \neq f^{d-1}$. Dann gibt es keine Jordanblöcke der Länge größer d und $\beta_d = \text{Rang}(f^{d-1})$ Jordanblöcke der Länge d . Seien $v_1, \dots, v_{\beta_d}$ linear unabhängig, so dass

$$V = (Kv_1 \oplus \dots \oplus Kv_{\beta_d}) \oplus \text{Kern}(f^{d-1}).$$

Setze

$$\begin{aligned} M_d &= \{v_1, \dots, v_{\beta_d}\}, \\ B_d &= M_d \cup f(M_d) \cup \dots \cup f^{d-1}(M_d). \end{aligned}$$

Dann ist B_d linear unabhängig und $B_d \cup \{0\}$ ist f -invariant. Der Untervektorraum

$$V_d := \text{Lin}(B_d)$$

ist daher f -invariant und $f|_{V_d} : V_d \rightarrow V_d$ hat bezüglich der Basis

$$(f^{d-1}(v_1), f^{d-2}(v_1), \dots, v_1, f^{d-1}(v_2), \dots, v_2, \dots, f^{d-1}(v_{\beta_d}), \dots, v_{\beta_d})$$

von V_d die darstellende Matrix

$$\begin{pmatrix} J(d) & & \\ & \ddots & \\ & & J(d) \end{pmatrix}$$

mit β_d Jordanblöcken. Betrachte nun ein Komplement zu $V_d + \text{Kern}(f^{d-2})$ in $\text{Kern}(f^{d-1})$. Die Dimension ist gerade β_{d-1} . Sei M_{d-1} eine Basis des Komplements. Setze

$$B_{d-1} := M_{d-1} \cup f(M_{d-1}) \cup \dots \cup f^{d-2}(M_{d-1}).$$

B_{d-1} ist linear unabhängig und $B_{d-1} \cup \{0\}$ ist f -invariant. Der Untervektorraum

$$V_{d-1} := \text{Lin}(B_{d-1})$$

ist daher f -invariant. Nach Umnummerierung von B_{d-1} hat die darstellende Matrix von $f|_{V_{d-1}} : V_{d-1} \rightarrow V_{d-1}$ die Gestalt

$$\begin{pmatrix} J(d-1) & & \\ & \ddots & \\ & & J(d-1) \end{pmatrix}$$

mit β_{d-1} Jordanblöcken. Betrachte nun ein Komplement zu $V_d + V_{d-1} + \text{Kern}(f^{d-3})$ in $\text{Kern}(f^{d-2})$.

...

Durch zusammenfügen von $B_d, B_{d-1}, \dots, B_1$ erhält man eine Jordanbasis.

Ein explizites Beispiel:

Betrachte die Matrix

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

Es ist

$$\begin{aligned} P_A(X) &= \det(X \cdot E - A) \\ &= (X - 3)^2 X + 4 + 6 + 2(X - 3) + 4(X - 3) - 3X \\ &= (X - 2)^3, \end{aligned}$$

das charakteristische Polynom zerfällt in Linearfaktoren und es gibt nur einen Eigenwert, nämlich 2. Hieraus folgt, dass der Hauptraum zum Eigenwert 2 der ganze Raum, K^3 , sein muss. Die Matrix

$$B := A - 2E = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix}$$

ist nilpotent. In der Tat:

$$\begin{aligned} B^2 &= \begin{pmatrix} 0 & 2 & 2 \\ 0 & -2 & -2 \\ 0 & 2 & 2 \end{pmatrix}, \\ B^3 &= 0. \end{aligned}$$

Es ist

$$\begin{aligned} \text{Rang}(B) &= 2, \\ \text{Rang}(B^2) &= 1, \\ \text{Rang}(B^3) &= 0. \end{aligned}$$

Daraus folgt $\beta_d = 0$ für alle $d > 3$ und $\beta_3 = \text{Rang}(B^2) = 1$. Es gibt daher genau einen Jordanblock der Länge 3, die Jordan-Normalform von B ist

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

Die Jordan-Normalform von $A = B + 2E$ ist dann

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

Kontrolle:

$$\begin{aligned} \beta_2 &= \text{Rang}(B) - 2\beta_3 = 2 - 2 \cdot 1 = 0, \\ \beta_1 &= \text{Rang}(B^0) - 2\beta_2 - 3\beta_3 = 3 - 0 - 3 = 0. \end{aligned}$$

Bestimmung einer Jordanbasis:

Es gilt $B^3 = 0 \neq B^2$ ($d = 3$) und

$$\text{Kern}(B^2) = \text{Kern} \begin{pmatrix} 0 & 2 & 2 \\ 0 & -2 & -2 \\ 0 & 2 & 2 \end{pmatrix} = \text{Lin} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right).$$

Durch $v_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ wird $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ zu einer Basis von K^3 ergänzt. Setze $C = (B^2v_1, Bv_1, v_1) = \left(\begin{pmatrix} 4 \\ -4 \\ 4 \end{pmatrix}, \begin{pmatrix} 7 \\ -3 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right)$. Es gilt

$$\mathcal{M}_C^C(B) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\mathcal{M}_C^C(A) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

7 Linear- und Bilinearformen

7.1 Linearformen und der Dualraum

Sei K ein Körper und V ein K -Vektorraum.

Definition 7.1. Eine lineare Abbildung $f : V \rightarrow K$ heißt *Linearform*. Die Menge $V^* := \text{Hom}(V, K)$ aller Linearformen auf V heißt *Dualraum* von V . Der Dualraum V^* ist ein K -Vektorraum.

Bemerkung 7.2. Ist V endlichdimensional, so gilt $\dim(V) = \dim(V^*)$.

Beweis. Sei $B = (b_1, \dots, b_n)$ eine Basis von V und nehme 1 als Basis des K -Vektorraums K . Zu einem Homomorphismus $f : V \rightarrow K$ betrachten wir die Koordinatenmatrix

$$A = \mathcal{M}_1^B(f) = (a_1, \dots, a_n) \in K^{1 \times n}$$

wobei

$$a_j = f(b_j).$$

Sei $E = (e_1, \dots, e_n)$ die Standardbasis von K^n und $\sigma_B : K^n \rightarrow V$ die lineare Abbildung mit $e_i \mapsto b_i$. Dann haben wir das kommutative Diagramm

$$\begin{array}{ccc} V & \xrightarrow{f} & K \\ \sigma_B \uparrow & & \uparrow \text{id}_K \\ K^n & \xrightarrow{x \mapsto Ax} & K \end{array}$$

und wir erhalten einen Isomorphismus

$$d_B : V^* \rightarrow K^{1 \times n}, \quad f \mapsto \mathcal{M}_1^B(f).$$

□

Definition 7.3. Die zur Basis B von V *duale Basis* $B^* = (b_1^*, \dots, b_n^*)$ besteht aus den eindeutig bestimmten Linearformen b_i^* , für die

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

für $i, j \in \{1, \dots, n\}$ gilt. Es gilt also

$$d_B(b_i^*) = e_i^t.$$

Bemerkung 7.4. Für $V = K^n = K^{n \times 1}$ und $B = E$ kommutiert das folgende Diagramm von Isomorphismen

$$\begin{array}{ccccc} K^n & \xrightarrow{e_i \mapsto e_i^*} & (K^n)^* & \xrightarrow{d_E : e_i^* \mapsto e_i^t} & K^{1 \times n}. \\ & \searrow & & \nearrow & \\ & & x \mapsto x^t & & \end{array}$$

Definition 7.5 (Dual einer linearen Abbildung). Sei $h : V \rightarrow W$ eine lineare Abbildung. Die zu h *duale (transponierte) Abbildung* ist definiert durch

$$h^* : W^* \rightarrow V^*, \quad h^*(f) = f \circ h.$$

Bemerkung 7.6. Sind $h_1 : V \rightarrow W$ und $h_2 : W \rightarrow X$ lineare Abbildungen, so gilt

$$(h_2 \circ h_1)^* = h_1^* \circ h_2^*.$$

Ist $h : V \rightarrow W$ ein Isomorphismus, so auch $h^* : W^* \rightarrow V^*$ und es gilt $(h^*)^{-1} = (h^{-1})^*$.

Beweis. Übung. □

Lemma 7.7. Sei $f : V \rightarrow W$ linear, $B = (b_1, \dots, b_n)$ eine Basis von V und $C = (c_1, \dots, c_n)$ eine Basis von W . Sei $f^* : W^* \rightarrow V^*$ die duale Abbildung zu f und C^*, B^* die dualen Basen. Ist

$$A = \mathcal{M}_C^B(f),$$

so gilt

$$A^t = \mathcal{M}_{B^*}^{C^*}(f^*).$$

Beweis. Übung. □

Wir erhalten als Korollar zwei Identitäten, die uns schon gut bekannt sind.

Korollar 7.8. Für Matrizen A und B gilt

$$(AB)^t = B^t A^t$$

und

$$(A^t)^{-1} = (A^{-1})^t$$

falls A invertierbar ist.

Definiere ein „Produkt“

$$\begin{aligned} \langle \cdot, \cdot \rangle : V^* \times V &\rightarrow K, \\ (f, x) &\mapsto \langle f, x \rangle := f(x). \end{aligned}$$

- Dies ist eine bilineare Abbildung, das heißt

$$\begin{aligned}\langle f_1 + f_2, x \rangle &= \langle f_1, x \rangle + \langle f_2, x \rangle, \\ \langle f, x_1 + x_2 \rangle &= \langle f, x_1 \rangle + \langle f, x_2 \rangle, \\ \langle \lambda f, x \rangle &= \lambda \langle f, x \rangle = \langle f, \lambda x \rangle\end{aligned}$$

für $f, f_1, f_2 \in V^*$, $x, x_1, x_2 \in V$ und $\lambda \in K$.

- Ist $h : V \rightarrow W$ linear und $f \in W^*$, so ist $h^*(f) \in V^*$ und

$$\langle h^*(f), x \rangle = \langle f, h(x) \rangle.$$

Definition 7.9. Sind $f \in V^*$ und $x \in V$ mit

$$\langle f, x \rangle = 0,$$

so heißt f *orthogonal* zu x . Schreibe „ $f \perp x$ “. Ist $M \subset V$ eine Teilmenge, so sei der Untervektorraum

$$M^0 = \{f \in V^* \mid \langle f, x \rangle = 0 \text{ für alle } x \in M\} \subset V^*$$

das *orthogonale Komplement* oder der *Annulator* von M . Ist $F \subset V^*$ eine Teilmenge, so sei der Untervektorraum

$$F^0 = \{x \in V \mid \langle f, x \rangle = 0 \text{ für alle } f \in F\} \subset V$$

das *orthogonale Komplement* oder *Annulator* von F .

Bemerkung 7.10. Für $M, M_1, M_2 \subset V$ gilt

- i) $M_1 \subset M_2 \Rightarrow M_2^0 \subset M_1^0$,
- ii) $M^0 = (\text{Lin}(M))^0$,
- iii) $M \subset M^{00}$,
- iv) $M^0 \subset M^{000}$.

Weiterhin folgt aus iv) zusammen mit i) und iii) bereits, dass $M^0 = M^{000}$ gilt.

Analoge Aussagen gelten für $F, F_1, F_2 \in V^*$. Für Teilräume $M_1, M_2 \subset V$ beziehungsweise $F_1, F_2 \in V^*$ gilt

$$\begin{aligned}(M_1 + M_2)^0 &= M_1^0 \cap M_2^0, \\ M_1^0 + M_2^0 &\subset (M_1 \cap M_2)^0.\end{aligned}$$

Lemma 7.11. Sei V ein endlichdimensionaler K -Vektorraum und $M \subset V$ eine Teilmenge. Es gilt

$$\dim(M^0) = \dim(V) - \dim(\text{Lin}(M)).$$

Beweis. Sei $U = \text{Lin}(M)$. Wegen $U^0 = M^0$ reicht es, $\dim(U^0) = \dim(V) - \dim(U)$ zu zeigen. Sei $m = \dim(U)$ und $n = \dim(V)$. Sei $(b_1, \dots, b_m)$ eine Basis von U . Setze diese zu einer Basis $(b_1, \dots, b_m, b_{m+1}, \dots, b_n)$ von V fort. Betrachte die duale Basis $(b_1^*, \dots, b_n^*)$ von V^* .

Behauptung: $b_{m+1}^*, \dots, b_n^*$ bilden eine Basis von U^0 .

Sei $f = \sum_{i=1}^n x_i b_i^* \in V^*$ beliebig. Es gilt

$$\begin{aligned} f \in U^0 &\Leftrightarrow \langle f, b_j \rangle = 0 \text{ für alle } j = 1, \dots, m \\ &\Leftrightarrow x_j = 0 \text{ für alle } j = 1, \dots, m, \end{aligned}$$

woraus die Behauptung folgt. □

Der Raum $V^{**} := (V^*)^*$ heißt *Bidualraum* von V . Wir haben die lineare Abbildung

$$V \rightarrow V^{**}, \quad x \mapsto \hat{x},$$

wobei $\hat{x}$ die Abbildung “Auswerten an x ” ist, d.h.

$$\hat{x} : V^* \rightarrow K, \quad \hat{x}(f) = f(x) = \langle f, x \rangle.$$

Bemerkung 7.12. Die Abbildung $V \rightarrow V^{**}$, $x \mapsto \hat{x}$ ist injektiv.

Beweis. Es reicht, $x \neq 0 \Rightarrow \hat{x} \neq 0$, zu zeigen. Sei dazu $b_1 = x \neq 0$ beliebig. Ergänze dies zu einer Basis B von V . Für $f = b_1^*$ aus der dualen Basis B^* von V^* gilt $\langle f, x \rangle = f(x) = b_1^*(b_1) = 1 \neq 0$, woraus $\hat{x} \neq 0$ folgt. □

Bemerkung 7.13. Ist $\dim(V) < \infty$, so ist $V \rightarrow V^{**}$, $x \mapsto \hat{x}$ ein kanonischer Isomorphismus.

Satz 7.14 (Dualitätssatz). *Sei V ein endlichdimensionaler K -Vektorraum. Dann ist*

$$\langle \cdot, \cdot \rangle : V^* \times V \rightarrow K, \quad (f, x) \mapsto \langle f, x \rangle = f(x)$$

nicht ausgeartet, das heißt

$$\begin{aligned} \langle f, x \rangle = 0 \text{ für alle } x \in V &\Rightarrow f = 0, \\ \langle f, x \rangle = 0 \text{ für alle } f \in V^* &\Rightarrow x = 0. \end{aligned}$$

Für einen Teilraum $M \subset V$ gilt

$$\dim(M^0) = \dim(V) - \dim(M).$$

Beweis. Falls $\langle f, x \rangle = 0$ ist für alle $x \in V$, so heißt dies $f(x) = 0$ für alle $x \in V$, also ist $f = 0$. Falls $\langle f, x \rangle = 0$ ist für alle $f \in V^*$, so ist $\hat{x} = 0$ und somit $x = 0$ nach Bemerkung 7.13. Die Zusatzaussage über die Dimensionen haben wir bereits in Lemma 7.11 bewiesen. □

Folgerung 7.15. *Ist V endlich-dimensional und $U \subset V$ ein Teilraum, so gilt $U = U^{00}$.*

7.2 Bilinearformen

Sei K ein Körper.

Definition 7.16. Seien V, W K -Vektorräume. Eine Abbildung $\beta : V \times W \rightarrow K$ heißt eine *Bilinearform*, falls β linear in beiden Variablen ist, das heißt

$$\begin{aligned}\beta(x + x', y) &= \beta(x, y) + \beta(x', y), \\ \beta(x, y + y') &= \beta(x, y) + \beta(x, y'), \\ \beta(\lambda x, y) &= \lambda\beta(x, y) = \beta(x, \lambda y)\end{aligned}$$

für $x, x' \in V$, $y, y' \in W$ und $\lambda \in K$. Die Bilinearform heißt *nicht ausgeartet in der ersten Variablen*, falls

$$\beta(x, y) = 0 \text{ für alle } y \in W \quad \Rightarrow \quad x = 0$$

und *nicht ausgeartet in der zweiten Variablen*, falls

$$\beta(x, y) = 0 \text{ für alle } x \in V \quad \Rightarrow \quad y = 0.$$

Ist $V = W$, so heißt $\beta : V \times V$ eine Bilinearform auf V . Mit $\text{Bil}(V, W) = \text{Bil}_K(V, W)$ bezeichnen wir die Menge aller Bilinearformen $\beta : V \times W \rightarrow K$. Die Menge $\text{Bil}_K(V, W)$ ist offenbar ein K -Vektorraum.

Beispiel 7.17.

- 1) Die Multiplikation eines Körpers,

$$\cdot : K \times K, \quad (x, y) \mapsto x \cdot y,$$

ist eine Bilinearform.

- 2) Sei $B = (b_{ij}) \in K^{n \times m}$. Dann ist die Abbildung β_B , gegeben durch

$$\beta_B : K^n \times K^m \rightarrow K, \quad (x, y) \mapsto x^t B y,$$

eine Bilinearform. Für $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$ ist

$$\beta_B(x, y) = \sum_{i=1}^n \sum_{j=1}^m x_i b_{ij} y_j.$$

Man nennt β_B die von B vermittelte Bilinearform.

