

LINEARE ALGEBRA I

OLIVER C. SCHNÜRER

ZUSAMMENFASSUNG. Bei diesem Manuskript handelt es sich um Notizen zu einer Vorlesung Lineare Algebra I (B1) an der Universität Konstanz im Wintersemester 2010/11.

INHALTSVERZEICHNIS

0. Vorbemerkungen	1
1. Lineare Gleichungssysteme	2
2. Körper und Vektorräume	9
3. Struktur von Vektorräumen	17
4. Lineare Abbildungen	31
5. Determinanten	44
6. Vektorräume mit Skalarprodukt	59
Anhang A. Mengenlehre und Logik	77
Literatur	79

0. VORBEMERKUNGEN

0.1. Lineare Algebra in der Mathematik. Lineare Algebra I ist eine Grundvorlesung, deren sämtliche Inhalte später in allen mathematischen Gebieten ständig auftauchen werden und daher ein guter Einstieg in die Mathematik.

Die lineare Algebra beschäftigt sich mit Objekten, die linear, d. h. gerade aussehen, also wie eine Gerade oder eine Ebene. Die Realität ist aber in der Regel nicht linear. Jedoch sehen Dinge aus der Nähe häufig linear aus, z. B. ein kleines Stück der Tischplatte eines Tisches.

0.2. Was kann man mit Mathematik anfangen?

Nach <http://www.maths.monash.edu.au/>

- Understanding our world (Bewegung der Sterne, Sonnenprotuberanzen, schwarze Löcher, Wasserwellen, Windhose, Buschbrände)
- Modelling and improving systems (Verkehrsleitsysteme, Logistik für Containerschiffe, Börse, Produktion, Medizin)
- Studying the beauty of perfection (Seifenblasen, Symmetrien in Sonnenblumen oder geometrischen Mustern, Fraktale, Wassertropfen)

Date: 7. September 2018.

1991 *Mathematics Subject Classification.* 15-01.

Vielen Dank an Dieter Hoffmann, Matthias Makowski, Reinhard Racke, Olaf Schnürer und an die Hörer der Vorlesung für Korrekturen.

0.3. **Literatur.** Die meisten Bücher über lineare Algebra sind geeignet, z. B. die Bücher von U. Stambach [5] (häufig die Grundlage dieses Skriptes), G. Fischer [1] (auch häufiger benutzt), S. Lang [3] oder F. Lorenz [4], alle etwa mit dem Titel „Lineare Algebra“ und (mit Vorsicht; auch im Skript verwandt) Wikipedia (<http://www.wikipedia.org/>) für Definitionen.

1. LINEARE GLEICHUNGSSYSTEME

1.1. Lineare Gleichungssysteme.

Beispiel 1.1.1. Wir wollen das folgende Gleichungssystem lösen:

$$\begin{aligned} 2x + 6y - 4z &= 10, \\ -x + 5y + 2z &= 11, \\ 3x + 7y + z &= -3. \end{aligned}$$

Dazu multiplizieren wir die erste Zeile mit $\frac{1}{2}$.

$$\begin{aligned} x + 3y - 2z &= 5, \\ -x + 5y + 2z &= 11, \\ 3x + 7y + z &= -3. \end{aligned}$$

Wir addieren die erste Zeile zur zweiten Zeile und subtrahieren sie, mit 3 multipliziert, von der dritten Zeile

$$\begin{aligned} x + 3y - 2z &= 5, \\ &8y &= 16, \\ &- 2y + 7z &= -18. \end{aligned}$$

Multipliziere die zweite Zeile mit $\frac{1}{8}$ und addiere das Resultat, mit 2 multipliziert, zur dritten Zeile

$$\begin{aligned} x + 3y - 2z &= 5, \\ &y &= 2, \\ &7z &= -14. \end{aligned}$$

Multipliziere die dritte Zeile mit $\frac{1}{7}$ und addiere das Resultat, mit 2 multipliziert, zur ersten Zeile. Addiere die zweite Zeile, mit -3 multipliziert, zur ersten Zeile und erhalte

$$\begin{aligned} x &= -5, \\ y &= 2, \\ z &= -2. \end{aligned}$$

Wir erhalten die Lösung $(x, y, z) = (-5, 2, -2)$. D. h. für $x = -5$, $y = 2$ und $z = -2$ sind sämtliche Gleichungen des obigen Gleichungssystems erfüllt.

Wir wollen jedoch nicht in erster Linie spezielle Probleme lösen, sondern Aussagen über allgemeine Probleme machen.

Definition 1.1.2 (Lineares Gleichungssystem). Ein reelles lineares Gleichungssystem in n Unbekannten x^1, \dots, x^n und m Gleichungen ist ein System von Gleichungen der Form

$$(1.1) \quad \begin{aligned} a_1^1 x^1 + a_2^1 x^2 + \dots + a_n^1 x^n &= b^1 \\ a_1^2 x^1 + a_2^2 x^2 + \dots + a_n^2 x^n &= b^2 \\ \vdots &\vdots \\ a_1^m x^1 + a_2^m x^2 + \dots + a_n^m x^n &= b^m \end{aligned}$$

Dabei sind a_j^i und b^i , $1 \leq i \leq m$, $1 \leq j \leq n$, reelle Zahlen, also $a_j^i, b^i \in \mathbb{R}$.

(Anders als in vielen Büchern über lineare Algebra schreiben wir einen Index der Koeffizienten a_j^i nach oben. (a_j^i hier entspricht a_{ij} in solchen Büchern.) Dies

bedeutet keine Potenzierung und ist eine in der Differentialgeometrie und Physik (Einsteinsche Summenkonvention) übliche Schreibweise.)

Das System heißt homogen, falls $b^1 = \dots = b^m = 0$ gilt und sonst inhomogen.

Ein n -Tupel (ξ^1, \dots, ξ^n) reeller Zahlen, also $(\xi^1, \dots, \xi^n) \in \mathbb{R}^n$, heißt Lösung des linearen Gleichungssystems, wenn sämtliche Gleichungen des Gleichungssystems erfüllt sind, wenn wir die Zahlen ξ^i statt x^i für $1 \leq i \leq n$ einsetzen.

Theorem 1.1.3. *Sei (H) ein homogenes lineares Gleichungssystem wie in (1.1). Seien (ξ^1, \dots, ξ^n) und (η^1, \dots, η^n) Lösungen von (H). Sei $\lambda \in \mathbb{R}$. Dann sind $(\xi^1 + \eta^1, \dots, \xi^n + \eta^n)$ und $(\lambda\xi^1, \dots, \lambda\xi^n)$ ebenfalls Lösungen von (H).*

Beweis. Nach Voraussetzung gilt für alle $1 \leq i \leq m$

$$(1.2) \quad a_1^i \xi^1 + a_2^i \xi^2 + \dots + a_n^i \xi^n = 0,$$

$$(1.3) \quad a_1^i \eta^1 + a_2^i \eta^2 + \dots + a_n^i \eta^n = 0.$$

Wir addieren (1.2) und (1.3) und erhalten

$$a_1^i (\xi^1 + \eta^1) + a_2^i (\xi^2 + \eta^2) + \dots + a_n^i (\xi^n + \eta^n) = 0.$$

Somit ist $(\xi^1 + \eta^1, \dots, \xi^n + \eta^n)$ ebenfalls eine Lösung von (H).

Multipliziere (1.2) mit λ und erhalte

$$a_1^i (\lambda\xi^1) + a_2^i (\lambda\xi^2) + \dots + a_n^i (\lambda\xi^n) = 0.$$

Daher ist $(\lambda\xi^1, \dots, \lambda\xi^n)$ wie behauptet auch eine Lösung von (H). \square

Bemerkung 1.1.4. Sei (H) ein homogenes Gleichungssystem. Dann besitzt (H) stets die triviale Lösung $(0, \dots, 0)$.

Sei (L) ein lineares Gleichungssystem wie in (1.1). Dann bezeichnen wir das homogene lineare Gleichungssystem, das aus (1.1) entsteht, wenn wir die Koeffizienten b^i durch 0 ersetzen als das zugehörige homogene System (HL).

Theorem 1.1.5. *Sei (L) ein lineares Gleichungssystem wie in (1.1) und sei (HL) das zugehörige homogene lineare Gleichungssystem.*

- (i) *Sei $(\zeta^1, \dots, \zeta^n)$ eine Lösung von (L) und (ξ^1, \dots, ξ^n) eine Lösung von (HL), so ist $(\zeta^1 + \xi^1, \dots, \zeta^n + \xi^n)$ eine Lösung von (L).*
- (ii) *Seien $(\zeta^1, \dots, \zeta^n)$ und (η^1, \dots, η^n) Lösungen von (L), so ist $(\zeta^1 - \eta^1, \dots, \zeta^n - \eta^n)$ eine Lösung von (HL).*

Beweis.

- (i) Addiere

$$a_1^i \zeta^1 + a_2^i \zeta^2 + \dots + a_n^i \zeta^n = b^i$$

und

$$a_1^i \xi^1 + a_2^i \xi^2 + \dots + a_n^i \xi^n = 0$$

und erhalte

$$a_1^i (\zeta^1 + \xi^1) + a_2^i (\zeta^2 + \xi^2) + \dots + a_n^i (\zeta^n + \xi^n) = b^i.$$

Die erste Behauptung folgt.