- 3) Aus der Analysis:

Sei $f : G \rightarrow \mathbb{R}$ eine auf einer offenen Teilmenge $G \subset \mathbb{R}^n$ definierte (total) differenzierbare Funktion. Für jedes $x \in G$ ist dann die totale Ableitung (das Differential) $f'(x) = df(x)$ eine Linearform auf $\mathbb{R}^n$. Ist $df : G \rightarrow (\mathbb{R}^n)^*$ wieder (total) differenzierbar, so kann man $f''(x)$ für alle $x \in G$ als Bilinearform sehen.

4) Die Abbildung $\langle \cdot, \cdot \rangle : V^* \times V \rightarrow K$, $(f, x) \mapsto \langle f, x \rangle = f(x)$ ist eine Bilinearform.

Bemerkung 7.18. Seien V und W zwei endlichdimensionale K -Vektorräume mit $\dim(V) = \dim(W)$. Dann ist jede Bilinearform $\beta : V \times W \rightarrow K$ genau dann nicht ausgeartet in der ersten Variablen, wenn sie nicht ausgeartet in der zweiten Variablen ist.

Beweis. Wir zeigen eine Richtung, die andere geht analog. Angenommen $\beta : V \times W \rightarrow K$ sei in der ersten Variablen ausgeartet. Dann existiert ein $v_1 \in V$, $v_1 \neq 0$, sodass $\beta(v_1, w) = 0$ ist für alle $w \in W$. Ergänze v_1 zu einer Basis $v_1, \dots, v_n$ von V und sei $w_1, \dots, w_n$ eine Basis von W .

Da $\beta(v_1, w) = 0$ für alle w ist, müssen wir also nur noch das lineare Gleichungssystem

$$\sum_{i=1}^n x_i \beta(v_j, w_i) = 0 \text{ für } j = 2, \dots, n,$$

lösen. Dies sind $n - 1$ Gleichungen in n Variablen, d.h. das homogene System besitzt eine nicht-triviale Lösung $(\lambda_1, \dots, \lambda_n) \in K^n \setminus \{0\}$ und $w = \lambda_1 w_1 + \dots + \lambda_n w_n$ erfüllt $\beta(v, w) = 0$ für alle $v \in V$. $\square$

Sei $\beta : V \times W \rightarrow K$ eine Bilinearform. Zu einem festen $x \in V$ betrachten wir die Abbildung

$$\beta_1(x) : W \rightarrow K, \quad y \mapsto \beta(x, y).$$

Diese ist linear, weshalb $\beta_1(x)$ für jedes $x \in V$ eine Linearform ist. Es ist also

$$\beta_1 : V \rightarrow W^*$$

eine lineare Abbildung von V in den Dualraum W^* von W . Es gilt

$$\beta(x, y) = \beta_1(x)(y) = \langle \beta_1(x), y \rangle$$

für alle $x \in V$ und $y \in W$.

Satz 7.19. Die Abbildung

$$\text{Bil}_K(V, W) \rightarrow \text{Hom}_K(V, W^*), \quad \beta \mapsto \beta_1$$

ist ein Isomorphismus von Vektorräumen. Weiter gilt:

a) Ist $\dim(V) = n$ und $\dim(W) = m$, dann ist $\dim(\text{Bil}(V, W)) = mn$.

b) β ist genau dann nicht ausgeartet in der ersten Variablen, wenn β_1 injektiv ist.

Beweis. Zu $h \in \text{Hom}(V, W^*)$ definieren wir eine Bilinearform $\beta_h : V \times W \rightarrow K$, $\beta_h(x, y) = \langle h(x), y \rangle$. Der Homomorphismus $h \mapsto \beta_h$ ist eine rechts- und linksseitige Umkehrabbildung des Homomorphismus $\beta \mapsto \beta_1$. Die Eigenschaften a) und b) sind klar. $\square$

Bemerkung 7.20. 1) Die Abbildung $\beta \mapsto \beta_1$ ist ein basisunabhängiger Isomorphismus von Vektorräumen. Deswegen können wir Bilinearformen $V \times W \rightarrow K$ mit den entsprechenden linearen Abbildungen $V \rightarrow W^*$ identifizieren.

2) Völlig analog kann man eine Abbildung $\beta_2 : W \rightarrow V^*$, $y \mapsto \beta_2(y) = \beta(\cdot, y)$ definieren.

Definition 7.21. Seien V und W endlichdimensionale K -Vektorräume. Es sei $\dim(V) = n$ und $\dim(W) = m$. Sei $\beta : V \times W \rightarrow K$ eine Bilinearform und sei $(v_1, \dots, v_n)$ eine Basis von V , $(w_1, \dots, w_m)$ eine Basis von W . Dann heißt die $n \times m$ -Matrix $B = (b_{ij})$ mit $b_{ij} = \beta(v_i, w_j)$ die *Gram-Matrix (oder Strukturmatrix) der Bilinearform β* in Bezug auf die Basen $(v_1, \dots, v_n)$ von V und $(w_1, \dots, w_m)$ von W .

Satz 7.22. Ist B die Gram-Matrix der Bilinearform $\beta : V \times W \rightarrow K$ bezüglich der Basen $(v_1, \dots, v_n)$ von V und $(w_1, \dots, w_m)$ von W , so ist B^t die Koordinatenmatrix der linearen Abbildung $\beta_1 : V \rightarrow W^*$ bezüglich der Basen $(v_1, \dots, v_n)$ von V und $(w_1^*, \dots, w_m^*)$ von W^* .

Beweis. Für $i = 1, \dots, n$ und $j = 1, \dots, m$ gilt nach Definition

$$\beta_1(v_i)(w_j) = b_{ij}.$$

Daraus folgt

$$\beta_1(v_i) = \sum_{j=1}^m b_{ij} w_j^*.$$

□

Bemerkung 7.23. Analog zeigt man, dass B die Koordinatenmatrix der linearen Abbildung $\beta_2 : W \rightarrow V^*$ bezüglich der Basen $(w_1, \dots, w_m)$ von W und $(v_1^*, \dots, v_n^*)$ von V^* ist.

Bemerkung 7.24. Seien β und B wie oben.

1) Für $x = x_1 v_1 + \dots + x_n v_n$ und $y = y_1 w_1 + \dots + y_m w_m$ gilt

$$\beta(x, y) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} B \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Dies folgt direkt aus der Bilinearität von β .

2) β ist durch B (und die gewählten Basen) eindeutig festgelegt. Dies erklärt die Bezeichnung Strukturmatrix.

3) Die Abbildung

$$\text{Bil}_K(V, W) \rightarrow K^{n \times m}, \quad \beta \mapsto B$$

ist ein (nicht-kanonischer) Isomorphismus. Nach Satz 7.19 entspricht dieser Isomorphismus der Abbildung, die man erhält, indem man zunächst β auf β_2 abbildet und dann β_2 auf ihre Darstellungsmatrix bzgl. der gewählten Basen.

Satz 7.25. Sei B die Gram-Matrix der Bilinearform $\beta : V \times W \rightarrow K$ bezüglich der Basen C von V und D von W . Seien C' und D' weitere Basen von V , beziehungsweise W , mit zugehöriger Gram-Matrix B' . Dann gilt

$$B' = \left(\mathcal{M}_C^{C'}(\text{id}_V) \right)^t \cdot B \cdot \mathcal{M}_D^{D'}(\text{id}_W).$$

Beweis. Übung. □

Lemma 7.26. Sei V ein n -dimensionaler K -Vektorraum und $\beta : V \times V \rightarrow K$ eine Bilinearform auf V . Seien B eine Gram-Matrix von β bezüglich einer Basis von V und B' eine Gram-Matrix bezüglich einer weiteren Basis von V . Dann gilt

$$\det(B') = a^2 \det(B)$$

mit einem $a \in K^\times$.

Beweis. Dies folgt aus Satz 7.25 mit $C = D$ und $C' = D'$. □

Definition 7.27. Seien $B, C \in K^{n \times n}$. Dann heißt C kongruent zu B (über K), falls es ein $S \in \text{GL}(n, K)$ gibt, so dass

$$C = S^t B S.$$

Schreibe $C \simeq B$.

Bemerkung 7.28. Es ist $C \simeq B$ genau dann, wenn es einen K -Vektorraum V und eine Bilinearform β auf V gibt, so dass sowohl B als auch C als Strukturmatrizen (Gram-Matrizen) von β (bezüglich geeigneter Basen) auftreten.

Sei $\beta : V \times W \rightarrow K$ eine Bilinearform und $f : V \rightarrow V$ ein Endomorphismus. Dann ist auch

$$\beta_f : V \times W \rightarrow K, \quad \beta_f(x, y) = \beta(f(x), y)$$

eine Bilinearform.

Frage. Erhält man jede Bilinearform $\gamma : V \times W \rightarrow K$ als ein β_f ?

Satz 7.29. Seien V, W endlichdimensionale K -Vektorräume mit $\dim(V) = \dim(W)$ und sei $\beta : V \times W \rightarrow K$ eine nicht ausgeartete Bilinearform. Zu jeder Bilinearform $\gamma : V \times W \rightarrow K$ existiert dann genau ein Endomorphismus $f : V \rightarrow V$, so dass $\gamma = \beta_f$, das heißt $\gamma(x, y) = \beta(f(x), y)$ für alle $x \in V$ und $y \in W$. (Analog gilt dies für die zweite Komponente.)

Beweis. Betrachte die Abbildung

$$\beta_1 : V \rightarrow W^*, \quad x \mapsto \underbrace{\beta(x, \cdot)}_{y \mapsto \beta(x, y)}.$$

Da β nicht ausgeartet ist, ist β_1 injektiv. Wegen $\dim(V) = \dim(W) = \dim(W^*)$ muss β_1 dann bereits ein Isomorphismus sein. Es gilt

$$\begin{aligned}\gamma(x, y) &= \beta(f(x), y) \quad \text{für alle } x \in V, y \in W \\ \Leftrightarrow \gamma_1 &= \beta_1 \circ f \\ \Leftrightarrow (\beta_1)^{-1} \circ \gamma_1 &= f.\end{aligned}$$

Der Endomorphismus $f := (\beta_1)^{-1} \circ \gamma_1$ hat daher die geforderten Eigenschaften und ist aufgrund der Äquivalenz eindeutig bestimmt. $\square$

Satz 7.30 (und Definition). *Seien V und W endlichdimensionale K -Vektorräume der gleichen Dimension, $\beta : V \times W \rightarrow K$ eine nicht ausgeartete Bilinearform. Zu jedem Endomorphismus $h : V \rightarrow V$ existiert genau ein Endomorphismus $h^* : W \rightarrow W$ mit*

$$\beta(h(x), y) = \beta(x, h^*(y))$$

für alle $x \in V$ und $y \in W$. Man nennt h^ den zu h (rechts-)adjungierten Endomorphismus (bezüglich β). Analog gibt es zu jedem Endomorphismus $g : W \rightarrow W$ genau einen Endomorphismus ${}^*g : V \rightarrow V$ mit*

$$\beta(x, g(y)) = \beta({}^*g(x), y)$$

für alle $x \in V$ und $y \in W$. Dies ist der zu g (links-)adjungierte Endomorphismus (bezüglich β).

Beweis. Nach Satz 7.29 für die zweite Komponente existiert zu $\beta_h : V \times W \rightarrow K$ ein $h^* : W \rightarrow W$ mit

$$\beta_h(x, y) = \beta(x, h^*(y)),$$

das heißt

$$\beta(h(x), y) = \beta(x, h^*(y)),$$

für alle $x \in V$ und $y \in W$. $\square$

Bemerkung 7.31. Die adjungierte Abbildung kann man als eine Verallgemeinerung der dualen Abbildung ansehen, welche man als Spezialfall erhält, wenn man die nicht ausgeartete Bilinearform die Paarung $V^* \times V \rightarrow K$, $(f, x) \mapsto \langle f, x \rangle = f(x)$ verwendet.

Bemerkung 7.32. In der obigen Situation gilt ${}^*(h^*) = h$.

Satz 7.33. *Seien V, W, β wie im vorherigen Satz. Sei $v_1, \dots, v_n$ eine Basis von V , $w_1, \dots, w_n$ eine Basis von W und B die zugehörige Strukturmatrix von β . Hat ein Endomorphismus h von V bezüglich $v_1, \dots, v_n$ die Koordinatenmatrix A , so hat h^* bezüglich $w_1, \dots, w_n$ die Koordinatenmatrix*

$$B^{-1}A^tB.$$

Beweis. Die Bedingung

$$\beta(h(x), y) = \beta(x, h^*(y))$$

für alle $x \in V$ und $y \in W$ ist äquivalent zu

$$h^t \circ \beta_2 = \beta_2 \circ h^*.$$

In dieser Gleichung bezeichnet h^t die zu h duale Abbildung $h^t : V^* \rightarrow V^*$. Da β nicht ausgeartet ist und V und W die gleiche Dimension haben, ist β_2 ein Isomorphismus. Es folgt

$$h^* = \beta_2^{-1} \circ h^t \circ \beta_2.$$

Die Koordinatenmatrix von β_2 ist B . Die von h^t ist A^t . Damit folgt die Behauptung. $\square$

Definition 7.34. Sei $\beta : V \times W \rightarrow K$ eine Bilinearform. Zwei Vektoren $x \in V$ und $y \in W$ heißen orthogonal zueinander bezüglich β , falls $\beta(x, y) = 0$. Sind $X \subset V$ und $Y \subset W$ Teilmengen, so definiert man deren *orthogonale Komplemente* durch

$$\begin{aligned} X^\perp &= \{y \in W \mid \beta(x, y) = 0 \text{ für alle } x \in X\}, \\ {}^\perp Y &= \{x \in V \mid \beta(x, y) = 0 \text{ für alle } y \in Y\}. \end{aligned}$$

Bemerkung 7.35.

- i) Die Mengen $X^\perp \subset W$ und ${}^\perp Y \subset V$ sind Untervektorräume.
- ii) Es gilt $X^\perp = (\text{Lin}(X))^\perp$ und ${}^\perp Y = {}^\perp(\text{Lin}(Y))$.
- iii) Es gilt ${}^\perp W = \text{Kern}(\beta_1)$ und $V^\perp = \text{Kern}(\beta_2)$.

Beweis. Übung. $\square$

Satz 7.36 (Dimensionsformel für orthogonale Komplemente). *Seien V, W endlichdimensionale Vektorräume und $\beta : V \times W \rightarrow K$ eine Bilinearform. Für jeden Teilraum $X \subset V$ gilt*

$$\dim(X) + \dim(X^\perp) = \dim(W) + \dim({}^\perp W \cap X).$$

Für jeden Teilraum $Y \subset W$ gilt

$$\dim(Y) + \dim({}^\perp Y) = \dim(V) + \dim(V^\perp \cap Y).$$

Beweis. Betrachte den Homomorphismus $\beta_1 : V \rightarrow W^*$ und seine Einschränkung

$$\beta_1|_X : X \rightarrow W^*.$$

Wegen ${}^\perp W = \text{Kern}(\beta_1)$ gilt

$${}^\perp W \cap X = \text{Kern}(\beta_1|_X).$$

Aus der Dimensionsformel für lineare Abbildungen folgt

$$\dim(\text{Bild}(\beta_1|_X)) = \dim(X) - \dim({}^\perp W \cap X),$$

nach Lemma 7.11 gilt

$$\dim(\text{Bild}(\beta_1|_X)) + \dim((\text{Bild}(\beta_1|_X))^0) = \dim(W^*) = \dim(W).$$

Es ist $(\text{Bild}(\beta_1|_X))^0 = X^\perp$, woraus

$$\dim(\text{Bild}(\beta_1|_X)) = \dim(W) - \dim(X^\perp)$$

und schließlich die Behauptung folgt. $\square$

Korollar 7.37. *Es gilt*

$$\dim(V) + \dim(V^\perp) = \dim(W) + \dim({}^\perp W).$$

Beweis. Satz 7.36 mit $X = V$. $\square$

Korollar 7.38. *Falls $\dim(V) = \dim(W)$, so ist $\dim(V^\perp) = \dim({}^\perp W)$. In diesem Fall ist die Bilinearform β genau dann nicht ausgeartet bezüglich der ersten Variable, wenn sie bezüglich der zweiten Variable nicht ausgeartet ist (vergleiche Bemerkung 7.18).*

Korollar 7.39. *Seien V, W endlichdimensionale K -Vektorräume der gleichen Dimension und $\beta : V \times W \rightarrow K$ eine nicht ausgeartete Bilinearform. Dann gilt für Teilräume $X \subset V$ und $Y \subset W$*

$$\begin{aligned} \dim(X) + \dim(X^\perp) &= \dim(V), \\ \dim(Y) + \dim({}^\perp Y) &= \dim(W). \end{aligned}$$

Außerdem gilt ${}^\perp(X^\perp) = X$ und $({}^\perp Y)^\perp = Y$.

Korollar 7.40. *Sei V endlich-dimensional, β eine nicht-ausgeartete Bilinearform auf V und $U \subset V$ ein Teilraum. Dann gilt $\beta_1(U^\perp) = U^0 \subset V^*$.*

Beweis. Es ist $U^0 = \{f \in V^* \mid f(x) = 0 \text{ für alle } x \in U\}$ und klarerweise ist $\beta_1(U^\perp) = \{y \mapsto \beta_1(x)(y) \mid x \in U^\perp\} \subset U^0$. Da nach den Dimensionsformeln (Satz 7.14 und Korollar 7.39) gilt $\dim(U^\perp) + \dim(U) = \dim(V) = \dim(U) + \dim(U^0)$, wonach $\dim(U^\perp) = \dim(U^0)$ ist und somit die Gleichheit folgt (denn β_1 ist injektiv nach Annahme). $\square$

Symmetrische und schiefsymmetrische Bilinearformen

Definition 7.41. Eine Bilinearform $\beta : V \times V \rightarrow K$ heißt *symmetrisch*, falls

$$\beta(x, y) = \beta(y, x)$$

für alle $x, y \in V$. Sie heißt *schiefsymmetrisch*, falls

$$\beta(x, y) = -\beta(y, x)$$

für alle $x, y \in V$.

In beiden Fällen gilt $X^\perp = {}^\perp X$ für $X \subset V$, und wir schreiben stets $X^\perp$ dafür.

Bemerkung 7.42. Sei $\beta : V \times V \rightarrow K$ eine Bilinearform und sei B die Gram-Matrix von β bezüglich einer Basis von V . Es gilt

$$\begin{aligned}\beta \text{ ist symmetrisch} &\Leftrightarrow B = B^t, \\ \beta \text{ ist schiefsymmetrisch} &\Leftrightarrow B = -B^t.\end{aligned}$$

Beweis. Übung. □

Bemerkung 7.43. Sei $\text{char}(K) \neq 2$ (das heißt $1 + 1 \neq 0 \in K$). Für eine Bilinearform $\beta : V \times V \rightarrow K$ sind dann äquivalent:

- i) β ist schiefsymmetrisch,
- ii) β ist symplektisch, das heißt $\beta(x, x) = 0$ für alle $x \in V$.

Beweis. (i) $\Rightarrow$ (ii): Es gilt

$$\begin{aligned}\beta(x, x) &= -\beta(x, x) \text{ für alle } x \in V \\ \Rightarrow 2\beta(x, x) &= 0 \text{ für alle } x \in V \\ \Rightarrow \beta(x, x) &= 0 \text{ für alle } x \in V.\end{aligned}$$

(ii) $\Rightarrow$ (i): Es gilt

$$0 = \beta(x + y, x + y) = \underbrace{\beta(x, x)}_{=0} + \beta(x, y) + \beta(y, x) + \underbrace{\beta(y, y)}_{=0}$$

für alle $x, y \in V$, woraus

$$\beta(x, y) = -\beta(y, x)$$

für alle $x, y \in V$ folgt. □

Satz 7.44. Sei $\beta : V \times V \rightarrow K$ eine nicht ausgeartete symplektische Bilinearform. Dann existiert eine symplektische Basis, das heißt eine Basis $u_1, v_1, \dots, u_m, v_m$ von V bezüglich der die Strukturmatrix die Gestalt

$$B = \begin{pmatrix} H & & & \\ & H & & \\ & & \ddots & \\ & & & H \end{pmatrix} \in K^{2m \times 2m}$$

hat, wobei $H = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Insbesondere ist $\dim(V) = 2m$ gerade und die Determinante einer beliebigen Gram-Matrix von β ein Quadrat.

Beweis. Sei $n = \dim(V)$. Sei $u_1 \neq 0$ ein beliebiger Vektor aus V . Da β nicht ausgeartet ist, gibt es einen Vektor v_1 mit $\beta(u_1, v_1) = 1$. Sei $V_0 = Ku_1 + Kv_1 \subset V$ und $W = V_0^\perp \subset V$. Es ist $\dim(V_0) = 2$ und $\beta|_{V_0}$ hat die Strukturmatrix H bezüglich der Basis u_1, v_1 . Da β nicht ausgeartet ist gilt nach der Dimensionsformel für orthogonale Komplemente (Satz 7.36)

$$\dim(W) = \dim(V) - \dim(V_0) = n - 2.$$

Die Einschränkung $\beta|_W : W \times W \rightarrow K$ ist eine nicht ausgeartete symplektische Bilinearform. Die Behauptung folgt per Induktion nach $\dim(V)$. Weiter gilt $\det(H) = 1$ und daher $\det(B) = 1$. Daraus folgt mit Lemma 7.26, dass die Determinante jeder Gram-Matrix von β ein Quadrat ist. $\square$

Definition 7.45. Sei $\beta : V \times V \rightarrow K$ eine Bilinearform. Eine lineare Abbildung $s : V \rightarrow V$ heißt *Isometrie*, falls

$$\beta(s(x), s(y)) = \beta(x, y)$$

für alle $x, y \in V$.

Bemerkung 7.46. Sei V ein K -Vektorraum der Dimension n und $\beta : V \times V \rightarrow K$ eine nicht ausgeartete Bilinearform. Für einen Endomorphismus $s : V \rightarrow V$ sind äquivalent:

- i) s ist eine Isometrie (bezüglich β).
- ii) Es gilt $s^* \circ s = \text{id}_V$.
- iii) s ist invertierbar und $s^{-1} = s^* = {}^*s$.
- iv) Ist $b_1, \dots, b_n$ eine Basis von V , B die Strukturmatrix von β und S die darstellende Matrix von s bezüglich dieser Basis, so gilt $B = S^t B S$.

Beweis. Übung. $\square$

Korollar 7.47. Sei V ein K -Vektorraum der Dimension $2m$ und $\beta : V \times V \rightarrow K$ eine nicht ausgeartete symplektische Bilinearform. Dann ist $(V, \beta) \simeq (K^{2m}, \beta_B)$ mit B wie in Satz 7.44 und vermittelter Bilinearform β_B , welche B wie in Beispiel 7.17 (2) zugeordnet ist.

8 Euklidische und unitäre Vektorräume

8.1 Euklidische Vektorräume

Auf $\mathbb{R}^n$ ist das Standard-Skalarprodukt durch

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R},$$

$$\langle x, y \rangle = x^t \cdot y = \sum_{i=1}^n x_i y_i,$$

definiert. Diese Abbildung ist linear in jeder Variablen, symmetrisch (das heißt $\langle x, y \rangle = \langle y, x \rangle$) und es gilt $\langle x, x \rangle > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$.

Wir wollen vergleichbare Abbildungen allgemeiner betrachten.

Definition 8.1. Sei V ein $\mathbb{R}$ -Vektorraum. Ein (*euklidisches*) *Skalarprodukt* auf V ist eine positiv definite symmetrische Bilinearform $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$. Dabei heißt *positiv definit*: Es gilt $\langle x, x \rangle > 0$ für alle $x \in V \setminus \{0\}$.

Ein *euklidischer Vektorraum* ist ein Paar $(V, \langle \cdot, \cdot \rangle)$ bestehend aus einem $\mathbb{R}$ -Vektorraum V und einem euklidischen Skalarprodukt auf V . Wegen ii) ist auch für festes $x \in V$ die Abbildung $y \mapsto \langle x, y \rangle$ linear.

Notation.

- Zwei Vektoren $x, y \in V$ heißen *orthogonal* zueinander, falls $\langle x, y \rangle = 0$. Schreibe $x \perp y$.
- Die *euklidische Norm* von $x \in V$ ist definiert als

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

- Die Vektoren $x_1, \dots, x_n \in V$ heißen *Orthogonalsystem*, falls $x_i \neq 0$ für alle $i = 1, \dots, m$ und

$$x_i \perp x_j$$

für alle $i \neq j$.

- Die Vektoren $x_1, \dots, x_n \in V$ heißen *Orthonormalsystem*, falls

$$\langle x_i, x_j \rangle = \delta_{ij}$$

für alle i, j .

Bemerkung 8.2. Jedes Orthogonalsystem ist linear unabhängig.

Beweis. Sei $x_1, \dots, x_m$ ein Orthogonalsystem und $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ mit

$$\lambda_1 x_1 + \dots + \lambda_m x_m = 0.$$

Nimmt man auf beiden Seiten das Skalarprodukt mit x_j so erhält man

$$\lambda_j \underbrace{\langle x_j, x_j \rangle}_{\neq 0} = \langle \lambda_1 x_1 + \dots + \lambda_m x_m, x_j \rangle = \langle 0, x_j \rangle = 0,$$

woraus

$$\lambda_j = 0$$

für alle $j = 1, \dots, m$ folgt. □

Bemerkung 8.3. Wir erinnern an das *orthogonale Komplement* $W^\perp$ eines Untervektorraums W :

$$W^\perp = \{a \in V \mid \langle a, x \rangle = 0 \text{ für alle } x \in W\}.$$

Es gelten:

- i) $W^\perp \subset V$ ist ein Untervektorraum von V .
- ii) $W \cap W^\perp = \{0\}$.

Proof. Der Punkt i) ist genau Bemerkung 7.35.

Zu ii): Sei $x \in W \cap W^\perp$. Dann ist $\langle x, x \rangle = 0$ und somit $x = 0$. □

Bemerkung 8.4. Im Gegensatz zum allgemeinen Fall benötigen wir keine separate Notation ${}^\perp W$, da das Skalarprodukt symmetrisch ist.

Bemerkung 8.5. Sei V ein endlichdimensionaler euklidischer Vektorraum. Seien $v_1, \dots, v_m \in V$, wobei $m < \dim(V)$. Dann existiert ein $v \in V \setminus \{0\}$ mit $\langle v, v_i \rangle = 0$ für alle $i = 1, \dots, m$.

Beweis. Sei $W = \text{Lin}(v_1, \dots, v_m)$. Es ist $\dim(W) < \dim(V)$. Nach Korollar 7.39 ist $\dim(W^\perp) = \dim(V) - \dim(W) > 0$, und deshalb existiert ein $v \in W^\perp$ mit $v \neq 0$. □

Korollar 8.6. Jeder endlichdimensionale euklidische Vektorraum besitzt eine Orthonormalbasis.

Beweis. Mit Bemerkung 8.5 finden wir induktiv eine Orthogonalbasis $v_1, \dots, v_n$. Dann ist $\frac{1}{\|v_1\|}v_1, \dots, \frac{1}{\|v_n\|}v_n$ eine Orthonormalbasis. □

Satz 8.7. Sei $W \subset V$ ein Untervektorraum. Es gilt

$$V = W \oplus W^\perp,$$

das heißt V ist die direkte Summe von W und $W^\perp$ (jedes $v \in V$ lässt sich eindeutig als $v = w + w'$ mit $w \in W$ und $w' \in W^\perp$ schreiben).

Beweis. W und $W^\perp$ sind Untervektorräume mit $\dim(W) + \dim(W^\perp) = \dim(V)$ (nach der Dimensionsformel in Korollar 7.39) sowie $W \cap W^\perp = \{0\}$ (nach Bemerkung 8.3). Damit gilt, dass $V = W \oplus W^\perp$. □

Das Gram-Schmidt-Verfahren

Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler euklidischer Vektorraum. Das heißt V ist ein $\mathbb{R}$ -Vektorraum der Dimension $n < \infty$ und $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ ist ein euklidisches Skalarprodukt.

Satz 8.8. V besitzt eine Orthonormalbasis.

Dieser Satz entspricht Korollar 8.6. Wir geben hier einen neuen Beweis mit Hilfe des folgenden Satzes, welcher das „Gram-Schmidtsche Orthonormalisierungsverfahren“ beschreibt.

Satz 8.9. *Sei $W \subset V$ ein Untervektorraum. Jede Orthonormalbasis $(w_1, \dots, w_m)$ von W lässt sich zu einer Orthonormalbasis $(w_1, \dots, w_m, w_{m+1}, \dots, w_n)$ von V ergänzen.*

Beweis. Falls $V = W$, so sind wir bereits fertig.

Angenommen $V \neq W$. Dann gibt es einen Vektor $v \in V \setminus W$. Sei

$$\tilde{v} = \langle w_1, v \rangle w_1 + \dots + \langle w_m, v \rangle w_m.$$

Dies ist das Bild von v unter der orthogonalen Projektion

$$P_W : V \rightarrow W$$

von V auf W . Es gilt dann

$$w := v - \tilde{v} \in W^\perp,$$

denn

$$\langle w_k, w \rangle = \langle w_k, v \rangle - \langle w_k, \tilde{v} \rangle = \langle w_k, v \rangle - \langle w_k, v \rangle \underbrace{\langle w_k, w_k \rangle}_{=1} = 0$$

für $k = 1, \dots, m$. Es ist $w \neq 0$ da $v \neq \tilde{v}$. Setze

$$w_{m+1} := \frac{1}{\|w\|} w.$$

Dann ist $(w_1, \dots, w_{m+1})$ eine Orthonormalbasis von $\tilde{W} = \text{Lin}(w_1, \dots, w_{m+1})$. Dieses Verfahren wiederholt man nun für den Untervektorraum $\tilde{W} \subset V$. $\square$

Beweis von Satz 8.8. Wende Satz 8.9 auf $W = \{0\} \subset V$ an. $\square$

Bemerkung 8.10. In der Praxis ergänzt man $(w_1, \dots, w_m)$ zunächst zu einer Basis $(w_1, \dots, w_m, v_{m+1}, \dots, v_n)$ von V . Dann beginnt man mit $v = v_{m+1} \in V \setminus W$ und macht danach mit den Vektoren $v_{m+2}, \dots, v_n$ weiter. Es gilt dann

$$\text{Lin}(w_1, \dots, w_m, w_{m+1}, \dots, w_{m+k}) = \text{Lin}(w_1, \dots, w_m, v_{m+1}, \dots, v_{m+k})$$

für alle $k = 1, \dots, n - m$.

Algorithmus 8.11 (Gram-Schmidtsches Orthonormalisierungsverfahren). Der folgende Algorithmus berechnet zu linear unabhängigen Vektoren $w_1, \dots, w_m \in V$ ein Orthonormalsystem von m normierten, paarweise orthogonalen Vektoren, das denselben Untervektorraum erzeugt. Die einzelnen Vektoren $v_1, \dots, v_m$ des Orthonormalsystems erhält man,

indem zuerst jeweils ein orthogonaler Vektor berechnet und anschließend normalisiert wird:

$$\begin{aligned}
v_1 &= \frac{1}{\|w_1\|} w_1 && \text{(Normalisieren des ersten Vektors } w_1) \\
v'_2 &= w_2 - \langle v_1, w_2 \rangle \cdot v_1 && \text{(Orthogonalisieren des zweiten Vektors } w_2) \\
v_2 &= \frac{1}{\|v'_2\|} v'_2 && \text{(Normalisieren des Vektors } v'_2) \\
v'_3 &= w_3 - \langle v_1, w_3 \rangle \cdot v_1 - \langle v_2, w_3 \rangle \cdot v_2 && \text{(Orthogonalisieren des dritten Vektors } w_3) \\
v_3 &= \frac{1}{\|v'_3\|} v'_3 && \text{(Normalisieren des Vektors } v'_3) \\
&\vdots \\
v'_j &= w_j - \sum_{i=1}^{j-1} \langle v_i, w_j \rangle \cdot v_i && \text{(Orthogonalisieren des } j\text{-ten Vektors } w_j) \\
v_j &= \frac{1}{\|v'_j\|} v'_j && \text{(Normalisieren des Vektors } v'_j) \\
&\vdots \\
v'_m &= w_m - \sum_{i=1}^{m-1} \langle v_i, w_m \rangle \cdot v_i && \text{(Orthogonalisieren des } m\text{-ten Vektors } w_m) \\
v_m &= \frac{1}{\|v'_m\|} v'_m && \text{(Normalisieren des Vektors } v'_m)
\end{aligned}$$

Beispiel 8.12. Sei $\mathbb{R}^4$ mit dem Standardskalarprodukt versehen und sei $U \subset \mathbb{R}^4$ der von den Vektoren

$$b_1 := \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \quad b_2 := \begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \\ \frac{1}{2} \end{pmatrix}, \quad \text{und} \quad b_3 := \begin{pmatrix} 1 \\ \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

aufgespannte Untervektorraum. Mit dem Gram-Schmidtschen Orthonormalisierungsverfahren bestimmen wir eine Orthonormalbasis von U .

Es ergeben sich die folgenden Schritte:

$$u_1 = b_1 = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix},$$

$$v_1 = \frac{1}{\|u_1\|} u_1 = \frac{1}{\sqrt{(-1)^2 + 1^2 + (-1)^2 + 1^2}} \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix},$$

$$u_2 = b_2 - \langle b_2, v_1 \rangle v_1 = \begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \\ \frac{1}{2} \end{pmatrix} - \left(-\frac{1}{2} + \frac{1}{4} - \frac{1}{2} + \frac{1}{4}\right) \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \\ \frac{3}{4} \\ \frac{3}{4} \\ \frac{3}{4} \end{pmatrix},$$

$$v_2 = \frac{1}{\|u_2\|} u_2 = \frac{1}{\sqrt{\frac{9}{16} + \frac{9}{16} + \frac{9}{16} + \frac{9}{16}}} \begin{pmatrix} \frac{3}{4} \\ \frac{3}{4} \\ \frac{3}{4} \\ \frac{3}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix},$$

$$\begin{aligned} u_3 &= b_3 - \langle b_3, v_1 \rangle v_1 - \langle b_3, v_2 \rangle v_2 \\ &= \begin{pmatrix} 1 \\ \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix} - \left(-\frac{1}{2} + \frac{1}{4} - \frac{1}{4} + \frac{1}{2}\right) \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} - \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{2}\right) \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \\ -\frac{1}{4} \\ -\frac{1}{4} \\ \frac{1}{4} \end{pmatrix}, \end{aligned}$$

$$v_3 = \frac{1}{\|u_3\|} u_3 = \frac{1}{\sqrt{\frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16}}} \begin{pmatrix} \frac{1}{4} \\ -\frac{1}{4} \\ -\frac{1}{4} \\ \frac{1}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

Folglich ist (v_1, v_2, v_3) eine Orthonormalbasis von $\text{Lin}(\{b_1, b_2, b_3\})$.