- (ii) Subtrahiere

$$a_1^i \eta^1 + a_2^i \eta^2 + \dots + a_n^i \eta^n = b^i$$

von

$$a_1^i \zeta^1 + a_2^i \zeta^2 + \dots + a_n^i \zeta^n = b^i$$

und erhalte

$$a_1^i(\zeta^1 - \eta^1) + a_2^i(\zeta^2 - \eta^2) + \dots + a_n^i(\zeta^n - \eta^n) = 0.$$

Wir erhalten die zweite Behauptung. \square

Bemerkung 1.1.6. Für n -Tupel verwenden wir die Abkürzungen

$$\begin{aligned}\eta &= (\eta^1, \eta^2, \dots, \eta^n), \\ \zeta &= (\zeta^1, \zeta^2, \dots, \zeta^n), \dots,\end{aligned}$$

und definieren

$$\eta + \zeta = (\eta^1 + \zeta^1, \eta^2 + \zeta^2, \dots, \eta^n + \zeta^n)$$

sowie für $\lambda \in \mathbb{R}$

$$\lambda\eta = (\lambda\eta^1, \lambda\eta^2, \dots, \lambda\eta^n).$$

Lemma 1.1.7. Sei (L) ein lineares Gleichungssystem. Dann ist die Lösungsmenge unter den folgenden Operationen invariant.

- (i) Vertauschen von Gleichungen.
- (ii) Multiplikation einer Gleichung mit $\lambda \in \mathbb{R}$, $\lambda \neq 0$.
- (iii) Wir addieren das λ -fache der j -ten Gleichung zur i -ten Gleichung.

Beweis. Sei L_1 die Lösungsmenge von (L) und sei L_2 die Lösungsmenge des Systems, das wir erhalten, wenn wir die jeweilige Operation anwenden.

- (i) Klar.
- (ii) Für jede Zahl $\lambda \in \mathbb{R}$ gilt $L_1 \subset L_2$. Ist $\lambda \neq 0$, so erhalten wir durch Multiplikation derselben Gleichung mit $\frac{1}{\lambda}$ gerade wieder (L) . Somit gilt auch $L_2 \subset L_1$, insgesamt also $L_1 = L_2$.
- (iii) Auch hier ist $L_1 \subset L_2$ wieder klar. Addition des $(-\lambda)$ -fachen der j -ten Gleichung zur i -ten Gleichung liefert wieder (L) , also gilt auch $L_2 \subset L_1$ und daher ist auch hier wieder $L_1 = L_2$. \square

1.2. Gaußsches Eliminationsverfahren. Das Gaußsche Eliminationsverfahren erlaubt es, allgemeine lineare Gleichungssysteme zu lösen oder nachzuweisen, dass sie keine Lösung besitzen.

Definition 1.2.1 (Zeilenstufenform). Sei (L) ein lineares Gleichungssystem in der Form (1.1). Dann ist (L) in Zeilenstufenform, falls „in jeder Zeile links mehr Nullen als in der Vorgängerzeile stehen“, d. h. setzen wir

$$A_j^i := \begin{cases} a_j^i, & 1 \leq j \leq n, 1 \leq i \leq m, \\ b_j^i, & j = n+1, 1 \leq i \leq m, \end{cases}$$

so ist die Anzahl der „links stehenden Nullen“ in Zeile i durch

$$N(i) := \max \{k \in \{0, 1, \dots, n+1\} : A_j^i = 0 \text{ für } 1 \leq j \leq k\}$$

definiert und wir fordern, dass es ein $i_0 \in \{0, \dots, m-1\}$ gibt, so dass $N(i) < N(i+1)$ für alle $1 \leq i \leq i_0$ und $N(i) = n+1$ für $i_0 < i \leq m$ gelten.

Wir verwenden

$$\mathbb{N} = \{0, 1, 2, \dots\} \quad \text{und} \quad \mathbb{N}^+ \equiv \mathbb{N}_{>0} \equiv \mathbb{N} \setminus \{0\}.$$

Theorem 1.2.2 (Gaußsches Eliminationsverfahren).

Sei (L) ein lineares Gleichungssystem wie in (1.1). Dann lässt sich (L) mit endlich vielen Operationen aus Lemma 1.1.7 in ein lineares Gleichungssystem (G) der Form (1.1) umformen, so dass (G) in Zeilenstufenform ist. (L) und (G) haben dieselbe Lösungsmenge.

Beweis. Die Gleichheit der Lösungsmengen folgt direkt nach Lemma 1.1.7.

Wir zeigen nun per Induktion nach der Anzahl der Spalten bzw. der Variablen, dass dies stets möglich ist.

Wir betrachten zuerst den Fall einer Variablen. Dann haben wir ein Gleichungssystem der Form

$$\begin{aligned} a_1^1 x^1 &= b^1, \\ &\vdots \\ a_1^m x^1 &= b^m. \end{aligned}$$

Ist ein $a_1^j \neq 0$, so dürfen wir nach Vertauschung der Zeilen annehmen, dass $a_1^1 \neq 0$ gilt. Addiere nun $-\frac{a_1^2}{a_1^1}$ -mal die erste Zeile zur zweiten, $-\frac{a_1^i}{a_1^1}$ -mal die erste Zeile zur i -ten. Damit haben wir das obige System auf eine Form mit $a_1^i = 0$ für $2 \leq i \leq m$ gebracht. Wir behalten die Bezeichnungen bei. Ist ein $b^i \neq 0$, so dürfen wir nach Zeilenvertauschung ohne Einschränkung annehmen, dass dies b^2 ist. Falls für alle $2 \leq i \leq m$ auch $b^i = 0$ gilt, so sind wir fertig. Nehme daher an, dass dieser Fall nicht eintritt. Wir addieren nun $-\frac{b^3}{b^2}$ -mal die zweite Zeile zur dritten und $-\frac{b^i}{b^2}$ -mal die zweite Zeile zur i -ten. Das Ergebnis ist ein Gleichungssystem in Zeilenstufenform.

Sei nun die Anzahl der Variablen n in (1.1) größer als eins. Nach Umordnen der Gleichungen dürfen wir ohne Einschränkung annehmen, dass $a_1^1 \neq 0$ gilt. (Gibt es nämlich keinen Koeffizienten a_1^i in der ersten Spalte, der nicht verschwindet, so folgt die Behauptung per Induktion.) Addiere nun $-\frac{a_1^i}{a_1^1}$ -mal die erste Zeile zur i -ten. Wir bezeichnen die Koeffizienten im neuen linearen Gleichungssystem wieder mit a_1^i . Dann gilt $a_1^i = 0$ für $2 \leq i \leq m$. Nach Induktionsvoraussetzung können wir den Block ab der zweiten Gleichung in einem System der Form

$$\begin{aligned} a_1^1 x^1 + a_2^1 x^2 + \dots + a_n^1 x^n &= b^1 \\ \tilde{a}_2^2 x^2 + \dots + \tilde{a}_n^2 x^n &= \tilde{b}^2 \\ &\vdots \\ \tilde{a}_2^m x^2 + \dots + \tilde{a}_n^m x^n &= \tilde{b}^m. \end{aligned}$$

in Zeilenstufenform bringen. Die Behauptung folgt daher.

Nach Multiplikation der i -ten Zeile mit $\frac{1}{A_{N(i)+1}^i}$ wie in Definition 1.2.1 dürfen wir zusätzlich ohne Einschränkung annehmen, dass das Gleichungssystem in Zeilenstufenform $A_{N(i)+1}^i = 1$ erfüllt (falls $N(i) \leq n$ ist), dass also der erste von Null verschiedene Koeffizient gleich 1 ist, falls die Gleichung nicht $0 = 0$ lautet. \square

Lemma 1.2.3. *Sei (L) ein lineares Gleichungssystem der Form (1.1) in Zeilenstufenform.*

- (i) *Dann besitzt (L) genau dann eine Lösung, wenn (L) keine Zeile der Form $0 = b^i$ mit $b^i \neq 0$ enthält.*
- (ii) *Die Lösung ist eindeutig, falls eine Lösung existiert, $N(i) = i-1$ für $1 \leq i \leq m$ mit $N(i)$ wie in Definition 1.2.1 und $n = m$ gelten.*

Beweis.

- (i) Damit (L) eine Lösung besitzen, darf keine Zeile der Form $0 \neq 0$ auftreten. Nehme dies an. Dann hat die letzte Zeile, die nicht die Form $0 = 0$ hat, die Form

$$a_k^i x^k + \dots + a_n^i x^n = b^i$$

mit $a_k^i \neq 0$. Wähle ξ^{k+1}, \dots, ξ^n beliebig und setze

$$\xi^k := \frac{1}{a_k^i} (b^i - (a_{k+1}^i \xi^{k+1} + \dots + a_n^i \xi^n)).$$

(ξ^k, \dots, ξ^n) löst diese letzte Gleichung. Ersetze nun x^k, \dots, x^n im restlichen Gleichungssystem durch ξ^k, \dots, ξ^n . Dieses Gleichungssystem hat nun eine Gleichung weniger als das bisher betrachtete. Per Induktion finden wir daher eine Lösung.