Bemerkung: Die lineare Unabhängigkeit der drei gegebenen Vektoren folgt daraus, dass man beim Gram-Schmidtschen Orthogonalisierungsverfahren drei von Null verschiedene Vektoren erhält. Alternativ kann man die lineare Unabhängigkeit auch vor dem Verfahren nachrechnen.

Definition 8.13. Zwei euklidische Vektorräume $(V, \langle \cdot, \cdot \rangle)$ und $(W, [\cdot, \cdot])$ heißen *isometrisch isomorph*, falls es einen $\mathbb{R}$ -Vektorraumisomorphismus $\sigma : V \rightarrow W$ mit der Eigenschaft

$$\langle v_1, v_2 \rangle = [\sigma(v_1), \sigma(v_2)]$$

für alle $v_1, v_2 \in V$ gibt.

Satz 8.14. Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum der Dimension $n < \infty$. Dann ist $(V, \langle \cdot, \cdot \rangle)$ isometrisch isomorph zu $\mathbb{R}^n$ mit dem Standardskalarprodukt.

Beweis. Sei $(e_1, \dots, e_n)$ die Standardbasis des $\mathbb{R}^n$. Dies ist eine Orthonormalbasis bezüglich des Standardskalarproduktes. Sei $(f_1, \dots, f_n)$ eine Orthonormalbasis von $(V, \langle \cdot, \cdot \rangle)$. Betrachte den Isomorphismus $\sigma : \mathbb{R}^n \rightarrow V$ definiert durch

$$\sigma(e_i) = f_i.$$

Es gilt

$$\langle \sigma(e_i), \sigma(e_j) \rangle = \langle f_i, f_j \rangle = \delta_{ij} = \langle e_i, e_j \rangle.$$

Daraus folgt, dass σ isometrisch ist. □

8.2 Unitäre Vektorräume

Auf dem Vektorraum $\mathbb{C}^n$ ist das Standardskalarprodukt durch

$$\begin{aligned} \langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n &\rightarrow \mathbb{C}, \\ \langle x, y \rangle &= x^t \cdot \bar{y} = \sum_{i=1}^n x_i \bar{y}_i, \end{aligned}$$

definiert. Diese Abbildung ist linear in der ersten und *semilinear* (auch antilinear oder halblinear genannt) in der zweiten Variable, wobei letzteres heißt, dass

$$\begin{aligned} \langle x + x', y \rangle &= \langle x, y \rangle + \langle x', y \rangle, \\ \langle \lambda x, y \rangle &= \lambda \langle x, y \rangle. \end{aligned}$$

Sie erfüllt $\langle x, y \rangle = \overline{\langle y, x \rangle}$ und ist positiv definit, das heißt $\langle x, x \rangle > 0$ für alle $x \in \mathbb{C}^n \setminus \{0\}$.

Wir wollen vergleichbare Abbildungen allgemeiner betrachten.

Definition 8.15. Sei V ein $\mathbb{C}$ -Vektorraum. Ein *hermitesches Skalarprodukt* auf V ist eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ mit

- i) Linearität: Für festes $y \in V$ ist $x \mapsto \langle x, y \rangle$ eine $\mathbb{C}$ -lineare Abbildung $V \rightarrow \mathbb{C}$.
- ii) Hermitezität: $\langle x, y \rangle = \overline{\langle y, x \rangle}$ für alle $x, y \in V$.
- iii) Positive Definitheit: Es gilt $\langle x, x \rangle > 0$ für alle $x \in V \setminus \{0\}$.

Das Paar $(V, \langle \cdot, \cdot \rangle)$ heißt *unitärer Vektorraum*.

- Für festes $x \in V$ ist $y \mapsto \langle x, y \rangle$ semilinear.
- Zwei Vektoren $x, y \in V$ heißen *orthogonal* zueinander, falls $\langle x, y \rangle = 0$. Schreibe $x \perp y$.
- Ist $W \subset V$ ein $\mathbb{C}$ -Untervektorraum, so heißt

$$W^\perp = \{x \in V \mid \langle x, w \rangle = 0 \text{ für alle } w \in W\}$$

das *orthogonale Komplement* von W .

Satz 8.16. *Jeder endlichdimensionale unitäre $\mathbb{C}$ -Vektorraum besitzt eine Orthonormalbasis.*

Beweis. Analog zu Bemerkung 8.5 und Korollar 8.6 □

Satz 8.17. *Sei $W \subset V$ ein Untervektorraum von V . Dann ist*

$$V = W \oplus W^\perp.$$

Beweis. Analog zu Satz 8.7 □

8.3 Normale Endomorphismen

Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler unitärer Vektorraum. (Das Skalarprodukt $\langle \cdot, \cdot \rangle$ ist $\mathbb{C}$ -linear im ersten Argument.)

Jedes $a \in V$ definiert eine lineare Abbildung

$$L_a : V \rightarrow \mathbb{C}, \quad x \mapsto \langle x, a \rangle.$$

Wir erhalten eine $\mathbb{C}$ -antilineare Abbildung

$$\ell : V \rightarrow V^*, \quad a \mapsto L_a.$$

Der folgende Satz zeigt, dass diese bijektiv ist.

Satz 8.18. *Sei $L : V \rightarrow \mathbb{C}$ eine lineare Abbildung. Es existiert ein eindeutig bestimmter Vektor $a \in V$, so dass*

$$L(x) = L_a(x) = \langle x, a \rangle$$

für alle $x \in V$.

Beweis. Sei $(e_1, \dots, e_n)$ eine Orthonormalbasis von V . Schreibe

$$a = \sum_{i=1}^n a_i e_i.$$

Es ist dann klar, dass die Abbildung L_a gegeben ist durch

$$L_a = \sum_{i=1}^n \overline{a_i} e_i^*,$$

wobei $(e_1^*, \dots, e_n^*)$ die Dualbasis bezeichne. Da $\sum_i a_i e_i \mapsto \sum_i a_i e_i^*$ einen Isomorphismus von V und V^* definiert, ist auch die antilineare Abbildung

$$a = \sum_i a_i e_i \mapsto L_a = \sum_i \overline{a_i} e_i^*$$

bijektiv. Explizit ist zu $L = \sum_{i=1}^n \lambda_i e_i^* \in V^*$ mit $\lambda_i \in \mathbb{C}$ der Vektor $a = \sum_{i=1}^n \overline{\lambda_i} e_i$ ein Urbild, d.h. $\ell(a) = L$. □

Satz 8.19. *Sei $f : V \rightarrow V$ ein Endomorphismus. Es existiert genau ein Endomorphismus $f^* : V \rightarrow V$ mit der Eigenschaft*

$$\langle f(a), b \rangle = \langle a, f^*(b) \rangle$$

für alle $a, b \in V$. Die Abbildung f^ heißt adjungierter Endomorphismus zu f .*

Beweis. Sei $f^T : V^* \rightarrow V^*$ die zu f duale Abbildung (die wir hier mal mit f^T bezeichnen, um Verwechslungen vorzubeugen). Sei $\ell : V \rightarrow V^*$ die $\mathbb{C}$ -antilineare Abbildung mit $\ell(a) = L_a$. Dann definieren wir $f^* = \ell^{-1} \circ f^T \circ \ell$. Da ℓ und ℓ^{-1} antilinear sind, ist f^* wieder linear und erfüllt

$$f^T \circ \ell = \ell \circ f^*,$$

d.h.

$$\langle a, f^*(b) \rangle = \ell(f^*(b))(a) = f^T(\ell(b))(a) = \ell(b)(f(a)) = \langle f(a), b \rangle,$$

wie gefordert. Hierbei haben wir die definierende Eigenschaft der dualen Abbildung benutzt, nämlich dass

$$f^T(\ell(b)) = \ell(b) \circ f$$

gilt. Die Eindeutigkeit folgt aus der Eindeutigkeit von f^T , denn wäre f' eine weitere Abbildung, welche

$$\langle f(a), b \rangle = \langle a, f'(b) \rangle$$

für alle $a, b \in V$ erfüllt, so erhalte man mit $g := \ell \circ f' \circ \ell^{-1}$ eine weitere Abbildung, die erfüllt $g(\ell(b)) = \ell(f'(b)) = \ell(b) \circ f$. Alternativ kann man auch folgern, dass

$$\langle a, (f' - f^*)(b) \rangle = 0$$

für alle $a, b \in V$ gilt und somit $f' - f^* = 0$. □

Bemerkung 8.20. Sei $E = (e_1, \dots, e_n)$ eine Orthonormalbasis von V , $f : V \rightarrow V$ ein Endomorphismus und $A = \mathcal{M}_E^E(f)$ seine darstellende Matrix. Dann ist die darstellende Matrix von f^* genau

$$A^* := \overline{A}^t.$$

Beweis. Es ist $A = (a_{ij})$ mit

$$f(e_j) = \sum_{i=1}^n a_{ij} e_i,$$

das heißt

$$a_{ij} = \langle e_i, f(e_j) \rangle.$$

Genauso ist $A^* = (a_{ij}^*)$ mit

$$a_{ij}^* = \langle e_i, f^*(e_j) \rangle.$$

Damit gilt

$$a_{ij}^* = \langle e_i, f^*(e_j) \rangle = \langle f(e_i), e_j \rangle = \overline{\langle e_j, f(e_i) \rangle} = \overline{a_{ji}}.$$

□

Definition 8.21. Ein Endomorphismus f von V heißt *normal*, falls er mit seinem Adjungierten vertauscht, das heißt

$$f \circ f^* = f^* \circ f.$$

Ein Endomorphismus f von V heißt *selbstadjungiert*, falls $f = f^*$ gilt.

Bemerkung 8.22. Ein Endomorphismus f von V ist also genau normal, wenn seine Darstellungsmatrix A bezüglich einer Orthonormalbasis die Gleichung

$$AA^* = A^*A$$

erfüllt. So eine Matrix nennt man ebenfalls normal.

Beispiel 8.23. • Selbstadjungierte Endomorphismen sind normal.

• $f : V \rightarrow V$ heißt *unitär*, falls

$$\langle a, b \rangle = \langle f(a), f(b) \rangle$$

für alle $a, b \in V$. In diesem Fall ist f invertierbar, denn aus $f(a) = 0$ folgt

$$0 = \langle f(a), f(a) \rangle = \langle a, a \rangle.$$

Dann muss bereits $a = 0$ gelten. (f ist also ein isometrischer Isomorphismus von V nach V .) Weiterhin gilt $f^* = f^{-1}$, denn

$$\langle a, f^*(b) \rangle = \langle f(a), b \rangle = \langle f(a), f \circ f^{-1}(b) \rangle = \langle a, f^{-1}(b) \rangle$$

für alle $a, b \in V$. Es folgt

$$f^* \circ f = \text{id}_V = f \circ f^*,$$

das heißt unitäre Endomorphismen sind normal.

Definition 8.24. Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt *hermitesch*, falls $\overline{A}^t = A$ gilt (d.h. $A^* = A$).

Bemerkung 8.25. Ein Endomorphismus $f : V \rightarrow V$ ist genau dann selbstadjungiert, wenn die darstellende Matrix A bezüglich einer Orthonormalbasis von V hermitesch ist. Hermitesche Matrizen sind normal.

Bemerkung 8.26. Ein Endomorphismus $f : V \rightarrow V$ ist genau dann unitär, wenn für die darstellende Matrix U von f bezüglich einer Orthonormalbasis von V die Gleichung $U^*U = E_n$ gilt. Dies ist äquivalent dazu, dass die Spalten von U eine Orthonormalbasis des $\mathbb{C}^n$ bzgl. des Standard-Skalarproduktes bilden.

Beweis. Übung. □

Definition 8.27. Die *unitäre Gruppe* $U(n, \mathbb{C})$ ist definiert als

$$U(n, \mathbb{C}) = \{U \in \text{GL}(n, \mathbb{C}) \mid U^*U = E_n\}.$$

Dies ist also die Gruppe der unitären Endomorphismen von $\mathbb{C}^n$ mit dem Standard-Skalarprodukt.

Lemma 8.28. *Sei $f : V \rightarrow V$ normal, $v \in V$ und $\lambda \in \mathbb{C}$ mit $f(v) = \lambda v$. Dann gilt $f^*(v) = \bar{\lambda}v$.*

Beweis. Falls $\lambda = 0$, dann ist $f(v) = 0$. Es ist zu zeigen, dass $f^*(v) = 0$.

$$0 = \langle f(v), f(v) \rangle = \langle f^* \circ f(v), v \rangle = \langle f \circ f^*(v), v \rangle = \langle f^*(v), f^*(v) \rangle$$

Dies beweist die Behauptung für $\lambda = 0$.

Sei nun $\lambda \in \mathbb{C}$ beliebig. Ersetze f durch $g = f - \lambda \text{id}_V$. Der Endomorphismus g ist ebenfalls normal mit $g^* = f^* - \bar{\lambda} \text{id}_V$ und erfüllt die Voraussetzungen des ersten Falles. Es folgt die Behauptung für f . $\square$

Satz 8.29 (Spektralsatz für normale Endomorphismen). *Zu jedem normalen Endomorphismus $f : V \rightarrow V$ gibt es eine Orthonormalbasis von V aus Eigenvektoren von f .*

Bemerkung 8.30. Normale Endomorphismen sind also insbesondere diagonalisierbar.

Beweis von Satz 8.29. Entscheidend ist:

- f hat einen Eigenvektor v mit $f(v) = \lambda v$ und $\|v\| = 1$ (da wir über $\mathbb{C}$ arbeiten).
- Setze $W = v^\perp \subset V$. Dann bildet f den Unterraum W auf sich selbst ab. Sei dazu $w \in W$. Es gilt

$$\langle f(w), v \rangle = \langle w, f^*(v) \rangle = \bar{\lambda} \langle w, v \rangle = 0,$$

das heißt $f(w) \in W$. Der Beweis folgt nun per Induktion nach $n = \dim(V)$: Ist $n = 1$ so sind wir fertig.

Andernfalls hat man $\dim(W) = n - 1$ und es ist also

- W ebenfalls ein unitärer Vektorraum mit der Einschränkung von $\langle \cdot, \cdot \rangle$ auf W ,
- und die Einschränkung $f|_W$ ein normaler Endomorphismus von W .

Damit W also nach Induktionsannahme eine Orthonormalbasis aus Eigenvektoren von $f|_W$. Zusammen mit v erhalten wir eine Orthonormalbasis von V .

$\square$

Es gilt auch die Umkehrung von Satz 8.29. Genauer haben wir:

Satz 8.31. *Sei $f : V \rightarrow V$ ein Endomorphismus. Es sind äquivalent:*

- f ist normal.
- V besitzt eine Orthonormalbasis aus Eigenvektoren von f .

Beweis. (i) $\Rightarrow$ (ii): Satz 8.29.

(ii) $\Rightarrow$ (i): Sei $B = (b_1, \dots, b_n)$ eine Orthonormalbasis von V aus Eigenvektoren und $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ die zugehörigen Eigenwerte von f , also

$$f(b_i) = \lambda_i b_i.$$

Es gilt

$$\langle b_j, f^*(b_i) \rangle = \langle f(b_j), b_i \rangle = \bar{\lambda}_j \delta_{ij},$$

woraus

$$f^*(b_i) = \bar{\lambda}_i b_i$$

folgt. Weiter gilt dann

$$f \circ f^*(b_i) = f(\bar{\lambda}_i b_i) = \bar{\lambda}_i f(b_i) = \bar{\lambda}_i \lambda_i b_i = \lambda_i \bar{\lambda}_i b_i = \lambda_i f^*(b_i) = f^*(\lambda_i b_i) = f^* \circ f(b_i),$$

das heißt $f \circ f^*$ und $f^* \circ f$ stimmen auf der Basis B überein. Es folgt $f \circ f^* = f^* \circ f$. $\square$

Satz 8.32. Sei $A \in \mathbb{C}^{n \times n}$. Es sind äquivalent:

i) A ist normal, das heißt $AA^* = A^*A$.

ii) Es gibt eine Matrix $U \in U(n, \mathbb{C})$ so dass $U^*AU = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ eine Diagonalmatrix ist. Die Einträge auf der Diagonalen sind die Eigenwerte von A .

Beweis. Folgt direkt aus Satz 8.31: Der erste Punkt besagt genau, dass die lineare Abbildung $f_A : \mathbb{C}^n \rightarrow \mathbb{C}^n, x \mapsto Ax$ normal ist. Nach Satz 8.31, ist dies äquivalent dazu, dass eine $\mathbb{C}^n$ eine Orthonormalbasis aus Eigenvektoren von f_A besitzt. Ist U die Basiswechselmatrix, so ist dies eine unitäre Matrix, denn die Spalten von U sind eine Orthonormalbasis von $\mathbb{C}^n$ (die Eigenvektoren von f_A). Somit ist $U^{-1} = U^*$ und es ist also U^*AU eine Diagonalmatrix mit den Eigenwerten auf der Diagonalen. $\square$

Bemerkung 8.33. Sei $f : V \rightarrow V$ normal. Es gilt:

i) f ist selbstadjungiert $\Leftrightarrow$ alle Eigenwerte sind reell.

ii) f ist unitär $\Leftrightarrow$ alle Eigenwerte haben Betrag 1.

Beweis. i): Sei $B = (b_1, \dots, b_n)$ eine Orthonormalbasis von V aus Eigenvektoren von f mit Eigenwerten $\lambda_1, \dots, \lambda_n$. Es gilt

$$\begin{aligned} & f \text{ ist selbstadjungiert} \\ \Leftrightarrow & \langle f(b_i), b_j \rangle = \langle b_i, f(b_j) \rangle \text{ für alle } i, j \\ \Leftrightarrow & \lambda_i \langle b_i, b_j \rangle = \bar{\lambda}_j \langle b_i, b_j \rangle \text{ für alle } i, j \\ \Leftrightarrow & \lambda_i = \bar{\lambda}_i \text{ für alle } i. \end{aligned}$$

ii): Übung. $\square$

Korollar 8.34. *Hermiteische Matrizen sind diagonalisierbar. Alle ihre Eigenwerte sind reell.*

Beweis. Hermiteische Matrizen sind normal, denn für sie gilt $A^* = A$. Damit sind sie diagonalisierbar. Mit der vorigen Bemerkung sind die Eigenwerte alle reell. $\square$

Bemerkung 8.35. Sei f ein normaler Endomorphismus eines unitären Vektorraums. Sind v und w Eigenvektoren zu den Eigenwerten $\lambda \neq \mu$, so gilt $v \perp w$.

Beweis. Dies folgt nun klar aus Satz 8.29. Man kann dies aber auch direkt einsehen: Es ist einerseits

$$\langle f(v), w \rangle = \langle \lambda v, w \rangle = \lambda \langle v, w \rangle$$

und andererseits

$$\langle f(v), w \rangle = \langle v, f^*(w) \rangle = \langle v, \bar{\mu} w \rangle = \bar{\mu} \langle v, w \rangle.$$

Zusammen ergibt sich $\langle v, w \rangle = 0$, falls $\lambda \neq \mu$ ist. $\square$

Satz 8.36. *Sei $f : V \rightarrow V$ ein Endomorphismus endlicher Ordnung (das heißt es existiert ein $m \in \mathbb{N}$ mit $f^m = \text{id}_V$). Dann ist f diagonalisierbar.*

Beweis. Definiere ein neues Skalarprodukt auf V durch

$$[\cdot, \cdot] : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}, \quad [a, b] = \sum_{i=1}^m \langle f^i(a), f^i(b) \rangle.$$

Das Skalarprodukt $[\cdot, \cdot]$ ist auch hermitesch. Der Endomorphismus f ist unitär bezüglich dieses Skalarproduktes, denn

$$[f(a), f(b)] = \sum_{i=1}^m \langle f^{i+1}(a), f^{i+1}(b) \rangle = \sum_{i=2}^{m+1} \langle f^i(a), f^i(b) \rangle = [a, b].$$

Hier haben wir benutzt, dass $\langle f^{m+1}(a), f^{m+1}(b) \rangle = \langle f(a), f(b) \rangle$. Mit Satz 8.29 folgt, dass $(V, [\cdot, \cdot])$ eine Orthonormalbasis von Eigenvektoren von f besitzt. $\square$

Bemerkung 8.37. Sei $f : V \rightarrow V$ ein Endomorphismus der Ordnung m (das heißt $f^m = \text{id}_V$). Dann sind die Eigenwerte von f m -te Einheitswurzeln (das heißt $\lambda^m = 1$, also $\lambda = e^{2\pi i b/m}$ für ein $b \in \mathbb{Z}$).

Beweis. Übung. $\square$

8.4 Anwendung: Selbstadjungierte Endomorphismen euklidischer Vektorräume

Lemma 8.38. *Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum. Sei $f : V \rightarrow V$ ein selbstadjungierter Endomorphismus von V , d.h. $f = f^*$. Sei A die darstellende Matrix von f bezüglich einer Orthonormalbasis von V . Dann gilt $A^t = A$, d.h. A ist symmetrisch.*

Beweis. Dies folgt direkt aus Satz 7.33, da die Strukturmatrix des Skalarproduktes bezüglich einer Orthonormalbasis von V die Einheitsmatrix ist (also $B = E_n$ in Satz 7.33). Es ist also die Darstellungsmatrix von f^* durch A^t gegeben, da aber $f^* = f$ gilt, so folgt $A^* = A$. $\square$

Satz 8.39. *Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum. Sei $f : V \rightarrow V$ ein selbstadjungierter Endomorphismus von V , d.h. $f = f^*$. Dann ist f diagonalisierbar (über $\mathbb{R}$).*

Beweis. Ist f selbstadjungiert, so ist die Darstellungsmatrix A von f bezüglich einer Orthonormalbasis symmetrisch. Da A nur reelle Einträge hat, gilt insbesondere $\overline{A}^t = A$ mit Lemma 8.38. Damit ist A also hermitesch und alle Eigenwerte von A sind reell. Also zerfällt das charakteristische Polynom von A schon über $\mathbb{R}$. Man kann jetzt den exakt gleichen induktiven Beweis wie bei Satz 8.29 durchführen. $\square$

Korollar 8.40. *Jede symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ ist diagonalisierbar.*

9 Quadratische Formen

9.1 Quadratische Formen

Sei K ein Körper.

Beispiel 9.1. In Dimension 1: Sei $c \in K$ fest. Dann ist die Abbildung

$$q : K \rightarrow K, \quad x \mapsto cx^2$$

eine quadratische Form.

Allgemeiner in Dimension n : Seien $c_{ij} \in K$ für $i, j = 1, \dots, n$ gegeben. Betrachte

$$q : K^n \rightarrow K, \quad x \mapsto \sum_{i,j=1}^n c_{ij}x_i x_j.$$

Fragen.

- Welche Werte nimmt q an?
- Klassifizierung (bis auf Isometrie)?
- Invarianten von q ?

Die Antworten hängen stark von K ab. „Schwer“ für $K = \mathbb{Q}$, „einfacher“ für $K = \mathbb{R}, \mathbb{C}$.

Definition 9.2. Eine quadratische Form auf V ist eine Abbildung $q : V \rightarrow K$, mit den Eigenschaften:

- i) Es gilt $q(ax) = a^2q(x)$ für alle $a \in K$ und $x \in V$.

ii) Die Abbildung

$$\beta_q : V \times V \rightarrow K, \quad (x, y) \mapsto q(x + y) - q(x) - q(y)$$

ist eine Bilinearform auf V .

Bemerkung 9.3. Sei $\beta = \beta_q$ die Bilinearform zu q . Für $x \in V$ gilt

$$\beta(x, x) = q(x + x) - q(x) - q(x) = q(2x) - 2q(x) = 2q(x).$$

Ist $\text{char}(K) = 2$, so folgt $\beta(x, x) = 0$ für alle $x \in V$.

Ist $\text{char}(K) \neq 2$, so folgt $q(x) = \frac{1}{2}\beta(x, x)$ für alle $x \in V$. Die quadratische Form q ist in diesem Fall eindeutig durch die Bilinearform β bestimmt und (i) folgt aus (ii).

Für den Rest des Kapitels sei $\text{char}(K) \neq 2$.

Bemerkung 9.4.

- i) β_q ist eine symmetrische Bilinearform.
- ii) Die Zuordnung $q \mapsto \beta_q$ definiert eine Bijektion zwischen der Menge der quadratischen Formen auf V und der Menge der symmetrischen Bilinearformen auf V .

Viele Begriffe der Theorie der Bilinearformen übertragen sich auf quadratische Formen.

Definition 9.5. Sei $\beta = \beta_q$.

- q heißt nicht ausgeartet, falls β nicht ausgeartet ist.
- Ist $b_1, \dots, b_n$ eine Basis von V , so heißt

$$B = \begin{pmatrix} \beta(b_1, b_1) & \dots & \beta(b_1, b_n) \\ \vdots & & \vdots \\ \beta(b_n, b_1) & \dots & \beta(b_n, b_n) \end{pmatrix}$$

Strukturmatrix oder *Gram-Matrix* von q bezüglich $b_1, \dots, b_n$. Es gilt $B = B^t$.

- Ist $X \subset V$ eine Teilmenge, so ist $X^\perp$ bezüglich β definiert.

Bemerkung 9.6. Sei $q : V \rightarrow K$ eine quadratische Form mit Strukturmatrix B bezüglich einer Basis $b_1, \dots, b_n$. Dann ist q genau dann nicht ausgeartet, wenn B invertierbar ist.

Beweis. Es gilt

$$q \text{ ist nicht ausgeartet} \quad \Leftrightarrow \quad V^\perp = \{0\}.$$

Für $x = \sum_{j=1}^n x_j b_j \in V$ gilt

$$\begin{aligned}
 x \in V^\perp &\Leftrightarrow \beta(b_i, x) = 0 \text{ für alle } i = 1, \dots, n \\
 &\Leftrightarrow \sum_{j=1}^n \beta(b_i, b_j) x_j = 0 \text{ für alle } i = 1, \dots, n \\
 &\Leftrightarrow B \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0 \\
 &\Leftrightarrow \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \text{Kern}(B).
 \end{aligned}$$

Damit ist q genau dann nicht ausgeartet, wenn der Kern von B trivial ist. Es folgt die Behauptung. $\square$

Bemerkung 9.7. Sei $q : V \rightarrow K$ eine quadratische Form. Sei $W \subset V$ ein Untervektorraum, so dass $V = V^\perp \oplus W$. Dann ist $q|_W : W \rightarrow K$ eine nicht ausgeartete quadratische Form.

Beweis. Übung. $\square$

9.2 Orthogonalbasen und der Kürzungssatz von Witt

Sei K ein Körper.

Definition 9.8. Ein *quadratischer Raum* über K ist ein Paar (V, q) bestehend aus einem K -Vektorraum V und einer quadratischen Form $q : V \rightarrow K$.

Definition 9.9. Eine *Isometrie* zwischen zwei quadratischen Räumen (V, q) und (V', q') ist eine lineare Abbildung $s : V \rightarrow V'$ mit der Eigenschaft $q'(s(x)) = q(x)$ für alle $x \in V$. Zwei quadratische Räume (V, q) und (V', q') heißen *äquivalent*, wenn es einen isometrischen Isomorphismus $s : V \rightarrow V'$ gibt. In diesem Fall schreiben wir $(V, q) \simeq (V', q')$.

Bemerkung 9.10. Ist $s : (V, q) \rightarrow (V', q')$ eine Isometrie, so gilt für die zugehörigen Bilinearformen

$$\beta'(s(x), s(y)) = \beta(x, y)$$

für alle $x, y \in V$.

Im Folgenden nehmen wir stets $\text{char}(K) \neq 2$ an.

Satz 9.11. Sei (V, q) ein quadratischer Raum über K . Dann besitzt V eine Orthogonalbasis $b_1, \dots, b_n$, das heißt eine Basis mit

$$\beta(b_i, b_j) = 0$$

für alle $i \neq j$. Setzt man $a_i = \beta(b_i, b_i)$, so gilt

$$q\left(\sum_{i=1}^n x_i b_i\right) = \frac{1}{2} \sum_{i=1}^n a_i x_i^2.$$

Beweis. Induktion nach $n = \dim(V)$.

$n = 1$: Wähle einen beliebigen Vektor $b_1 \in V \setminus \{0\}$.

$n > 1$: Ist $q(v) = 0$ für alle $v \in V$, so ist jede Basis eine Orthogonalbasis und wir sind fertig. Angenommen es gibt ein $b \in V$ mit $q(b) \neq 0$. Sei

$$W = b^\perp = \{w \in V \mid \beta(b, w) = 0\}.$$

Die Dimensionsformel für orthogonale Komplemente (Satz 7.36) liefert

$$\underbrace{\dim(Kb)}_{=1} + \dim(W) = \dim(V) + \underbrace{\dim(V^\perp \cap Kb)}_{=0 \text{ wegen } q(b) \neq 0}.$$

Daher ist $(W, q|_W)$ ein quadratischer Raum der Dimension $n - 1$. (Alternativ sieht man dies, indem man W als Kern der linearen Abbildung $V \rightarrow K$, $x \mapsto \beta(b, x)$ auffasst.) Nach der Induktionsannahme hat W eine Orthogonalbasis $b_2, \dots, b_n$. Die Vektoren $b, b_2, \dots, b_n$ bilden dann eine Orthogonalbasis von V . $\square$

Bemerkung 9.12. Bezüglich einer Orthogonalbasis $b_1, \dots, b_n$ hat die Strukturmatrix Diagonalgestalt

$$B = \begin{pmatrix} \beta(b_1, b_1) & & \\ & \ddots & \\ & & \beta(b_n, b_n) \end{pmatrix}.$$

Definition 9.13. Ist $A \in K^{n \times n}$ eine symmetrische Matrix, so schreiben wir q_A für die quadratische Form

$$q_A : K^n \rightarrow K, \quad x \mapsto \frac{1}{2} x^t A x.$$

Sind $a_1, \dots, a_n \in K$, so schreiben wir $[a_1, \dots, a_n]$ für die quadratische Form zu

$$A = 2 \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix},$$

also $q(x) = a_1 x_1^2 + \dots + a_n x_n^2$. Solche Formen heißen *Diagonalformen*.

Bemerkung 9.14. Ist (V, q) ein quadratischer Raum und A die Gram-Matrix bezüglich einer Basis $b_1, \dots, b_n$, so ist

$$s : (V, q) \rightarrow (K^n, q_A), \quad b_i \mapsto e_i$$

eine isometrische Isomorphie. Die quadratischen Räume (V, q) und (K^n, q_A) sind damit äquivalent. Satz 9.11 besagt dann, dass jede quadratische Form zu einer Diagonalform äquivalent ist.

Achtung. Die in der Diagonalform $[a_1, \dots, a_n]$ auftretenden Koeffizienten sind durch q *nicht* eindeutig bestimmt.

Frage. Wann sind zwei Diagonalformen $[a_1, \dots, a_n]$, $[b_1, \dots, b_n]$ auf K^n äquivalent über K ? Äquivalent: Wann gibt es ein $S \in \text{GL}(n, K)$ mit

$$\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} = S^t \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_n \end{pmatrix} S?$$

Antwort. Dies hängt von K ab. „Einfach“ über $\mathbb{R}$ oder $\mathbb{C}$ (siehe unten), „schwierig“ über $\mathbb{Q}$ (Satz von Hasse-Minkowski).

Über $\mathbb{C}$:

Bemerkung 9.15. Sei q eine nicht ausgeartete quadratische Form auf dem $\mathbb{C}$ -Vektorraum V . Dann gibt es eine Basis $b_1, \dots, b_n$, so dass die Strukturmatrix

$$B = 2E_n$$

ist, das heißt q ist äquivalent zur Einheitsform $[1, 1, \dots, 1]$ auf $\mathbb{C}^n$.

Beweis. Sei $c_1, \dots, c_n$ eine Orthogonalbasis von V und $\gamma_i \in \mathbb{C}$ mit $\gamma_i^2 = q(c_i) = \frac{1}{2}\beta(c_i, c_i)$. Die Existenz solcher γ_i folgt aus der algebraischen Abgeschlossenheit von $\mathbb{C}$. Es gilt $\gamma_i \neq 0$ da q nicht ausgeartet ist. Setze $b_i = \frac{1}{\gamma_i}c_i$. Dann ist $b_1, \dots, b_n$ eine Orthogonalbasis mit $q(b_i) = \frac{1}{\gamma_i^2}q(c_i) = 1$ für $i = 1, \dots, n$. $\square$

Satz 9.16 (Kürzungssatz von Witt). Seien $k, n \in \mathbb{N}$ mit $1 \leq k < n$. Seien $a_1, \dots, a_n \in K^\times$ und $b_1, \dots, b_n \in K^\times$. Falls

$$\begin{cases} [a_1, \dots, a_k, a_{k+1}, \dots, a_n] \simeq [b_1, \dots, b_k, b_{k+1}, \dots, b_n] & \text{und} \\ [a_1, \dots, a_k] \simeq [b_1, \dots, b_k], \end{cases} \quad (*)$$

so folgt

$$[a_{k+1}, \dots, a_n] \simeq [b_{k+1}, \dots, b_n].$$

Beweis. Schritt 1: Aus $(*_1)$ folgt direkt

$$[a_1, \dots, a_k, a_{k+1}, \dots, a_n] \simeq [a_1, \dots, a_k, b_{k+1}, \dots, b_n], \quad (*_2)$$

denn wenn es Matrizen $S \in \text{GL}(n, K)$ und $T_0 \in \text{GL}(k, K)$ mit

$$S^t \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} S = \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_n \end{pmatrix}$$

und

$$T_0^t \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_k \end{pmatrix} T_0 = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_k \end{pmatrix}$$

so folgt mit $T = \begin{pmatrix} T_0 & \\ & E_{n-k} \end{pmatrix}$, dass

$$(ST)^t \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} ST = \begin{pmatrix} a_1 & & & & \\ & \ddots & & & \\ & & a_k & & \\ & & & b_{k+1} & \\ & & & & \ddots \\ & & & & & b_n \end{pmatrix}.$$

Schritt 2: Aus $(*_2)$ folgt die Behauptung, dies zeigen wir per Induktion nach k .