- (ii) Gilt $N(1) = 0, N(2) = 1, \dots$, so sind sämtliche ξ^i beim Lösen des Gleichungssystems von unten eindeutig bestimmt. Daher ist die Lösung eindeutig. \square

Korollar 1.2.4. Sei (H) ein homogenes lineares Gleichungssystem der Form (1.1) mit $m < n$, also mit weniger Zeilen als Variablen. Dann besitzt (H) eine nichttriviale Lösung, d. h. eine Lösung $(\xi^1, \dots, \xi^n) \neq (0, \dots, 0)$.

Beweis. Wir bringen das Gleichungssystem in Zeilenstufenform. Wegen $m < n$ gibt es ein $i < m$ mit $N(i) + 2 \leq N(i + 1)$ und $N(i) + 2 \leq n$ oder es ist $N(m) \leq n - 2$. Lösen wir das Gleichungssystem also wie in Lemma 1.2.3, so ist $\xi^{N(i)+2}$ bzw. ξ^n frei wählbar. \square

Korollar 1.2.5. Sei (L) ein beliebiges lineares Gleichungssystem wie in (1.1) mit $m = n$, also mit genausovielen Gleichungen wie Variablen. Dann gilt

- (i) Besitzt (HL) nur die triviale Lösung, so besitzt (L) genau eine Lösung.
(ii) Besitzt (HL) eine nichttriviale Lösung, so besitzt (L) entweder keine oder unendlich viele Lösungen.

(Dieser Teil gilt mit unverändertem Beweis auch für $m \neq n$.)

Äquivalent dazu ist, dass (L) entweder eindeutig lösbar ist, oder (HL) eine nichttriviale Lösung besitzt.

Beweis. Verfahren wir wie in Theorem 1.2.2, so erhalten wir dasselbe homogene lineare Gleichungssystem, egal ob wir zunächst (L) auf Zeilenstufenform bringen und dann das homogene lineare Gleichungssystem dazu betrachten oder ob wir (HL) auf Zeilenstufenform bringen. Daher dürfen wir ohne Einschränkung annehmen, dass (L) bereits in Zeilenstufenform ist.

- (i) Besitzt (HL) nur die triviale Lösung, so hat (L) die Gestalt

$$\begin{array}{cccccc} a_1^1 x^1 & + & a_2^1 x^2 & + & \dots & + & a_n^1 x^n & = & b^1 \\ & & a_2^2 x^2 & + & \dots & + & a_n^2 x^n & = & b^2 \\ & & & & \ddots & & \vdots & & \vdots \\ & & & & & & a_n^n x^n & = & b^n. \end{array}$$

mit $a_i^i \neq 0$ für $1 \leq i \leq n$, hat also „obere Dreiecksgestalt“. Nach Lemma 1.2.3 besitzt (L) daher eine eindeutige Lösung.

- (ii) Sei $\eta = (\eta^1, \dots, \eta^n)$ eine nichttriviale Lösung des homogenen linearen Gleichungssystems (HL) , also eine Lösung mit $(\eta^1, \dots, \eta^n) \neq (0, \dots, 0)$. Nehme an, dass (L) eine Lösung $\xi = (\xi^1, \dots, \xi^n)$ besitzt. Dann sind $\xi + \lambda\eta$ für alle $\lambda \in \mathbb{R}$ Lösungen von (L) . Die Behauptung folgt. \square

1.3. Matrizen.

Definition 1.3.1. Eine $(m \times n)$ -Matrix A ist ein Element in $\mathbb{R}^{m \cdot n}$. Wir schreiben $A \in \mathbb{R}^{m \times n}$. A hat $m \cdot n$ reelle Komponenten, die wir mit a_j^i , $1 \leq i \leq m$, $1 \leq j \leq n$, bezeichnen: $A = (a_j^i)$ oder genauer $A = (a_j^i)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$. Wir stellen Matrizen wie folgt graphisch dar:

$$A = \begin{pmatrix} a_1^1 & a_2^1 & a_3^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & a_3^2 & \dots & a_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ a_1^m & a_2^m & a_3^m & \dots & a_n^m \end{pmatrix}.$$

Der obere Index bezieht sich auf die Zeilen, der untere auf die Spalten der Matrix. a_j^i steht daher in Zeile i und Spalte j .

Seien A, B zwei $(m \times n)$ -Matrizen mit Komponenten a_j^i bzw. b_j^i . Wir definieren die Summe durch

$$A + B := (a_j^i + b_j^i),$$

also durch komponentenweise Addition. Sei $\lambda \in \mathbb{R}$. Wir definieren das Produkt einer Matrix mit einer reellen Zahl komponentenweise durch

$$\lambda A := (\lambda a_j^i).$$

Bemerkung 1.3.2. Wir können n -Tupel $\xi \in \mathbb{R}^n$ als $(1 \times n)$ -Matrizen auffassen. Dann stimmen die Operationen Summieren und Produktbildung mit einer reellen Zahl mit den für n -Tupel bereits definierten Operationen überein.

Definition 1.3.3. Eine $(n \times n)$ -Matrix heißt quadratische Matrix. Die $(n \times n)$ -Matrix

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

heißt Einheitsmatrix der Ordnung n . Man schreibt auch $I = \mathbb{1}$. Wir definieren das Kroneckersymbol durch

$$\delta_j^i := \begin{cases} 1, & i = j, \\ 0, & \text{sonst.} \end{cases}$$

Dann gilt $I = (\delta_j^i)$.

Eine quadratische Matrix heißt Diagonalmatrix, falls $a_j^i = 0$ für $i \neq j$ gilt. Die Elemente a_i^i heißen Diagonalelemente.

Definition 1.3.4 (Matrixmultiplikation). Sei A eine $(m \times n)$ -Matrix und B eine $(n \times p)$ -Matrix, wie bisher mit Einträgen (a_j^i) und (b_k^j) . Wenn die Spaltenzahl von A mit der Zeilenzahl von B wie angegeben übereinstimmt, so definieren wir das Produkt $C = AB$, eine $(m \times p)$ -Matrix (c_k^i) , für $1 \leq i \leq m$ und $1 \leq k \leq p$ durch

$$c_k^i := \sum_{j=1}^n a_j^i b_k^j \equiv a_1^i b_k^1 + a_2^i b_k^2 + \dots + a_n^i b_k^n.$$

(Physiker benutzen die Einsteinsche Summenkonvention und schreiben $c_k^i = a_j^i b_k^j$.)

Beispiele 1.3.5. Es gilt

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

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

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

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

Wie das letzte Beispiel zeigt, gilt im allgemeinen nicht $AB = BA$, selbst, wenn beide Verknüpfungen definiert sind. Gilt $AB = BA$, so sagen wir, dass die beiden Matrizen vertauschen oder kommutieren.

Lemma 1.3.6. Sei A eine $(m \times n)$ -Matrix, B, C seien $(n \times p)$ -Matrizen, D eine $(p \times q)$ -Matrix und $\lambda \in \mathbb{R}$. Dann gelten die Rechenregeln

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

Beweis. Es gilt

$$\begin{aligned} A(B + C) &= \left(\sum_{j=1}^n a_j^i (b_k^j + c_k^j) \right)_{\substack{1 \leq i \leq m \\ 1 \leq k \leq p}} = \left(\sum_{j=1}^n a_j^i b_k^j \right)_{\substack{1 \leq i \leq m \\ 1 \leq k \leq p}} + \left(\sum_{j=1}^n a_j^i c_k^j \right)_{\substack{1 \leq i \leq m \\ 1 \leq k \leq p}} \\ &= AB + AC, \\ (\lambda A)B &= \left(\sum_{j=1}^n (\lambda a_j^i) b_k^j \right)_{\substack{1 \leq i \leq m \\ 1 \leq k \leq p}} = \underbrace{\left(\sum_{j=1}^n a_j^i (\lambda b_k^j) \right)_{\substack{1 \leq i \leq m \\ 1 \leq k \leq p}}}_{=A(\lambda B)} = \lambda \underbrace{\left(\sum_{j=1}^n a_j^i b_k^j \right)_{\substack{1 \leq i \leq m \\ 1 \leq k \leq p}}}_{=\lambda(AB)}, \\ A(BD) &= \left(\sum_{j=1}^n a_j^i \left(\sum_{k=1}^p b_k^j d_l^k \right) \right)_{\substack{1 \leq i \leq m \\ 1 \leq l \leq q}} = \left(\sum_{k=1}^p \left(\sum_{j=1}^n a_j^i b_k^j \right) d_l^k \right)_{\substack{1 \leq i \leq m \\ 1 \leq l \leq q}} \\ &= (AB)D. \quad \square \end{aligned}$$

Bemerkung 1.3.7. Sei

$$\sum_{k=1}^n a_k^i x^k = b^i, \quad 1 \leq i \leq m,$$

ein lineares Gleichungssystem. Wir definieren die Koeffizientenmatrix A und die augmentierte Koeffizientenmatrix A' zu diesem linearen Gleichungssystem durch

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^m & a_2^m & \dots & a_n^m \end{pmatrix} \quad \text{und} \quad A' = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 & b^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 & b^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^m & a_2^m & \dots & a_n^m & b^m \end{pmatrix}.$$

(In A' verwendet man gelegentlich auch einen senkrechten Strich vor der letzten Spalte.) Definiere weiterhin

$$b = \begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^m \end{pmatrix} \quad \text{und} \quad x = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}.$$

Dann können wir das Gleichungssystem in der Form

$$Ax = b$$

schreiben.