Sei $k = 1$ (und n beliebig). Dann bedeutet $(*_2)$, dass es ein $T \in \text{GL}(n, K)$ mit

$$T^t \begin{pmatrix} a_1 & & \\ & a_2 & \\ & & \ddots \\ & & & a_n \end{pmatrix} T = \begin{pmatrix} a_1 & & \\ & b_2 & \\ & & \ddots \\ & & & b_n \end{pmatrix}$$

gibt. Schreibe

$$A = \begin{pmatrix} a_2 & & \\ & \ddots & \\ & & a_n \end{pmatrix}, \quad B = \begin{pmatrix} b_2 & & \\ & \ddots & \\ & & b_n \end{pmatrix}, \quad T = \left(\begin{array}{c|c} c & v^t \\ \hline u & S \end{array} \right)$$

mit $c \in K$, $u, v \in K^{n-1}$ und $S \in K^{(n-1) \times (n-1)}$. Dann gilt

$$\begin{aligned} \begin{pmatrix} a_1 & 0 \\ 0 & B \end{pmatrix} &= \begin{pmatrix} c & u^t \\ v & S^t \end{pmatrix} \begin{pmatrix} a_1 & 0 \\ 0 & A \end{pmatrix} \begin{pmatrix} c & v^t \\ u & S \end{pmatrix} \\ &= \begin{pmatrix} ca_1 & u^t A \\ a_1 v & S^t A \end{pmatrix} \begin{pmatrix} c & v^t \\ u & S \end{pmatrix} \\ &= \begin{pmatrix} c^2 a_1 + u^t A u & ca_1 v^t + u^t A S \\ ca_1 v + S^t A u & a_1 v v^t + S^t A S \end{pmatrix}. \end{aligned}$$

Dies ist äquivalent zu

$$\begin{cases} u^t Au = (1 - c^2)a_1, \\ S^t Au = -ca_1 v, \\ S^t AS + a_1 v v^t = B. \end{cases}$$

Gesucht: Eine Matrix $M \in K^{(n-1) \times (n-1)}$ mit

$$M^t AM = B. \quad (*_3)$$

Ist $(*_3)$ für ein $M \in K^{(n-1) \times (n-1)}$ erfüllt, so ist $M \in \text{GL}(n-1, K)$, da $\det(M)^2 = \det(A)^{-1} \det(B) \in K^\times$.

Ansatz: Sei $M = S + \lambda uv^t$ mit $\lambda \in K$. Dann gilt

$$\begin{aligned} (*_3) &\Leftrightarrow (S^t + \lambda v u^t) A (S + \lambda u v^t) = B \\ &\Leftrightarrow \underbrace{S^t AS}_{=B-a_1 v v^t} + \lambda v \underbrace{u^t AS}_{=-ca_1 v^t} + \lambda \underbrace{S^t Au}_{=-ca_1 v} v^t + \lambda^2 v \underbrace{u^t Au}_{=(1-c^2)a_1} v^t = B \\ &\Leftrightarrow B - a_1 v v^t - 2\lambda ca_1 v v^t + \lambda^2 (1 - c^2) a_1 v v^t = B \\ &\Leftrightarrow a_1 (-1 - 2\lambda c + \lambda^2 (1 - c^2)) = 0 \quad \text{oder} \quad v = 0 \\ &\Leftrightarrow (c^2 - 1)\lambda^2 + 2c\lambda + 1 = 0. \end{aligned}$$

Für $c^2 = 1$ ist die Gleichung $(c^2 - 1)\lambda^2 + 2c\lambda + 1 = 0$ lösbar. Für $c^2 \neq 1$ gilt

$$\begin{aligned} &(c^2 - 1)\lambda^2 + 2c\lambda + 1 = 0 \\ &\Leftrightarrow \lambda^2 + \frac{2c}{c^2 - 1}\lambda + \frac{1}{c^2 - 1} = 0 \\ &\Leftrightarrow \left(\lambda + \frac{c}{c^2 - 1}\right)^2 - \frac{c^2}{(c^2 - 1)^2} + \frac{1}{c^2 - 1} = 0 \\ &\Leftrightarrow \left(\lambda + \frac{c}{c^2 - 1}\right)^2 - \frac{c^2}{(c^2 - 1)^2} + \frac{c^2 - 1}{(c^2 - 1)^2} = 0 \\ &\Leftrightarrow \left(\lambda + \frac{c}{c^2 - 1}\right)^2 = \frac{1}{(c^2 - 1)^2} \\ &\Leftrightarrow \lambda + \frac{c}{c^2 - 1} = \pm \frac{1}{c^2 - 1} \\ &\Leftrightarrow \lambda = -\frac{c}{c^2 - 1} \pm \frac{1}{c^2 - 1} = -\frac{c \mp 1}{c^2 - 1} = \frac{-1}{c \pm 1}. \end{aligned}$$

Dies liefert auch im Fall $c^2 \neq 1$ eine Lösung der Gleichung $(c^2 - 1)\lambda^2 + 2c\lambda + 1 = 0$. Es gibt daher in beiden Fällen eine Matrix M welche $(*_3)$ erfüllt, dies zeigt Behauptung für $k = 1$.

Sei nun $k > 1$ und die Behauptung für $k - 1$ (und alle n) bewiesen. Nach Schritt 1 gilt

$$[a_1, \dots, a_k, a_{k+1}, \dots, a_n] \simeq [a_1, \dots, a_k, b_{k+1}, \dots, b_n], \quad (*_2)$$

woraus nach der Induktionsannahme

$$[a_k, a_{k+1}, \dots, a_n] \simeq [a_k, b_{k+1}, \dots, b_n]$$

folgt. Wie im Fall $k = 1$ folgt hieraus

$$[a_{k+1}, \dots, a_n] \simeq [b_{k+1}, \dots, b_n],$$

die Behauptung ist somit insgesamt bewiesen. $\square$

9.3 Die orthogonale Gruppe

Definition 9.17. Sei (V, q) ein quadratischer Raum. Ein Vektor $v \in V$ heißt *isotrop*, falls $q(v) = 0$. Er heißt *anisotrop*, falls $q(v) \neq 0$. Die quadratische Form q heißt isotrop, falls V einen isotropen Vektor $v \neq 0$ enthält. Ein Teilraum $U \subset V$ heißt *total isotrop*, falls $q(v) = 0$ für alle $v \in U$.

Beispiel 9.18.

- i) Sei $A = (a_{ij}) \in K^{n \times n}$ eine symmetrische Matrix. Dann ist die quadratische Form $q_A : K^n \rightarrow K$ genau dann isotrop, wenn die Gleichung

$$\sum_{i,j=1}^n a_{ij} x_i x_j = 0$$

eine nicht-triviale Lösung hat.

- ii) Die Diagonalform $q = [1, -1]$ auf K^2 ist nicht ausgeartet und isotrop, denn

$$q\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = 0.$$

Bemerkung 9.19. Es gibt keinen nicht ausgearteten isotropen quadratischen Raum der Dimension 1.

Sei nun (V, q) ein nicht ausgearteter quadratischer Raum über einem Körper K mit Charakteristik ungleich 2. Für $x, y \in V$ misst $q(x - y) \in K$ den „Abstand“ zwischen x und y . Wir interessieren uns für Abbildungen, die diesen Abstand invariant lassen.

Satz 9.20. Sei $f : V \rightarrow V$ eine distanztreue Abbildung, das heißt es gelte

$$q(f(x) - f(y)) = q(x - y) \tag{*_1}$$

für alle $x, y \in V$. Dann ist f eine affine Abbildung, das heißt es existieren $v \in V$ und eine lineare Abbildung $s : V \rightarrow V$, so dass

$$f(x) = s(x) + v$$

für alle $x \in V$. Hierbei sind die Abbildung s und der Vektor v eindeutig bestimmt und s ist eine Isometrie.

Beweis. Schritt 1: Setze $v := f(0)$ und definiere $s : V \rightarrow V$, $s(x) := f(x) - v$. Es gilt dann $s(0) = 0$.

Schritt 2: Behauptung: Die Abbildung s ist linear.

Schritt 2.1: Aus $(*)_1$ folgt, dass s auch distanztreu ist, also

$$q(s(x) - s(y)) = q(f(x) - f(y)) = q(x - y)$$

für alle $x, y \in V$. Mit $y = 0$ folgt daraus

$$q(s(x)) = q(x)$$

für alle $x \in V$. Es gilt dann

$$\begin{aligned} q(s(x) - s(y)) &= q(s(x)) - \beta(s(x), s(y)) + q(s(y)) \\ \Rightarrow -\beta(s(x), s(y)) &= q(x - y) - q(x) - q(y) = \beta(x, -y) \\ \Rightarrow \beta(s(x), s(y)) &= \beta(x, y). \end{aligned} \quad (*)_2$$

Schritt 2.2: Behauptung: Es gilt

$$s(ax + by) - as(x) - bs(y) \perp s(z)$$

für alle $x, y, z \in V$ und $a, b \in K$.

In der Tat:

$$\begin{aligned} &\beta(s(ax + by) - as(x) - bs(y), s(z)) \\ &= \beta(s(ax + by), s(z)) - a\beta(s(x), s(z)) - b\beta(s(y), s(z)) \\ &= \beta(ax + by, z) - a\beta(x, z) - b\beta(y, z) \\ &= 0 \end{aligned}$$

für alle $x, y, z \in V$ und $a, b \in K$.

Schritt 2.3: Behauptung: Das Bild $s(V)$ der Abbildung s erzeugt V .

Sei dazu $b_1, \dots, b_n$ eine Orthogonalbasis von V . Es gilt

$$\beta(s(b_i), s(b_j)) = \beta(b_i, b_j) = \begin{cases} 0, & i \neq j \\ \neq 0, & i = j. \end{cases}$$

Hieraus folgt, dass $s(b_1), \dots, s(b_n)$ eine Orthogonalbasis von V ist, insbesondere gilt

$$\text{Lin}(s(V)) = V.$$

Schritt 2.4: Aus den Schritten 2.2 und 2.3 folgt (da q nicht ausgeartet ist)

$$s(ax + by) - as(x) - bs(y) \in s(V)^\perp = \text{Lin}(s(V))^\perp = V^\perp = \{0\}$$

für alle $x, y \in V$ und $a, b \in K$, das heißt die Abbildung s ist linear.

Schritt 3: $(*)_2$ zeigt, dass s eine Isometrie ist. Der Vektor v ist eindeutig als $f(0)$ bestimmt. Die Abbildung $s : V \rightarrow V$ ist dann durch $s(x) = f(x) - v$ auch eindeutig bestimmt. $\square$

Definition 9.21. Sei (V, q) ein quadratischer Raum. Die Gruppe

$$\mathrm{O}(V, q) := \{s : V \rightarrow V \mid s \text{ ist eine Isometrie}\} \subset \mathrm{GL}(V)$$

heißt *orthogonale Gruppe* von (V, q) . Die *spezielle orthogonale Gruppe* von (V, q) ist

$$\mathrm{SO}(V, q) := \{s \in \mathrm{O}(V, q) \mid \det(s) = 1\} \subset \mathrm{O}(V, q).$$

Beispiel 9.22. Für $V = K^n$ mit $q = q_{2E} = [1, \dots, 1]$ schreibt man einfach

$$\begin{aligned} \mathrm{O}(n, K) &:= \mathrm{O}(K^n, q), \\ \mathrm{SO}(n, K) &:= \mathrm{SO}(K^n, q). \end{aligned}$$

Bemerkung 9.23. Es gilt

$$\begin{aligned} \mathrm{O}(n, K) &= \{S \in K^{n \times n} \mid S^t S = E_n\}, \\ \mathrm{SO}(n, K) &= \{S \in K^{n \times n} \mid S^t S = E_n \text{ und } \det(S) = 1\}. \end{aligned}$$

Bemerkung 9.24. Sei $A \in K^{n \times n}$ symmetrisch. Dann gilt

$$\mathrm{O}(K^n, q_A) = \{S \in \mathrm{GL}(n, K) \mid S^t A S = A\}.$$

Bemerkung 9.25. Sei (V, q) ein nicht ausgearteter quadratischer Raum und $s : V \rightarrow V$ eine Isometrie. Dann gilt

$$\det(s) = \pm 1.$$

Beweis. Es genügt $(V, q) = (K^n, q_A)$ zu betrachten, wobei $\det(A) \neq 0$. Es folgt

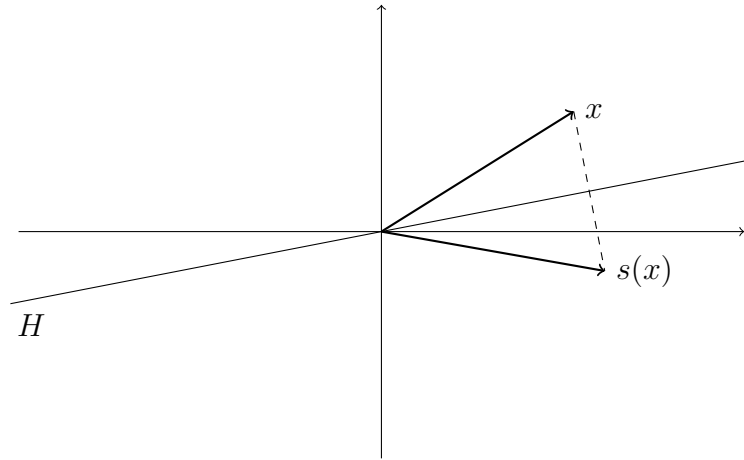
$$\begin{aligned} \det(S^t A S) &= \det(A) \\ \Rightarrow \det(S)^2 \cdot \det(A) &= \det(A) \\ \Rightarrow \det(S)^2 &= 1. \end{aligned}$$

□

Beispiele von Isometrien erhält man durch Spiegelungen. Sei (V, q) ein nicht ausgearteter quadratischer Raum der Dimension n . Sei $H \subset V$ eine Hyperebene, also ein $(n-1)$ -dimensionaler Teilraum. Sei H nicht ausgeartet. Dann ist $V = H \oplus H^\perp$. Sei $s = s_H$ die lineare Abbildung $s : V \rightarrow V$, definiert durch

$$\begin{aligned} s(x) &= x \quad \text{für alle } x \in H, \\ s(y) &= -y \quad \text{für alle } y \in H^\perp. \end{aligned}$$

Die Abbildung s heißt *Spiegelung* an der Hyperebene H .



- Es gilt $s^2 = \text{id}_V$.
- Die Spiegelung s ist eine Isometrie von (V, q) .
- Falls $\dim(V) = 1$, so ist $H = \{0\}$, $H^\perp = V$, $s_H = -\text{id}_V$ zugelassen.

Bemerkung 9.26. Sei $s : V \rightarrow V$ eine Isometrie, so dass

$$H = \{x \in V \mid s(x) = x\} \subset V$$

eine nicht ausgeartete Hyperebene ist. Dann ist $s = s_H$ die Spiegelung an H .

Beweis. Es ist $\dim(H^\perp) = 1$. Außerdem gilt $s(H^\perp) \subset H^\perp$, denn für $y \in H^\perp$ und $x \in H$ ist

$$\beta(s(y), x) = \beta(s(y), s(x)) = \beta(y, x) = 0.$$

Für $0 \neq y \in H^\perp$ ist daher $s(y) = \lambda y$ mit $\lambda \in K$. Es ist $q(y) \neq 0$ da (V, q) nicht ausgeartet ist. Daraus folgt

$$\begin{aligned} 0 \neq q(y) &= q(s(y)) = q(\lambda y) = \lambda^2 q(y) \\ \Rightarrow \lambda^2 &= 1 \\ \stackrel{s \neq \text{id}_V}{\Rightarrow} \lambda &= -1 \\ \Rightarrow s &= s_H. \end{aligned}$$

□

Bemerkung 9.27. Seien $V, H, s = s_H$ wie oben. Sei $b_1, \dots, b_{n-1}$ eine Basis von H und b_n eine Basis von $H^\perp$. Dann ist $b_1, \dots, b_n$ eine Basis von V und s hat die darstellende Matrix

$$\begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & -1 \end{pmatrix}.$$

Insbesondere ist $\det(s) = -1$.

Bemerkung 9.28. Seien $V, H, s = s_H$ wie oben. Sei $0 \neq y_0 \in H^\perp$. Dann gilt

$$s(v) = v - 2 \frac{\beta(v, y_0)}{\beta(y_0, y_0)} y_0$$

für alle $v \in V$.