Alle Aussagen dieses Kapitels lassen sich nicht nur für lineare Gleichungssysteme und Matrizen im Reellen zeigen, sondern auch für beliebige Körper (die wie im nächsten Kapitel definieren werden). Lediglich die Aussagen über unendlich viele Lösungen sind in mehrere Lösungen abzuändern.

2. KÖRPER UND VEKTORRÄUME

2.1. Körper. Bereits bekannte Beispiele von Körpern sind die reellen Zahlen \mathbb{R} , die komplexen Zahlen \mathbb{C} oder die rationalen Zahlen \mathbb{Q} . Allgemein definieren wir

Definition 2.1.1 (Körper). Eine Menge F (engl.: field) mit einer Addition „+“ und einer Multiplikation „ \cdot “ heißt Körper, wenn für alle $a, b, c \in F$ die folgenden Axiome erfüllt sind. (Die Verknüpfungen Addition und Multiplikation sind Abbildungen $F \times F \rightarrow F$, die wir in der Form $a + b$ oder $a \cdot b$ schreiben.)

- (F1) $a + b = b + a$ (für alle $a, b \in F$) (Kommutativgesetz der Addition)
 (F2) $a + (b + c) = (a + b) + c$ (Assoziativgesetz der Addition)
 (F3) Es gibt ein neutrales Element $0 \in F$ der Addition mit $a + 0 = a$ für alle $a \in F$.
 (F4) Zu jedem $a \in F$ gibt es ein (eindeutig bestimmtes) Element $-a \in F$, so dass $a + (-a) = 0$ gilt.
 (F5) $a \cdot b = b \cdot a$ (Kommutativgesetz der Multiplikation)
 (F6) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (Assoziativgesetz der Multiplikation)
 (F7) Es gibt ein neutrales Element $1 \in F$ der Multiplikation, $0 \neq 1$, mit $a \cdot 1 = a$ für alle $a \in F$.
 (F8) Zu jedem $a \in F$ mit $a \neq 0$ gibt es ein (eindeutig bestimmtes) Element $a^{-1} \in F$ mit $a \cdot a^{-1} = 1$.
 (F9) $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ (Distributivgesetz)

Die eingeklammerten Eindeutigkeitsforderungen kann man auch beweisen.

Beispiele 2.1.2.

- (i) Die Mengen $\mathbb{R}, \mathbb{Q}, \mathbb{C}$ mit der üblichen Addition und Multiplikation bilden jeweils einen Körper.
 (ii) $F := \{a + b\sqrt{3} \equiv a + (b\sqrt{3}) : a, b \in \mathbb{Q}\} \equiv \mathbb{Q}[\sqrt{3}] \subset \mathbb{R}$ mit der üblichen Addition und Multiplikation von reellen Zahlen ist ein Körper.
 (iii) Sei p eine Primzahl. Definiere die folgenden Teilmengen von \mathbb{Z}

$$\begin{aligned}\bar{0} &:= \{\dots, -3p, -2p, -p, 0, p, 2p, 3p, \dots\}, \\ \bar{1} &:= \{\dots, 1 - 3p, 1 - 2p, 1 - p, 1, 1 + p, 1 + 2p, 1 + 3p, \dots\}, \\ \bar{2} &:= \{\dots, 2 - 3p, 2 - 2p, 2 - p, 2, 2 + p, 2 + 2p, 2 + 3p, \dots\}, \\ \bar{a} &:= \{\dots, a - 3p, a - 2p, a - p, a, a + p, a + 2p, a + 3p, \dots\}, \quad a \in \mathbb{Z}.\end{aligned}$$

Wir benutzen hier die Konvention $kp + q := (k \cdot p) + q$. Setze

$$\mathbb{Z}/p\mathbb{Z} := \{\bar{0}, \bar{1}, \bar{2}, \dots, \overline{p-1}\}.$$

Auf $\mathbb{Z}/p\mathbb{Z}$ definieren wir Addition und Multiplikation durch

$$\bar{a} + \bar{b} := \overline{a + b} \quad \text{und} \quad \bar{a} \cdot \bar{b} := \overline{a \cdot b}.$$

(Hier ist noch die Wohldefiniertheit nachzuweisen, d. h. wir müssen zeigen, dass $\overline{a + kp} + \overline{b + lp} = \overline{a + b}$ und $\overline{a + kp} \cdot \overline{b + lp} = \overline{a \cdot b}$ für alle $k, l \in \mathbb{Z}$ gelten.)

Wir multiplizieren also jeweils vertreterweise. Dann ist

$$\mathbb{F}_p := \{\bar{0}, \bar{1}, \bar{2}, \dots, \overline{p-1}\}$$

mit der oben definierten Addition und Multiplikation ein Körper. (Wir schreiben die Elemente auch wieder als $0, 1, 2, \dots, p - 1$, falls keine Verwechslungsgefahr besteht.)

Beweis. Übung. □

Im Beispiel \mathbb{F}_p haben wir implizit bereits einen Quotientenraum betrachtet. Wir wollen dies nun auch noch allgemein untersuchen.

Definition 2.1.3 (Relationen). Sei A eine Menge. Eine binäre Relation R ist eine Teilmenge von $A \times A$. Seien $a, b \in A$. Wir schreiben aRb und definieren, dass aRb genau dann wahr ist, wenn $(a, b) \in R$ ist.

Häufige Eigenschaften von Relationen sind

- (i) Reflexivität: Es gilt aRa oder $(a, a) \in R$ für alle $a \in R$.
- (ii) Symmetrie: Gilt aRb , so auch bRa .
- (iii) Antisymmetrie: Gilt aRb und bRa , so ist $a = b$.
- (iv) Transitivität: Gilt aRb und bRc , so auch aRc .
- (v) Eine Relation heißt total oder linear, falls aRb oder bRa für beliebige $a, b \in A$ gilt.
- (a) Eine reflexive, symmetrische und transitive Relation heißt Äquivalenzrelation.
- (b) Eine reflexive, transitive, antisymmetrische und totale/lineare Relation heißt Ordnung oder Totalordnung.
- (c) Eine reflexive, transitive, antisymmetrische Relation heißt Halbordnung.

Beispiele 2.1.4.

- (i) Auf den reellen Zahlen gibt es folgende wichtige Relationen:
 - (a) „ $<$ “ ($a < b$ ist genau dann wahr, falls a strikt kleiner als b ist),
 - (b) „ $=$ “ ($a = b$ ist genau dann wahr, falls a und b gleich sind),
 - (c) „ $>$ “ ($a > b$ ist genau dann wahr, falls a strikt größer als b ist),
 - (d) „ \leq “ ($a \leq b$ ist genau dann wahr, falls $a <$ oder $a = b$ wahr ist),
 - (e) „ \geq “ ($a \geq b$ ist genau dann wahr, falls $a >$ oder $a = b$ wahr ist) und
 - (f) „ \neq “ ($a \neq b$ ist genau dann wahr, falls a und b verschiedene reelle Zahlen sind).

(Genauer werden diese Relationen auf den reellen Zahlen in der Analysis im Zusammenhang mit geordneten Körpern besprochen.)

$=$ ist eine Äquivalenzrelation, \leq und \geq sind Ordnungen.

- (ii) Sei $0 < q \in \mathbb{Z}$. Auf \mathbb{Z} definieren wir eine Äquivalenzrelation \sim durch $n \sim m$ für $n, m \in \mathbb{Z}$, wenn $n - m$ durch q teilbar ist, d. h. wenn es ein $r \in \mathbb{Z}$ gibt, so dass $n - m = qr$ ist. Wir schreiben auch

$$n \equiv m \pmod{q}.$$

(Rechnen modulo 24 verwenden wir bei der Uhr: 2 Stunden nach 23 Uhr ist es 1 Uhr.)

- (iii) Sei A eine Menge. Dann ist „ \subset “ eine Halbordnung auf $\mathcal{P}(A)$.

Definition 2.1.5 (Äquivalenzklassen, Partition). Sei A eine Menge und \sim eine Äquivalenzrelation auf A . Sei $a \in A$. Dann definieren wir die Äquivalenzklasse von a , \bar{a} oder $[a]$ durch

$$\bar{a} \equiv [a] := \{x \in A : x \sim a\}.$$

Die Menge aller Äquivalenzklassen wird mit A/\sim bezeichnet:

$$A/\sim := \{[a] : a \in A\}.$$

Sei A eine Menge. Eine Partition ist eine Teilmenge P der Potenzmenge von A , $P \subset \mathcal{P}(A)$, also eine Menge von Teilmengen von A deren Elemente nichtleere, disjunkte Teilmengen von M sind, so dass jedes Element von A in einem Element von P als Element enthalten ist:

- (i) $\emptyset \notin P$,
- (ii) $U, V \in P \implies U \cap V = \emptyset$ oder $U = V$,
- (iii) $\bigcup_{U \in P} U = A$.

Lemma 2.1.6. Sei A eine Menge und \sim eine Äquivalenzrelation auf A . Dann ist A/\sim eine Partition von A .

Beweis. Übung. □

Theorem 2.1.7. *Sei F ein Körper. Dann gilt für alle $a, b \in F$*

- (i) $0 \cdot a = a \cdot 0 = 0$,
- (ii) $a \cdot (-b) = -(a \cdot b) = (-a) \cdot b$,
- (iii) $-(-a) = a$,
- (iv) $(a^{-1})^{-1} = a$ für $a \neq 0$,
- (v) $(-1) \cdot a = -a$,
- (vi) $(-a) \cdot (-b) = a \cdot b$,
- (vii) $(-a)^{-1} = -(a^{-1})$ für $a \neq 0$.

Beweis.

- (i) Es gilt $0 \cdot a \stackrel{(F3)}{=} (0+0) \cdot a \stackrel{(F5),(F9)}{=} 0 \cdot a + 0 \cdot a$. Die rechte Seite lesen wir als $(0 \cdot a) + (0 \cdot a)$ nach der Regel „Punkt vor Strich“. Addiere nun auf beiden Seiten $-(0 \cdot a)$. Es folgt $0 = 0 \cdot a$. Nun folgt $a \cdot 0 = 0$ aus der Kommutativität der Multiplikation.
- (ii) Es gilt $0 = a \cdot (b + (-b)) = a \cdot b + a \cdot (-b)$. Addiere nun auf beiden Seiten $-(a \cdot b)$ und erhalte $a \cdot (-b) = -(a \cdot b)$. Die andere Gleichheit folgt analog.
- (iii) Es gelten $0 = a + (-a)$ und $0 = (-a) + (-(-a))$, also auch $(-a) + (-(-a)) = a + (-a)$. Durch Addition von a dürfen wir auf beiden Seiten den Term $(-a)$ streichen. Die Behauptung folgt.
- (iv) Verfahre wie beim Beweis von (iii) und verwende Körperaxiom (F8) statt (F4) und Multiplikation statt Addition.
- (v) Es gilt $a + (-1) \cdot a = 1 \cdot a + (-1) \cdot a = (1 + (-1)) \cdot a = 0 \cdot a = 0$ nach (i) für die letzte Gleichheit. Addiere nun auf beiden Seiten $(-a)$ und erhalte $(-1) \cdot a = -a$.
- (vi) Wir multiplizieren $a + (-a) = 0$ mit $(-b)$ und erhalten

$$a \cdot (-b) + (-a) \cdot (-b) = 0 \cdot (-b) = 0.$$

Andererseits folgt aus $(-b) + b = 0$ durch Multiplikation mit a

$$a \cdot (-b) + a \cdot b = a \cdot 0 = 0.$$

Gleichsetzen und auf beiden Seiten $a \cdot (-b)$ abziehen liefert $(-a) \cdot (-b) = a \cdot b$.

- (vii) Wegen $a \neq 0$ und somit $-a \neq 0$ gibt es $(-a)^{-1}$. Es gilt

$$1 = a \cdot a^{-1} \stackrel{(vi)}{=} (-a) \cdot (-a^{-1}).$$

Multipliziere nun beide Seiten von links mit $(-a)^{-1}$. Die Behauptung folgt. □

2.2. Vektorräume. Die Lösungsmenge eines homogenen linearen Gleichungssystems ist ein Beispiel für einen Vektorraum.

Definition 2.2.1. Sei F ein Körper. Ein Vektorraum V über F (oder ein F -Vektorraum) ist eine nichtleere Menge zusammen mit einer Addition „+“ und einer Multiplikation „ \cdot “ mit den folgenden Eigenschaften:

- Die Addition $+$: $V \times V \rightarrow V$ erfüllt die folgenden Axiome
 - (V1) $a + b = b + a$ für alle $a, b \in V$, (Kommutativität)
 - (V2) $a + (b + c) = (a + b) + c$ für alle $a, b, c \in V$ (Assoziativität)
 - (V3) Es gibt ein (eindeutig bestimmtes) Element, den Nullvektor 0 , mit $a + 0 = a$ für alle $a \in V$. (neutrales Element)
 - (V4) Zu jedem $a \in V$ gibt es ein (eindeutig bestimmtes) Inverses $(-a) \in V$, so dass $a + (-a) = 0$ gilt. (additives Inverses)
- Die skalare Multiplikation \cdot : $F \times V \rightarrow V$ erfüllt
 - (V5) $\lambda \cdot (a + b) = (\lambda \cdot a) + (\lambda \cdot b)$ für alle $\lambda \in F, a, b \in V$

- (V6) $(\lambda + \mu) \cdot a = (\lambda \cdot a) + (\mu \cdot a)$ für alle $\lambda, \mu \in F, a \in V$ (Distributivgesetze)
 (V7) $(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot a)$ für alle $\lambda, \mu \in F, a \in V$ (Assoziativität)
 (V8) $1 \cdot a = a$ für alle $a \in V$.

Üblicherweise schreibt man (nicht nur hier) λa statt $\lambda \cdot a$. Die eingeklammerten Eindeutigkeitsforderungen kann man auch hier wieder beweisen.

Bemerkung 2.2.2. Ohne (V8) könnte man auch $\lambda \cdot a := 0$ für alle $\lambda \in F$ und alle $a \in V$ setzen.

Beispiele 2.2.3.

- (i) \mathbb{R}^n mit der üblichen Addition und Skalarmultiplikation, in Koordinaten

$$\begin{aligned} (\xi^1, \dots, \xi^n) + (\eta^1, \dots, \eta^n) &:= (\xi^1 + \eta^1, \dots, \xi^n + \eta^n), \\ \lambda \cdot (\xi^1, \dots, \xi^n) &:= (\lambda \xi^1, \dots, \lambda \xi^n), \end{aligned}$$

ist ein \mathbb{R} -Vektorraum.

- (ii) \mathbb{C}^n , der Raum der komplexen n -Tupel mit ebensolchen Verknüpfungen ist ein \mathbb{C} -Vektorraum.
 (iii) \mathbb{Q}^n ist ein \mathbb{Q} -Vektorraum.
 (iv) Sei F ein Körper. Dann ist F^n ein F -Vektorraum.
 (v) Sei $F = \mathbb{R}$. Dann ist die Menge der Folgen reeller Zahlen

$$\mathbb{R}^{\mathbb{N}} = \{(\xi^0, \xi^1, \xi^2, \dots) : \xi^i \in \mathbb{R} \text{ für alle } i \in \mathbb{N}\}$$

mit den Verknüpfungen

$$\begin{aligned} (\xi^0, \xi^1, \xi^2, \dots) + (\eta^0, \eta^1, \eta^2, \dots) &:= (\xi^0 + \eta^0, \xi^1 + \eta^1, \xi^2 + \eta^2, \dots), \\ \lambda \cdot (\xi^0, \xi^1, \xi^2, \dots) &:= (\lambda \xi^0, \lambda \xi^1, \lambda \xi^2, \dots) \end{aligned}$$

ein \mathbb{R} -Vektorraum. Die Teilmengen der konvergenten Folgen oder der Nullfolgen bilden ebenfalls einen Vektorraum.

Folgenräume können wir auch für andere Körper F definieren: $F^{\mathbb{N}}$, den Raum aller Folgen mit Werten in F . Dies ist ebenfalls ein Vektorraum.

- (vi) Die Menge der Lösungen eines linearen homogenen Gleichungssystems wie in (1.1) bilden einen Vektorraum. (Für nicht homogene lineare Gleichungssysteme ist dies nicht richtig.)
 (vii) \mathbb{C}^n ist ein \mathbb{R} -Vektorraum mit der üblichen Addition und der Skalarmultiplikation

$$\lambda (\xi^1, \dots, \xi^n) := (\lambda \xi^1, \dots, \lambda \xi^n).$$

- (viii) Sei A eine Menge. Die Menge der Funktionen $f: A \rightarrow \mathbb{R}$ bildet einen \mathbb{R} -Vektorraum. Die Verknüpfungen sind

$$\begin{aligned} (f + g)(x) &:= f(x) + g(x) \quad \text{für } x \in A, \\ (\lambda f)(x) &:= \lambda(f(x)) \quad \text{für } x \in A. \end{aligned}$$

- (ix) Die Menge der stetigen Funktionen $f: [0, 1] \rightarrow \mathbb{R}$ bildet einen \mathbb{R} -Vektorraum $C^0([0, 1])$ mit wie für beliebige Funktionen definierter Addition und Skalarmultiplikation.

- (x)

$$\{f \in C^0([0, 1]) : f(0) = 0\}$$

bildet mit der üblichen Addition und Multiplikation einen Vektorraum.