Beweis. Übung. □

Erinnerung. Die Symmetrische Gruppe n -ten Grades S_n wird durch Transpositionen erzeugt.

Ziel. Wir zeigen nun, dass die Orthogonale Gruppe durch Spiegelungen erzeugt wird.

Zum Beweis benötigen wir das folgende Lemma.

Lemma 9.29. *Sei $s : V \rightarrow V$ eine Isometrie. Sei $v \in V$ ein anisotroper Vektor, so dass auch $s(v) - v$ anisotrop ist. Dann existiert eine Spiegelung $s_0 : V \rightarrow V$, so dass der Vektor v ein Fixpunkt von $s_0 \circ s$ ist.*

Beweis. Sei $H \subset V$ die Hyperebene $(s(v) - v)^\perp$. Da $s(v) - v$ anisotrop ist, ist H nicht ausgeartet und

$$V = H \oplus K(s(v) - v).$$

Sei $s_0 = s_H$. Dann gilt

$$\beta(s(v) + v, s(v) - v) = \beta(s(v), s(v)) - \beta(v, v) = 0,$$

das heißt $s(v) + v \in H$. Hieraus folgt

$$s_0(s(v) + v) = s(v) + v,$$

andererseits gilt

$$s_0(s(v) - v) = -s(v) + v.$$

Addiert man diese Gleichungen, so erhält man

$$2(s_0 \circ s)(v) = 2v.$$

Da $\text{char}(K) \neq 2$ folgt $(s_0 \circ s)(v) = v$. □

Satz 9.30. *Sei (V, q) ein nicht ausgearteter quadratischer Raum der Dimension n . Dann ist jede Isometrie $s \in O(V, q)$ ein Produkt von höchstens n Spiegelungen.*

Beweis. Wir beweisen hier nur den Fall, dass (V, q) anisotrop ist durch Induktion nach $\dim(V)$. Für den allgemeinen Fall siehe [4] Satz 4, Seite 56.

Fall 1: Angenommen es existiert ein $v \in V$ mit $q(v) \neq 0$ und $s(v) = v$. Sei $H = v^\perp \subset V$. Dann ist $V = H \oplus Kv$ und $\dim(H) = n - 1$. Es gilt $s(H) \subset H$ und $s'|_H$ ist eine

Isometrie des nicht ausgearteten quadratischen Raums (H, q_H) . Per Induktion existieren Spiegelungen $t_1, \dots, t_r \in O(H, q_H)$, mit $r \leq n - 1$, an Hyperebenen von H , so dass

$$s' = t_1 \circ \dots \circ t_r.$$

Definiere lineare Abbildungen $s_i : V \rightarrow V$ für $i = 1, \dots, r$ durch

$$\begin{aligned} s_i(x) &= t_i(x) \text{ für alle } x \in H, \\ s_i(v) &= v. \end{aligned}$$

Diese sind Isometrien von (V, q) die eine Hyperebene von V fest lassen und nicht die Identität sind. Nach Bemerkung 9.26 sind $s_1, \dots, s_r \in O(V)$ Spiegelungen. Es gilt $s = s_1 \circ \dots \circ s_r$, die Isometrie s ist damit ein Produkt von höchstens $n - 1$ Spiegelungen.

Fall 2: Angenommen es existiert kein $v \in V$ mit $q(v) \neq 0$ und $s(v) = v$. Falls $V = \{0\}$, so ist nichts zu zeigen.

Sei daher $V \neq \{0\}$ und $0 \neq v \in V$. Da (V, q) anisotrop ist, gilt $q(v) \neq 0$. Aufgrund der Annahme gilt auch $s(v) - v \neq 0$ und $q(s(v) - v) \neq 0$. Der quadratische Raum (V, q) enthält daher einen anisotropen Vektor v , so dass $s(v) - v$ anisotrop ist. Nach Lemma 9.29 gibt es eine Spiegelung s_0 , so dass $s_0 \circ s$ den Vektor v fest lässt. Aus Fall 1 folgt, dass $s_0 \circ s = s_1 \circ \dots \circ s_r$ für ein $r \leq n - 1$ und Spiegelungen $s_1, \dots, s_r$. Da $s_0^2 = \text{id}_V$ folgt $s = s_0 \circ \dots \circ s_r$, die Isometrie s ist damit ein Produkt aus höchstens n Spiegelungen. $\square$

Bemerkung 9.31. Sei $\dim(V) = 2$. Für $\text{id}_V \neq s \in O(V, q)$ sind äquivalent:

- i) $s \in \text{SO}(V, q)$, das heißt $\det(s) = 1$.
- ii) s ist das Produkt von 2 Spiegelungen.
- iii) 0 ist der einzige Fixvektor von s .

Beweis. Übung. $\square$

Hieraus kann man folgern, dass $\text{SO}(V, q)$ für $\dim(V) = 2$ abelsch ist. Für $\text{char}(K) \neq 2, 3$ ist $O(V, q)$ nicht abelsch.

9.4 Quadratische Formen über $\mathbb{R}$

Bisher haben wir quadratische Räume (V, q) über einem Körper K mit $\text{char}(K) \neq 2$ betrachtet. Nun betrachten wir speziell $K = \mathbb{R}$. Im Folgenden bezeichnet V einen $\mathbb{R}$ -Vektorraum mit $\dim(V) = n$ und q eine nicht ausgeartete quadratische Form auf V .

Bemerkung 9.32. Unter diesen Voraussetzungen gilt

$$q \simeq \underbrace{[1, \dots, 1]}_r, \underbrace{[-1, \dots, -1]}_{n-r}$$

mit einer ganzen Zahl $r \geq 0$. Für Matrizen formuliert heißt dies:

Jede symmetrische Matrix $A \in \text{GL}(n, \mathbb{R})$ ist kongruent zu einer Diagonalmatrix der Form

$$\text{diag}(\underbrace{1, \dots, 1}_r, \underbrace{-1, \dots, -1}_{n-r}) = \begin{pmatrix} E_r & 0 \\ 0 & -E_{n-r} \end{pmatrix}.$$

Beweis. Nach Satz 9.11 gibt es eine Orthogonalbasis $(b_1, \dots, b_n)$ von V , das heißt $\beta(b_i, b_j) = 0$ für $i \neq j$. Die quadratische Form q ist dann eindeutig durch die reellen Zahlen

$$a_i := q(b_i), \quad i = 1, \dots, n,$$

bestimmt. Da q nicht ausgeartet ist, gilt $a_i \neq 0$, also entweder $a_i > 0$ oder $a_i < 0$ für jedes $i \in \{1, \dots, n\}$. Nach Umnummerieren der Basis kann man ohne Beschränkung der Allgemeinheit annehmen, dass

$$a_1, \dots, a_r > 0 \quad \text{und} \quad a_{r+1}, \dots, a_n < 0.$$

Setze

$$c_i = \frac{1}{\sqrt{a_i}} b_i, \quad \text{für } i = 1, \dots, r,$$

$$c_i = \frac{1}{\sqrt{-a_i}} b_i, \quad \text{für } i = r+1, \dots, n.$$

Dann ist $(c_1, \dots, c_n)$ eine Orthogonalbasis von (V, q) und

$$q(c_i) = \begin{cases} \left(\frac{1}{\sqrt{a_i}}\right)^2 a_i = +1, & i = 1, \dots, r, \\ \left(\frac{1}{\sqrt{-a_i}}\right)^2 a_i = -1, & i = r+1, \dots, n. \end{cases}$$

Folglich ist

$$q \simeq [\underbrace{1, \dots, 1}_r, \underbrace{-1, \dots, -1}_{n-r}].$$

Für Matrizen: Ist A die zu q gehörige Matrix, so folgt aus $q \simeq [1, \dots, 1, -1, \dots, -1]$, dass A kongruent zu

$$2 \begin{pmatrix} E_r & 0 \\ 0 & -E_{n-r} \end{pmatrix}$$

ist. Da $\sqrt{2} \in \mathbb{R}$, ist A dann auch zu

$$\begin{pmatrix} E_r & 0 \\ 0 & -E_{n-r} \end{pmatrix} = \left(\frac{1}{\sqrt{2}} E_n\right)^t \begin{pmatrix} 2E_r & 0 \\ 0 & -2E_{n-r} \end{pmatrix} \left(\frac{1}{\sqrt{2}} E_n\right)$$

kongruent. □

Nachfolgend zeigen wir nun, dass r nur von (V, q) abhängt und eindeutig bestimmt ist.

Satz 9.33 (Sylvesterscher Trägheitssatz). *Unter den obigen Voraussetzungen ist die Zahl r aus Bemerkung 9.32 eindeutig bestimmt. Das heißt: Ist V ein n -dimensionaler $\mathbb{R}$ -Vektorraum und q eine nicht ausgeartete quadratische Form auf V , dann gibt es genau ein $r \geq 0$, so dass $q \simeq [\underbrace{1, \dots, 1}_r, \underbrace{-1, \dots, -1}_{n-r}]$.*

Für Matrizen: Gilt für ganze Zahlen $0 \leq r \leq n$, $0 \leq r' \leq n$, dass

$$\begin{pmatrix} E_{r'} & 0 \\ 0 & -E_{n-r'} \end{pmatrix} = S^t \begin{pmatrix} E_r & 0 \\ 0 & -E_{n-r} \end{pmatrix} S$$

mit einer reellen $n \times n$ -Matrix S , so gilt $r = r'$.

Beweis. Nach Bemerkung 9.32 gilt $q \simeq [\underbrace{1, \dots, 1}_r, \underbrace{-1, \dots, -1}_{n-r}]$ mit einer ganzen Zahl $r \geq 0$.

Es gelte nun auch $q \simeq [\underbrace{1, \dots, 1}_{r'}, \underbrace{-1, \dots, -1}_{n-r'}]$ mit einer ganzen Zahl $r' \geq 0$, also

$$[\underbrace{1, \dots, 1}_r, \underbrace{-1, \dots, -1}_{n-r}] \simeq [\underbrace{1, \dots, 1}_{r'}, \underbrace{-1, \dots, -1}_{n-r'}].$$

Ohne Beschränkung der Allgemeinheit sei $r' \geq r$. Annahme: $r' > r$.

Wir haben dann

$$[\underbrace{1, \dots, 1}_r, \underbrace{1, \dots, 1}_{r'-r}, \underbrace{-1, \dots, -1}_{n-r'}] \simeq [\underbrace{1, \dots, 1}_r, \underbrace{-1, \dots, -1}_{n-r}].$$

Aus dem Kürzungssatz von Witt (Satz 9.16) folgt dann

$$[\underbrace{1, \dots, 1}_{r'-r}, \underbrace{-1, \dots, -1}_{n-r'}] \simeq [\underbrace{-1, \dots, -1}_{n-r}].$$

Die Diagonalform auf der rechten Seite stellt nur reelle Zahlen kleiner gleich Null dar (das heißt $q(v) \leq 0$ für alle $v \in V$), die Diagonalform auf der linken Seite stellt hingegen mindestens eine positive Zahl dar. Die quadratischen Formen können daher nicht äquivalent sein. Dies ist ein Widerspruch, es muss daher $r' = r$ gelten.

Da die zu

$$[\underbrace{1, \dots, 1}_r, \underbrace{-1, \dots, -1}_{n-r}]$$

gehörige Matrix

$$2 \begin{pmatrix} E_r & 0 \\ 0 & -E_{n-r} \end{pmatrix}$$

ist, folgt hieraus auch die Aussage über Matrizen. $\square$

Wir haben also gesehen, dass für einen nicht ausgearteten reellen quadratischen Raum (V, q) die Zahlen r und $n - r$ eindeutig festgelegt sind.

Definition 9.34. Sei V wie oben und q eine nicht ausgeartete quadratische Form auf V . Gilt $q \simeq [\underbrace{1, \dots, 1}_r, \underbrace{-1, \dots, -1}_s]$ mit ganzen Zahlen $r, s \geq 0$, so nennt man das Paar (r, s) die *Signatur* von q .

Definition 9.35. Eine quadratische Form q auf einem beliebigen $\mathbb{R}$ -Vektorraum V heißt *positiv semi-definit*, wenn $q(x) \geq 0$ für alle $x \in V$. Gilt sogar $q(x) > 0$ für alle $x \in V \setminus \{0\}$, so heißt q *positiv definit*. Gilt hingegen $q(x) \leq 0$ für jedes $x \in V$, so nennt man q *negativ semi-definit*. *Negativ definit* nennt man q , wenn $q(x) < 0$ für alle $x \in V \setminus \{0\}$. Ist q weder positiv noch negativ semi-definit, so heißt q *indefinit*.

Bemerkung 9.36. Sei (V, q) ein quadratischer Raum über $\mathbb{R}$. Ist q positiv definit, so gilt $\det(B) > 0$ für die Strukturmatrix von q bezüglich einer beliebigen Basis $b_1, \dots, b_n$ von V . Ist q positiv semi-definit, so gilt $\det(B) \geq 0$.

Beweis. Ist $c_1, \dots, c_n$ eine weitere Basis von V und C die zugehörige Strukturmatrix, so gilt $\det(C) = \alpha^2 \det(B)$ mit $\alpha \in \mathbb{R}^\times$. Das Vorzeichen der Determinante ist daher unabhängig von der Wahl der Basis. Sei $b_1, \dots, b_n$ eine Orthogonalbasis. Dann gilt

$$\begin{aligned} \det(B) &= \beta(b_1, b_1) \cdot \dots \cdot \beta(b_n, b_n) \\ &= 2^n \cdot q(b_1) \cdot \dots \cdot q(b_n) \\ &\begin{cases} > 0, & \text{falls } q \text{ positiv definit ist,} \\ \geq 0, & \text{falls } q \text{ positiv semi-definit ist.} \end{cases} \end{aligned}$$

□

Definition 9.37. Eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ heißt *positiv definit*, wenn die quadratische Form q_A auf $\mathbb{R}^n$ positiv definit ist, das heißt, falls

$$x^t A x > 0$$

für alle $x \in \mathbb{R}^n \setminus \{0\}$. Entsprechendes gilt für die Bezeichnungen positiv semi-definit, negativ definit und negativ semi-definit.

Bemerkung 9.38. Für eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ gilt

$$\begin{aligned} A \text{ positiv definit} &\Rightarrow \det(A) > 0, \\ A \text{ positiv semi-definit} &\Rightarrow \det(A) \geq 0. \end{aligned}$$

Achtung. Aus $\det(A) > 0$ folgt im Allgemeinen nicht, dass A positiv definit ist. Beispiel:

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

Satz 9.39 (Determinantenkriterium). *Sei $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix. Dann gilt*

$$A \text{ ist positiv definit} \Leftrightarrow \det \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix} > 0 \text{ f\"ur alle } k = 1, \dots, n.$$

Beweis. Sei $q = q_A$ die quadratische Form auf $V = \mathbb{R}^n$ mit Strukturmatrix A bezüglich der Standardbasis $e_1, \dots, e_n$.

„ $\Rightarrow$ “: Folgt leicht aus Bemerkung 9.36 angewandt auf (W, q_W) mit $W = \text{Lin}(e_1, \dots, e_k)$ für $k = 1, \dots, n$.

„ $\Leftarrow$ “: Beweis per Induktion nach n . Der Fall $n = 1$ ist klar, sei $n > 1$. Sei $W \subset \mathbb{R}^n$ der von $e_1, \dots, e_{n-1} \in \mathbb{R}^n$ erzeugte Teilraum. Der quadratische Raum (W, q_W) hat Strukturmatrix

$$A' = \begin{pmatrix} a_{1,1} & \dots & a_{1,n-1} \\ \vdots & \ddots & \vdots \\ a_{n-1,1} & \dots & a_{n-1,n-1} \end{pmatrix}.$$

Diese erfüllt das Determinantenkriterium, woraus nach der Induktionsannahme folgt, dass A' positiv definit ist. Es ist daher auch q_W positiv definit und insbesondere (W, q_W) nicht ausgeartet. Es gilt dann

$$V = W \oplus \mathbb{R}c$$

mit $0 \neq c \in W^\perp$. Denn: da (W, q_W) nicht ausgeartet ist, gilt $W \cap W^\perp = \{0\}$. Damit ist die Dimension von $W + W^\perp$ gleich $\dim(W) + \dim(W^\perp)$. Außerdem gilt nach der Dimensionsformel für orthogonale Komplemente

$$\dim(W^\perp) = \dim(V) - \dim(W) = n - (n - 1) = 1,$$

da (V, q) ebenfalls nicht ausgeartet ist. Damit ist $V = W + W^\perp$ und die Summe ist direkt, da der Schnitt leer ist und jedes $0 \neq c \in W^\perp$ spannt $W^\perp$ auf.

Um die positive Definitheit von q zu zeigen, betrachten wir die Basis $(e_1, \dots, e_{n-1}, c)$ von V . Die Gram-Matrix von q bezüglich dieser Basis ist

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

mit $a = \beta_q(c, c) = 2q(c)$. Da sich $\det(C)$ von $\det(A)$ nur um das Quadrat einer von Null verschiedenen reellen Zahl unterscheidet, und da nach Voraussetzung $\det(A) > 0$ ist, muss auch $\det(C) > 0$ sein. Andererseits gilt $\det(C) = \det(A') \cdot a$ und $\det(A') > 0$. Folglich ist $a > 0$ und somit $q_{W^\perp}$ positiv definit. Es folgt die positive Definitheit von q , und damit auch die positive Definitheit der Matrix A . $\square$

Achtung. Für positiv semi-definit gilt das Kriterium nicht analog: $A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$ erfüllt $a_{11} \geq 0$ und $\det(A) \geq 0$, aber A ist nicht positiv semi-definit.