- (xi) Polynome: Sei F ein Körper. Betrachte in $F^{\mathbb{N}}$ die Menge aller abbrechenden Folgen, die also die Form $(\xi_0, \xi_1, \dots, \xi_n, 0, 0, \dots)$ für ein $n \in \mathbb{N}$ haben. Dies ist mit der Addition und der Skalarmultiplikation wie bei Folgen in F ein Vektorraum. Wir schreiben auch mit Hilfe einer Variablen X

$$\xi_0 + \xi_1 X + \xi_2 X^2 + \dots + \xi_n X^n$$

und bezeichnen den Vektorraum als $F[X]$. Polynome können wir auch multiplizieren:

$$(\xi_0, \xi_1, \dots) \cdot (\eta_0, \eta_1, \dots) := (\xi_0 \cdot \eta_0, \xi_0 \cdot \eta_1 + \xi_1 \cdot \eta_0, \xi_0 \cdot \eta_2 + \xi_1 \cdot \eta_1 + \xi_2 \cdot \eta_0, \dots),$$

wobei allgemein an Stelle k der Eintrag $\sum_{i=0}^k \xi_i \eta_{k-i}$ steht. In vertrauterer Notation mit der Variablen X schreiben wir

$$\begin{aligned} & (\xi_0 + \xi_1 X + \xi_2 X^2 + \dots + \xi_m X^m) \cdot (\eta_0 + \eta_1 X + \eta_2 X^2 + \dots + \eta_n X^n) \\ &= \xi_0 \eta_0 + (\xi_0 \eta_1 + \xi_1 \eta_0) X + \dots + \left(\sum_{i=0}^k \xi_i \eta_{k-i} \right) X^k + \dots + \xi_m \eta_n X^{m+n}. \end{aligned}$$

Die Multiplikation von Polynomen werden eigentlich erst später im Zusammenhang mit „Ringen“ untersuchen. Wir schreiben z. B. $P(X)$ für ein Polynom und $P(a)$, wenn wir statt der Variablen X überall $a \in F$ einsetzen.

Theorem 2.2.4. *Sei V ein F -Vektorraum. Dann gilt für alle $a, b \in V$ und $\lambda, \mu \in F$*

- (i) $0 \cdot a = 0$
- (ii) $\lambda \cdot 0 = 0$
- (iii) Aus $\lambda \cdot a = 0$ folgt $a = 0$ oder $\lambda = 0$.
- (iv) $(-\lambda) \cdot a = \lambda \cdot (-a) = -(\lambda \cdot a)$
- (v) Definiere eine Subtraktion durch $a - b := a + (-b)$ und $\lambda - \mu := \lambda + (-\mu)$. Dann gelten

$$\begin{aligned} \lambda \cdot (a - b) &= \lambda \cdot a - \lambda \cdot b, \\ (\lambda - \mu) \cdot a &= \lambda \cdot a - \mu \cdot a. \end{aligned}$$

Beweis.

- (i) Nach (V6) gilt

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

Wir wenden (V4) auf $0 \cdot a$ an und erhalten

$$\begin{aligned} 0 &= 0 \cdot a + (-(0 \cdot a)) \\ &= (0 \cdot a + 0 \cdot a) + (-(0 \cdot a)) \\ &= 0 \cdot a + (0 \cdot a + (-(0 \cdot a))) && \text{nach (V2)} \\ &= 0 \cdot a + 0 \\ &= 0 \cdot a && \text{nach (V3)}. \end{aligned}$$

- (ii) Nach (V5) gilt

$$\lambda \cdot 0 = \lambda \cdot (0 + 0) = \lambda \cdot 0 + \lambda \cdot 0.$$

Wir addieren den zu $\lambda \cdot 0$ inversen Vektor auf beiden Seiten und erhalten

$$\begin{aligned} 0 &= \lambda \cdot 0 + (-(\lambda \cdot 0)) \\ &= (\lambda \cdot 0 + \lambda \cdot 0) + (-(\lambda \cdot 0)) = \lambda \cdot 0 + (\lambda \cdot 0 + (-(\lambda \cdot 0))) \\ &= \lambda \cdot 0 + 0 && \text{nach (V2)} \\ &= \lambda \cdot 0 && \text{nach (V3)}. \end{aligned}$$

- (iii) Sei $\lambda \cdot a = 0$ und gelte $\lambda \neq 0$. Zeige also, dass $a = 0$ gilt. Da $\lambda \neq 0$ ist, existiert $\frac{1}{\lambda} \equiv \lambda^{-1}$. Es gilt

$$\begin{aligned} 0 &= \frac{1}{\lambda} \cdot 0 && \text{nach (ii)} \\ &= \frac{1}{\lambda}(\lambda \cdot a) \\ &= \left(\frac{1}{\lambda} \cdot \lambda\right) \cdot a && \text{nach (V7)} \\ &= 1 \cdot a \\ &= a && \text{nach (V8)}. \end{aligned}$$

- (iv) Es gilt

$$\lambda \cdot a + (-\lambda) \cdot a = (\lambda + (-\lambda)) \cdot a = 0 \cdot a = 0 = \lambda \cdot a + (-\lambda \cdot a).$$

Wir addieren auf beiden Seiten das Inverse von $\lambda \cdot a$ und erhalten

$$(-\lambda) \cdot a = -(\lambda \cdot a).$$

Ebenso addieren wir zu beiden Seiten von

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

auf beiden Seiten das Inverse von $\lambda \cdot a$ und erhalten $\lambda \cdot (-a) = -(\lambda \cdot a)$. Die Behauptung folgt.

- (v) Unter Benutzung von (iv) erhalten wir

$$\lambda \cdot (a - b) = \lambda \cdot (a + (-b)) = \lambda \cdot a + \lambda \cdot (-b) = \lambda \cdot a + (-\lambda \cdot b) = \lambda \cdot a - \lambda \cdot b.$$

Ebenso folgt

$$(\lambda - \mu) \cdot a = (\lambda + (-\mu)) \cdot a = \lambda \cdot a + (-\mu) \cdot a = \lambda \cdot a + (-\mu \cdot a) = \lambda \cdot a - \mu \cdot a.$$

□

Bemerkung 2.2.5. Das Assoziativgesetz (V2) gilt auch für mehr als drei Summanden, ebenso gilt das Kommutativgesetz (V1) für mehr als zwei Summanden. Der Beweis ist nicht kompliziert, nur umständlich aufzuschreiben. Man benutzt Induktion nach der Anzahl der Summanden. Beispiele sind:

$$\begin{aligned} (a + (b + c)) + d &= ((a + b) + c) + d = (a + b) + (c + d) \\ &= a + (b + (c + d)) = a + ((b + c) + d), \\ a + b + c &= a + c + b = c + a + b = c + b + a = b + c + a = b + a + c. \end{aligned}$$

Daher wollen wir zukünftig Klammern weglassen und benutzen das Summenzeichen bei assoziativer und kommutativer Summation

$$\sum_{i=1}^n a_i = a_1 + a_2 + \dots + a_n.$$

Dann gelten die üblichen Rechenregeln

- (i) $\lambda \cdot \sum_{i=1}^n a_i = \sum_{i=1}^n \lambda \cdot a_i$ für alle $\lambda \in F$ und $a_1, \dots, a_n \in V$,
(ii) $\left(\sum_{i=1}^n \lambda_i\right) \cdot a = \sum_{i=1}^n \lambda_i \cdot a$ für alle $\lambda_1, \dots, \lambda_n \in F$ und $a \in V$,
(iii) $\sum_{i=1}^n \lambda^i a_i + \sum_{i=1}^n \mu^i a_i = \sum_{i=1}^n (\lambda^i + \mu^i) a_i$ für alle $\lambda^1, \dots, \lambda^n, \mu^1, \dots, \mu^n \in F$ und $a_1, \dots, a_n \in V$.

Beweis. Übung.

□

2.3. Linearkombinationen.

Definition 2.3.1 (Linearkombination). Sei V ein F -Vektorraum. Sei $S \subset V$ nicht-leer. Gilt

$$a = \sum_{i=1}^n \lambda^i a_i \quad \text{für } \lambda^1, \dots, \lambda^n \in F \quad \text{und } a_1, \dots, a_n \in S \quad \text{und ein } n \in \mathbb{N},$$

so sagen wir, dass a als Linearkombination von Vektoren aus S dargestellt ist. (Beachte, dass die Summen in Linearkombinationen stets endlich sind.)

Die Menge aller Vektoren, die sich als Linearkombination von Vektoren in S darstellen lässt, heißt lineare Hülle von S ; wir schreiben $\langle S \rangle$. Ist S eine endliche Menge, so schreiben wir auch $\langle a_1, \dots, a_m \rangle$ statt $\langle \{a_1, \dots, a_m\} \rangle$. Wir sagen dann auch, dass $\langle a_1, \dots, a_m \rangle$ aus den Linearkombinationen von a_1, \dots, a_m besteht.

Beispiele 2.3.2.

- (i) Ist V ein F -Vektorraum und $S = \{a_1, \dots, a_p\} \subset V$. Dann ist

$$\langle S \rangle = \{ \lambda^1 a_1 + \dots + \lambda^p a_p : \lambda^1, \dots, \lambda^p \in F \}.$$

- (ii) Im Vektorraum $V = \mathbb{R}[X]$, der reellen Polynome in X sei

$$S = \{1, X^2, X^4, \dots, X^{2n}, \dots\}.$$

Dann ist $\langle S \rangle = \{P(X) \in \mathbb{R}[X] : P(a) = P(-a) \text{ für alle } a \in \mathbb{R}\}$.

- (iii) Sei $V = \mathbb{R}^3$ und $S = \{(1, 1, 0), (1, 0, 1)\}$. Es folgt

$$\begin{aligned} \langle S \rangle &= \{a(1, 1, 0) + b(1, 0, 1) : a, b \in \mathbb{R}\} \\ &= \{(a+b, a, b) : a, b \in \mathbb{R}\} \\ &= \{(x, y, z) : x - y - z = 0, x, y, z \in \mathbb{R}\}. \end{aligned}$$

Ist $S' = \{(1, 1, 0), (2, 1, 1)\}$, so gilt $\langle S \rangle = \langle S' \rangle$.

2.4. Unterräume.

Definition 2.4.1. Sei V ein Vektorraum über F . Eine nichtleere Teilmenge $W \subset V$, so dass für alle $a, b \in W$ und alle $\lambda, \mu \in F$ auch $\lambda a + \mu b \in W$ gilt, heißt Unterraum von V .

Ein Unterraum eines F -Vektorraumes V ist mit der Addition und Skalarmultiplikation von V ein F -Vektorraum.

Beispiele 2.4.2.

- (i) Sei V ein Vektorraum. Dann sind V und $\{0\}$ (triviale) Unterräume von V .
(ii) Sei $V = F[X]$. Der Grad eines Polynomes $P(X) = a_0 + a_1 X + \dots + a_m X^m$ ist das kleinste n , so dass für alle $i > n$ für die Koeffizienten $a_i = 0$ gilt. Wir schreiben $\deg P = n$ und setzen $\deg 0 \equiv -\infty$. Sei $m \leq n$. Der Vektorraum der Polynome vom Grad kleiner oder gleich m ist ein Unterraum der Polynome vom Grad kleiner oder gleich n .
(iii) Sei $c \in [0, 1]$. Dann ist $\{f \in C([0, 1]) : f(c) = 0\}$ ein Unterraum von $C([0, 1])$.
(iv) Die Menge der Lösungen eines linearen homogenen Gleichungssystems in n Variablen sind ein Unterraum von F^n .

Hier und im folgenden betrachten wir Vektorräume über einem beliebigen aber festen Körper F , falls dies nicht ausdrücklich anders angegeben ist.

Theorem 2.4.3. Sei V ein Vektorraum, $S \subset V$. Dann ist $\langle S \rangle$ der kleinste Unterraum von V , der S enthält.

Beweis. Wegen $\sum_{\emptyset} = 0$ ist $\langle S \rangle \neq \emptyset$. $\langle S \rangle$ ist ein Unterraum, denn für $a, b \in \langle S \rangle$ mit $a = \sum_{i \in I} a^i s_i$ und $b = \sum_{j \in J} b^j s_j$ für endliche Mengen I, J und $\lambda, \mu \in F$ gilt auch

$$\lambda a + \mu b = \sum_{i \in I \cap J} (\lambda a^i + \mu b^i) s_i + \sum_{i \in I \setminus J} \lambda a^i s_i + \sum_{j \in J \setminus I} \mu b^j s_j \in \langle S \rangle.$$

Somit ist $\langle S \rangle$ ein Unterraum von V . $\langle S \rangle$ enthält insbesondere die Vektoren von S und daher auch alle Linearkombinationen von Vektoren aus S .

Sei $W \subset V$ ein beliebiger Unterraum von V mit $S \subset W$. Dann enthält W nach Definition auch alle Linearkombinationen von Vektoren aus S und somit $\langle S \rangle$. Die Behauptung folgt. \square

Theorem 2.4.4. *Sei V ein Vektorraum, $S \subset V$, $b \in V$. Sei b eine Linearkombination von Vektoren aus S . Dann gilt $\langle S \cup \{b\} \rangle = \langle S \rangle$.*

Beweis. Es gilt $S \cup \{b\} \subset \langle S \rangle$. $\langle S \rangle$ ist ein Vektorraum. Somit folgt nach Theorem 2.4.3 $\langle S \cup \{b\} \rangle \subset \langle S \rangle$.

$\langle S \rangle \subset \langle S \cup \{b\} \rangle$ ist klar. \square

Definition 2.4.5. Wir sagen, dass $\langle S \rangle$ von S erzeugt ist. Im Fall $\langle S \rangle = V$ heißt S ein Erzeugendensystem von V . Besitzt ein Vektorraum V ein endliches Erzeugendensystem, so heißt V endlich erzeugt bzw. endlich erzeugbar.

Beispiele 2.4.6.

- (i) Sei $A \in F^{m \times n}$. Fassen wir die Spalten von A als Vektoren in F^m auf, so nennen wir den von ihnen erzeugten Unterraum den Spaltenraum von A . Ebenso ist der Zeilenraum von A der von den Zeilen von A in F^n erzeugte Unterraum.
- (ii) Sei $V = \mathbb{R}^n$, $S = \{e_1, \dots, e_n\}$ mit

$$\begin{aligned} e_1 &= (1, 0, 0, \dots, 0), \\ e_2 &= (0, 1, 0, \dots, 0), \\ e_3 &= (0, 0, 1, \dots, 0), \\ &\vdots \\ e_n &= (0, 0, 0, \dots, 1). \end{aligned}$$

(Für später eigentlich besser als Spaltenvektoren geschrieben.) Dann ist $\langle S \rangle = V$ und somit ist S ein Erzeugendensystem von V . Für $\xi = (\xi^1, \dots, \xi^n)$ gilt

$$\xi = \sum_{i=1}^n \xi^i e_i. \text{ Somit ist } \mathbb{R}^n \text{ endlich erzeugbar.}$$

- (iii) Sei $V = \mathbb{R}[X]$. Dann ist $S = \{1, X, X^2, \dots\}$ ein Erzeugendensystem, da sich jedes Polynom als Linearkombination von Termen der Form X^i schreiben lässt. $\mathbb{R}[X]$ ist aber nicht endlich erzeugbar, da ein endliches Erzeugendensystem nicht Polynome beliebigen Grades erzeugen kann.

Theorem 2.4.7. *Sei $(W_i)_{i \in I}$ eine beliebige Familie von Unterräumen von V . Dann ist $W := \bigcap_{i \in I} W_i$ ein Unterraum von V .*

Beweis. Der Schnitt ist nicht leer, da jeder Unterraum die Null enthält. Seien $a, b \in W$, also $a, b \in W_i$ für alle $i \in I$. Seien $\lambda, \mu \in F$. Dann gilt $\lambda a + \mu b \in W_i$ für alle $i \in I$ und somit ist auch $\lambda a + \mu b \in W$. \square

Im Allgemeinen ist die Vereinigung von Unterräumen eines Vektorraumes kein Unterraum mehr. Die folgende Konstruktion macht aus mehreren Unterräumen

einen Unterraum, der die Vereinigung aller dieser Unterräume enthält. Dies funktioniert auch für beliebig viele Unterräume; dann erklärt man die Summe als die Menge beliebiger endlicher Summen von Vektoren aus den Unterräumen.

Definition 2.4.8. Seien W_1, W_2, \dots, W_p Unterräume eines F -Vektorraumes V . Definiere deren Summe durch

$$W_1 + W_2 + \dots + W_p := \{a_1 + a_2 + \dots + a_p : a_i \in W_i, i = 1, \dots, p\}.$$

Theorem 2.4.9. Sei V ein F -Vektorraum. Seien W_1, \dots, W_p Unterräume von V . Dann ist $W := W_1 + \dots + W_p$ der kleinste Unterraum von V , der jedes W_i enthält.

Beweis. Zeige zunächst, dass W ein Unterraum von V ist: Seien $a = a_1 + \dots + a_p, b = b_1 + \dots + b_p \in W$ mit $a_i, b_i \in W_i$. Seien $\lambda, \mu \in F$. Dann ist

$$\lambda a + \mu b = (\lambda a_1 + \mu b_1) + (\lambda a_2 + \mu b_2) + \dots + (\lambda a_p + \mu b_p).$$

Somit ist $\lambda a + \mu b \in W$ und W ist ein (nichtleerer) Unterraum.

Wegen $0 + \dots + 0 + a_i + 0 + \dots + 0 \in W$ für $a_i \in W_i$ gilt $W_i \subset W$. Sei W' ein weiterer Unterraum von V mit $W_i \subset W'$ für alle i . Zeige, dass $W \subset W'$ gilt: Sei $a_1 + \dots + a_p \in W$ beliebig. Da $a_i \in W_i \subset W'$ ist, folgt auch $a_1 + \dots + a_p \in W'$, denn W' ist ein Unterraum. Wir erhalten $W \subset W'$. \square

3. STRUKTUR VON VEKTORRÄUMEN

3.1. Lineare Unabhängigkeit.

Definition 3.1.1. Sei $\{a_1, \dots, a_n\}$ eine endliche Familie von Vektoren in einem F -Vektorraum. Dann heißt $\{a_1, \dots, a_n\}$ linear unabhängig, wenn aus

$$\sum_{i=1}^n \lambda^i a_i = 0$$

für $\lambda^i \in F$ bereits $\lambda^1 = \dots = \lambda^n = 0$ folgt. Andernfalls heißt diese Familie linear abhängig.

Wir sagen auch, dass die Vektoren a_1, \dots, a_n linear unabhängig sind.