Satz 9.40. Sei $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix. A ist genau dann positiv semi-definit, wenn alle Hauptunterdeterminanten von A größer gleich 0 sind. Das heißt falls

$$\det(A_S) := \det \begin{pmatrix} a_{i_1 i_1} & \cdots & a_{i_1, i_k} \\ \vdots & \ddots & \vdots \\ a_{i_k i_1} & \cdots & a_{i_k, i_k} \end{pmatrix} \geq 0$$

für alle $k = 1, \dots, n$ und alle k -elementigen Teilmengen $S = \{i_1, \dots, i_k\} \subset \{1, \dots, n\}$.

Beweis. Übung. Ähnlich wie der Beweis des Determinantenkriteriums (Satz 9.39). $\square$

Satz 9.41 (Variante des Sylvesterschen Trägheitssatzes). Sei (V, q) ein nicht ausgearteter quadratischer Raum über $\mathbb{R}$. Dann lässt sich V als orthogonale Summe

$$V = U_1 \oplus U_2$$

schreiben, wobei q_{U_1} positiv definit und q_{U_2} negativ definit ist. Die Größen

$$\begin{aligned} r &= \dim(U_1), \\ s &= \dim(U_2) \end{aligned}$$

sind durch (V, q) eindeutig bestimmt.

Beweis. Dies folgt aus dem Sylvesterschen Trägheitssatz (Satz 9.33). $\square$

Satz 9.42 (Determinantenbeschreibung der Signatur einer quadratischen Form). Sei (V, q) ein nicht ausgearteter quadratischer Raum über $\mathbb{R}$ mit Signatur (r, s) . Sei $A = (a_{ij})$ die Strukturmatrix von q bezüglich einer Basis von V , so dass die Bedingung

$$d_k := \det \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} \neq 0$$

für alle $k = 1, \dots, n$ erfüllt ist. Dann ist s gleich der Anzahl der Vorzeichenwechsel der Folge $(1, d_1, d_2, \dots, d_n)$.

Beweis. Dies beweist man wie das Determinantenkriterium (Satz 9.39) mit passenden Änderungen. Induktion nach n :

$n = 1$: Ist $d_1 > 0$, so hat die Folge $(1, d_1)$ keinen Vorzeichenwechsel. Es gilt $A = (a_{11}) = (d_1) \simeq (1) \in \mathbb{R}^{1 \times 1}$, das heißt $s = 0$. Ist $d_1 < 0$, so hat die Folge $(1, d_1)$ einen Vorzeichenwechsel. Es gilt $A = (a_{11}) = (d_1) \simeq (-1) \in \mathbb{R}^{1 \times 1}$, das heißt $s = 1$. In beiden Fällen ist s gleich der Anzahl der Vorzeichenwechsel.

$n > 1$: Sei $B = (b_1, b_2, \dots, b_n)$ eine Basis von V mit Strukturmatrix A . Definiere den Unterraum $W = \text{Lin}(b_1, \dots, b_{n-1}) \subset V$. Der quadratische Raum (W, q_W) hat Strukturmatrix

$$A' = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n-1} \\ \vdots & \ddots & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{pmatrix}$$

bezüglich der Basis $(b_1, \dots, b_{n-1})$. Nach der Induktionsannahme hat (W, q_W) die Signatur (r', s') , wobei s' gleich der Anzahl der Vorzeichenwechsel der Folge $(1, d_1, \dots, d_{n-1})$ ist. Da $\det(A') = d_{n-1} \neq 0$, ist (W, q_W) nicht ausgeartet. Es gilt dann wieder

$$V = W \oplus \text{Lin}(c)$$

mit eindimensionalem $W^\perp = \text{Lin}(c)$. Sei $c_1, \dots, c_{n-1}$ eine Orthogonalbasis von W mit Strukturmatrix

$$C' = 2 \begin{pmatrix} E_{r'} & 0 \\ 0 & -E_{s'} \end{pmatrix}.$$

Dann ist $c_1, \dots, c_{n-1}, c$ eine Orthogonalbasis von V mit Strukturmatrix

$$C = 2 \begin{pmatrix} E_{r'} & 0 & 0 \\ 0 & -E_{s'} & 0 \\ 0 & 0 & q(c) \end{pmatrix}.$$

Nach Lemma 7.26 gilt

$$\begin{aligned} q(c) &= 2^{-n} \cdot (-1)^{s'} \cdot \det(C) \\ &= a^2 \cdot 2^{-n} \cdot (-1)^{s'} \cdot \det(A) \\ &= a^2 \cdot 2^{-n} \cdot (-1)^{s'} \cdot d_n, \\ (-1)^{s'} &= 2^{1-n} \cdot \det(C') \\ &= a'^2 \cdot \det(A') = a'^2 \cdot d_{n-1} \end{aligned}$$

mit $a, a' \in \mathbb{R}^\times$. Schreiben wir „ $\simeq$ “ für die Kongruenz von Matrizen, so folgt für die Signatur (r, s) von (V, q)

$$\begin{aligned} (r, s) = (r', s' + 1) &\Leftrightarrow C \simeq \begin{pmatrix} E_{r'} & 0 \\ 0 & -E_{s'+1} \end{pmatrix} \\ &\Leftrightarrow q(c) < 0 \\ &\Leftrightarrow d_n \text{ und } d_{n-1} \text{ haben unterschiedliche Vorzeichen} \\ &\Leftrightarrow \text{Die Folge } (1, d_1, \dots, d_n) \text{ hat } s' + 1 \text{ Vorzeichenwechsel.} \end{aligned}$$

Analog gilt

$$\begin{aligned} (r, s) = (r' + 1, s') &\Leftrightarrow C \simeq \begin{pmatrix} E_{r'+1} & 0 \\ 0 & -E_{s'} \end{pmatrix} \\ &\Leftrightarrow q(c) > 0 \\ &\Leftrightarrow d_n \text{ und } d_{n-1} \text{ haben das gleiche Vorzeichen} \\ &\Leftrightarrow \text{Die Folge } (1, d_1, \dots, d_n) \text{ hat } s' \text{ Vorzeichenwechsel.} \end{aligned}$$

In beiden Fällen ist s gleich der Anzahl der Vorzeichenwechsel der Folge $(1, d_1, \dots, d_n)$. $\square$

Satz 9.43 (Cauchy-Schwarz'sche Ungleichung). *Sei (V, q) ein quadratischer Raum über $\mathbb{R}$ und q positiv semi-definit. Für alle $x, y \in V$ gilt*

$$\beta(x, y)^2 \leq \beta(x, x)\beta(y, y).$$

Beweis. Ohne Beschränkung der Allgemeinheit sei $V = \text{Lin}(x, y)$. Falls $\dim(V) = 0$, so folgt $x = y = 0$ und die Ungleichung ist erfüllt. Falls $\dim(V) = 1$ so sind x und y linear abhängig und es herrscht Gleichheit aufgrund der Bilinearität von β . Falls $\dim(V) = 2$, so ist x, y eine Basis von V . Bezüglich dieser hat β die Gram-Matrix

$$B = \begin{pmatrix} \beta(x, x) & \beta(x, y) \\ \beta(x, y) & \beta(y, y) \end{pmatrix}.$$

Nach Bemerkung 9.36 folgt $\beta(x, x)\beta(y, y) - \beta(x, y)^2 = \det(B) \geq 0$. $\square$

Bemerkung 9.44. Sei (V, q) ein quadratischer Raum über $\mathbb{R}$ und q positiv definit. Dann gilt in der Cauchy-Schwarz'schen Ungleichung genau dann Gleichheit, wenn die Vektoren x, y linear abhängig sind.

Beweis. „ $\Leftarrow$ “: Dies steht bereits im obigen Beweis.

„ $\Rightarrow$ “: Aus $\beta(x, y)^2 = \beta(x, x)\beta(y, y)$ folgt $\det(B) = 0$. Da q nicht ausgeartet ist, ist dies nur für linear abhängige Vektoren x, y möglich. $\square$

9.4.1 Matrix-Zerlegungen

Bemerkung 9.45. Ist q eine positiv definite quadratische Form auf einem $\mathbb{R}$ -Vektorraum V , so ist die zugehörige Bilinearform β ein *Skalarprodukt* auf V . In diesem Fall kann man zu (V, q) eine *Orthonormalbasis* finden. Zum Beispiel mit dem Gram-Schmidtschen Orthonormalisierungsverfahren. Hierzu ersetzt man in Algorithmus 8.11 $\|x\|$ durch $\sqrt{q(x)}$ und $\langle x, y \rangle$ durch $\frac{1}{2}\beta(x, y)$ für $x, y \in V$.

Für positiv definite symmetrische Matrizen bedeutet dies:

Satz 9.46. *Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische positiv definite Matrix. Dann existiert eine obere Dreiecksmatrix $C \in \mathbb{R}^{n \times n}$, so dass*

$$C^t A C = E_n.$$

Korollar 9.47 (Cholesky-Zerlegung). *Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische positiv definite Matrix. Dann lässt sich A in der Form*

$$A = B^t B$$

für eine obere Dreiecksmatrix B schreiben.

Lemma 9.48. *Sei $x \in \mathbb{R}^n$ mit $\|x\|^2 = x^t x = 1$. Dann existiert ein $S \in O(n, \mathbb{R})$ mit $Se_1 = x$, wobei e_1 den ersten Standardbasisvektor bezeichnet.*

Beweis. Ergänze x zu einer Orthonormalbasis $(x, b_2, \dots, b_n)$ des $\mathbb{R}^n$. Fassen wir diese Basisvektoren als Spalten einer Matrix $S \in \mathbb{R}^{n \times n}$ auf, so gilt $S \in O(n, \mathbb{R})$. Außerdem ist $Se_1 = x$. $\square$

Satz 9.49 (QR-Zerlegung). *Sei $A \in GL_n(\mathbb{R})$. Es existieren ein $Q \in O(n, \mathbb{R})$ und eine obere Dreiecksmatrix $R = (r_{ij}) \in \mathbb{R}^{n \times n}$ mit $r_{ii} > 0$ für $i = 1, \dots, n$, so dass*

$$A = QR.$$

Hierbei sind Q und R eindeutig bestimmt.

Beweis. Wir zeigen die Existenz durch Induktion nach n .

Für $n = 1$ ist die Behauptung klar.

Sei $n > 1$. Sei a die erste Spalte der Matrix A . Da A invertierbar ist, gilt $A \neq 0$. Der Vektor $a' = a/\|a\| \in \mathbb{R}^n$ hat Norm 1. Nach Lemma 9.48 existiert ein $S \in O(n, \mathbb{R})$ mit $Sa' = e_1$. Somit gilt $Sa = \|a\| \cdot e_1$ und

$$SA = \begin{pmatrix} \|a\| & * \\ 0 & A_1 \end{pmatrix}$$

mit einer Matrix $A_1 \in GL_{n-1}(\mathbb{R})$. Nach Induktionsannahme existieren ein $S_1 \in O(n-1, \mathbb{R})$ und eine obere Dreiecksmatrix $R_1 \in \mathbb{R}^{(n-1) \times (n-1)}$ mit $A_1 = S_1 R_1$. Es folgt, dass

$$A = S^{-1} \begin{pmatrix} 1 & 0 \\ 0 & S_1 \end{pmatrix} \begin{pmatrix} \|a\| & * \\ 0 & R_1 \end{pmatrix}$$

und damit die Existenz von Q und R .

Die Eindeutigkeit der Zerlegung folgt aus der Tatsache, dass eine obere Dreiecksmatrix mit positiven Diagonaleinträgen genau dann in $O(n, \mathbb{R})$ liegt, wenn sie die Einheitsmatrix ist. $\square$

Satz 9.50 (Iwasawa-Zerlegung). *Sei $A \in GL_n(\mathbb{R})$. Es existieren ein $Q \in O(n, \mathbb{R})$, eine Diagonalmatrix D mit positiven Diagonaleinträgen und eine unipotente obere Dreiecksmatrix*

$$N = \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}$$

so dass

$$A = QDN.$$

Hierbei sind Q , D und N eindeutig bestimmt.

Beweis. Sei $A = QR$ die Zerlegung nach Satz 9.49 mit $R = (r_{ij})$. Setze

$$D = \begin{pmatrix} r_{11} & & 0 \\ & \ddots & \\ 0 & & r_{nn} \end{pmatrix}$$

und $N = D^{-1}R$. Dann haben Q , D und R die gewünschten Eigenschaften.

Die Eindeutigkeit der Zerlegung folgt leicht aus der Eindeutigkeit in Satz 9.49. $\square$

Eine weitere Zerlegung ergibt sich leicht aus dem Gauß-Algorithmus.

Satz 9.51 (LR-Zerlegung). *Sei $A \in \mathbb{R}^{n \times n}$. Es existieren eine Permutationsmatrix P , eine untere Dreiecksmatrix L und eine obere Dreiecksmatrix R , so dass*

$$A = PLR.$$

Lemma 9.52 (Quadratwurzel-Lemma). *Sei $A \in \mathbb{R}^{n \times n}$ eine positiv semidefinite symmetrische Matrix. Es existiert eine eindeutig bestimmte positiv semidefinite symmetrische Matrix P mit*

$$A = P^2.$$

Falls A positiv definit ist, so ist auch P positiv definit.

Beweis. Nach dem Spektralsatz für symmetrische Matrizen kann A durch eine orthogonale Matrix diagonalisiert werden. Es gibt also ein $S \in O(n, \mathbb{R})$ und eine Diagonalmatrix D , so dass

$$A = S^t D S.$$

Da A positiv semidefinit ist, sind die Diagonaleinträge von D nicht-negativ. Sei $\sqrt{D}$ die eindeutig bestimmte Diagonalmatrix mit nicht-negativen Diagonaleinträgen, so dass $(\sqrt{D})^2 = D$. Dann hat

$$P = S^t \sqrt{D} S$$

die gewünschten Eigenschaften. Den Nachweis der Eindeutigkeit von P lassen wir als Übung. $\square$

Bemerkung 9.53. Sei $T \in GL(n, \mathbb{R})$. Dann ist $T^t T$ symmetrisch und positiv definit.

Beweis. Symmetrisch: klar.

Positiv definit: Sei $x \in \mathbb{R}^n$. Dann gilt

$$x^t T^t T x = (Tx)^t (Tx) = \|Tx\|_2^2 \geq 0.$$

Gleichheit gilt genau dann, wenn $Tx = 0$. Weiter gilt $Tx = 0$ genau dann, wenn $x = 0$. $\square$

Satz 9.54 (Polar-Zerlegung). *Zu jeder Matrix $A \in GL_n(\mathbb{R})$ existiert ein $S \in O(n, \mathbb{R})$ und eine positiv definite symmetrische Matrix P , so dass*

$$A = SP.$$

Hierbei sind S und P eindeutig bestimmt.

Beweis. Existenz: Die Matrix $B = A^t A$ ist symmetrisch und positiv definit. Nach Lemma 9.52 existiert eine eindeutig bestimmte positiv definite Matrix P mit $B = P^2$. Setze $S = AP^{-1}$. Dann gilt $S^t S = 1$ und $A = SP$.

Eindeutigkeit: Aus $A = SP$ folgt $A^t A = P^t S^t S P = P^2$. Damit ist P die eindeutig bestimmte Quadratwurzel von $A^t A$ und S ist ebenfalls eindeutig bestimmt. $\square$

Satz 9.55 (Cartan-Zerlegung). *Zu jeder Matrix $A \in \text{GL}_n(\mathbb{R})$ existieren $S_1, S_2 \in \text{O}(n, \mathbb{R})$ und eine Diagonalmatrix D mit positiven Diagonaleinträgen, so dass*

$$A = S_1 D S_2.$$

Beweis. Sei $A = SP$ die Polar-Zerlegung von A . Nach dem Spektralsatz für symmetrische Matrizen gibt es ein $T \in \text{O}(n, \mathbb{R})$ und eine Diagonalmatrix D , so dass

$$P = T^t D T.$$

Da P positiv definit ist, sind die Diagonaleinträge von D positiv. Durch Einsetzen erhält man

$$A = S T^t D T,$$

also die Behauptung mit $S_1 = S T^t$ und $S_2 = T$. □

Literatur

- [1] *S. Bosch*, Lineare Algebra, Springer-Verlag, 2008.
- [2] *G. Fischer*, Lineare Algebra, Vieweg-Verlag, 2009.
- [3] *M. Koecher*, Lineare Algebra und Analytische Geometrie, Springer-Verlag, 2003.
- [4] *F. Lorenz*, Lineare Algebra 1, Bibliographisches Institut, Mannheim, 1992.
- [5] *E. Freitag*, Lineare Algebra, Vorlesungsskript, Universität Heidelberg.
- [6] *M. Otto*, Linear Algebra, Vorlesungsskript, TU Darmstadt.