Eine beliebige Familie heißt linear unabhängig, wenn dies für jede endliche Teilfamilie $\{a_1, \dots, a_n\}$ gilt. Andernfalls heißt sie linear abhängig.

Beispiele 3.1.2.

- (i) $\{a\}$ ist genau dann linear unabhängig, wenn $a \neq 0$ gilt.
- (ii) Ist S linear abhängig und $S \subset T$, so ist T linear abhängig.
- (iii) $\{a + tb : t \in \mathbb{R}\}$ beschreibt genau dann eine Gerade, die den Ursprung 0 nicht enthält, wenn a, b linear unabhängig sind.
- (iv) Die Menge $\{(2, -1, 1), (1, 0, 0), (0, 2, 1)\} \equiv \{a_1, a_2, a_3\}$ ist linear unabhängig, denn für eine Linearkombination der Null gilt $0 = \lambda^1 a_1 + \lambda^2 a_2 + \lambda^3 a_3$, also

$$\begin{array}{rcccc} 2\lambda^1 & + & \lambda^2 & + & & = & 0, \\ -\lambda^1 & & & + & 2\lambda^3 & = & 0, \\ \lambda^1 & & & + & \lambda^3 & = & 0 \end{array}$$

und diese lineare Gleichungssystem besitzt nur die Nulllösung.

- (v) Sei $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ mit Eintrag 1 an der Stelle i . Dann sind die Vektoren $\{e_1, \dots, e_n\}$ in F^n linear unabhängig.
- (vi) $\{1, X, X^2, X^3, \dots\}$ sind in $\mathbb{R}[X]$ linear unabhängig.

Theorem 3.1.3. Eine Familie von Vektoren $\{a_1, \dots, a_n\}$ ist genau dann linear unabhängig, wenn sich keiner der Vektoren als Linearkombination der anderen schreiben lässt.

Beweis.

(i) Sei zunächst ohne Einschränkung

$$a_1 = \sum_{i=2}^n \lambda^i a_i$$

für $\lambda^i \in F$, $i = 2, \dots, n$. Setze $\lambda^1 := -1$. Dann gilt

$$\sum_{i=1}^n \lambda^i a_i = 0$$

mit $\lambda^1 = -1 \neq 0$. Somit ist die Familie $\{a_1, \dots, a_n\}$ linear abhängig.

(ii) Gelte

$$\sum_{i=1}^n \lambda^i a_i = 0$$

und sei ohne Einschränkung $\lambda^1 \neq 0$. Dann folgt

$$a_1 = -\frac{1}{\lambda^1} \sum_{i=2}^n \lambda^i a_i.$$

Somit haben wir a_1 als Linearkombination der übrigen Vektoren dargestellt. \square

Theorem 3.1.4. *Eine Familie S von Vektoren ist genau dann linear unabhängig, wenn jedes $a \in \langle S \rangle$ nur auf genau eine Art und Weise linear aus den Vektoren aus S kombiniert werden kann.*

Beweis.

- (i) Insbesondere ist $0 \in \langle S \rangle$. Der Nullvektor lässt sich als Linearkombination mit Koeffizienten 0 schreiben. Da dies nach Voraussetzung die einzige Möglichkeit ist, sind die Vektoren in S linear unabhängig.
- (ii) Sei ein Vektor auf zwei verschiedene Arten dargestellt, gelte also

$$\sum_{i=1}^n \lambda^i a_i = \sum_{i=1}^n \mu^i a_i$$

mit $(\lambda^1, \dots, \lambda^n) \neq (\mu^1, \dots, \mu^n)$. Dann ist

$$\sum_{i=1}^n (\lambda^i - \mu^i) a_i = 0$$

eine nichttriviale Linearkombination der Null. Die Vektoren sind also linear abhängig. \square

3.2. Basen.

Definition 3.2.1. Sei V ein Vektorraum. Dann ist $S \subset V$ eine Basis von V , wenn S linear unabhängig ist und $\langle S \rangle = V$ gilt.

Beispiele 3.2.2.

- (i) $\{e_1, \dots, e_n\}$ ist eine Basis des F^n , die Standardbasis von F^n .
- (ii) $\{1, X, X^2, \dots\}$ ist eine Basis von $F[X]$.

Als Korollar zu Theorem 3.1.4 erhalten wir

Korollar 3.2.3. $S \subset V$ ist genau dann eine Basis von V , wenn sich jeder Vektor in eindeutiger Weise als Linearkombination von Vektoren aus S schreiben lässt.

Theorem 3.2.4. Sei $S \subset V$ linear unabhängig. Sei $b \in V$. Ist $b \notin \langle S \rangle$, so ist $S \cup \{b\}$ linear unabhängig.

Beweis. Sei $\{a_1, \dots, a_n\}$ eine endliche Teilmenge von S . Zeige, dass $\{a_1, \dots, a_n, b\}$ linear unabhängig ist: Gelte

$$\mu b + \sum_{i=1}^n \lambda^i a_i = 0.$$

Dann folgt zunächst $\mu = 0$, denn sonst wäre b als Linearkombination der $\{a_i\}$ darstellbar. Aufgrund der linearen Unabhängigkeit von S folgt dann aber auch $\lambda^i = 0$ für alle i . Somit ist $S \cup \{b\}$ linear unabhängig. \square

Mit Hilfe dieses Satzes kann man linear unabhängige Teilmengen sukzessive vergrößern.

Definition 3.2.5. Sei $S \subset V$ linear unabhängig. Dann heißt S maximal oder maximal linear unabhängig, wenn $S \cup \{b\}$ für jedes $b \in V \setminus S$ linear abhängig ist.

Theorem 3.2.6. Sei $S \subset V$. Dann ist S genau dann eine Basis von V , wenn S linear unabhängig und maximal ist.

Beweis.

- (i) Sei S maximal und linear unabhängig. Gäbe es $b \in V \setminus \langle S \rangle$, so wäre $S \cup \{b\}$ nach Theorem 3.2.4 auch linear unabhängig. Dies widerspricht der Maximalität. Somit gibt es kein solches b und wir erhalten $\langle S \rangle = V$. Da S nach Voraussetzung linear unabhängig ist, handelt es sich um eine Basis.
- (ii) Ist S eine Basis, so ist S nach Definition linear unabhängig. Da jedes $b \in V$ eine Linearkombination von Vektoren aus S ist, ist $S \cup \{b\}$ für beliebiges $b \in V$ nach Theorem 3.1.3 linear abhängig. Somit ist S bereits maximal. \square

Das Zornsche Lemma erlaubt uns nun, zu zeigen, dass jeder Vektorraum eine Basis besitzt. Das Zornsche Lemma ist äquivalent zum Auswahlaxiom. Dies wird in einer Vorlesung über Logik oder Mengenlehre bewiesen.

Lemma 3.2.7 (Zornsches Lemma). Sei A nichtleer und \leq eine Halbordnung auf A . Besitzt jede total geordnete Teilmenge B eine obere Schranke in A , also ein $\beta \in A$ mit $b \leq \beta$ für alle $b \in B$, so besitzt A mindestens ein maximales Element m . (m ist maximales Element von A , falls aus $m \leq a$, $a \in A$ stets $a = m$ folgt.)

Theorem 3.2.8. Sei V ein Vektorraum. Dann besitzt V eine Basis.

Nach Konvention ist \emptyset die Basis von $\{0\}$, da $\sum_{\emptyset} = 0$ gilt.

Beweis. Die Relation \subset ist eine Halbordnung auf $\mathcal{P}(V)$ und daher auch auf der Menge aller linear unabhängigen Teilmengen \mathcal{U} von V . Sei also $\{A_i\}_{i \in I}$ eine total geordnete Teilmenge von \mathcal{U} , gelte also stets $A_i \subset A_j$ oder $A_j \subset A_i$. Setze $A := \bigcup_{i \in I} A_i$. Wir behaupten, dass A eine obere Schranke für $\{A_i\}_{i \in I}$ ist. $A_i \subset A$ ist klar. $i \in I$

Zur linearen Unabhängigkeit: Seien $a_1, \dots, a_n \in A$ und gelte $\sum_{i=1}^n \lambda^i a_i = 0$. Dann gibt es Mengen $A_{j(i)}$ mit $a_i \in A_{j(i)}$. Es gilt $A_{j(1)} \subset A_{j(2)}$ oder $A_{j(2)} \subset A_{j(1)}$. Setze $k(2) := j(2)$, falls $A_{j(1)} \subset A_{j(2)}$ gilt und sonst $k(2) := j(1)$. Dann gilt $A_{j(i)} \subset A_{k(2)}$ für alle $i = 1, 2$. Gelte per Induktion bereits $A_{j(i)} \subset A_{k(p)}$ für $1 \leq i \leq p$ und für ein $k(p) \in \{j(1), \dots, j(p)\}$. Gilt auch $A_{j(p+1)} \subset A_{k(p)}$, so setze $k(p+1) := k(p)$, sonst $k(p+1) := j(p+1)$. Es folgt $A_{j(i)} \subset A_{k(p)}$ für $1 \leq i \leq p+1$. Somit gibt es einen Index $k(n) \in \{j(1), \dots, j(n)\}$ mit $a_i \in A_{k(n)}$ für alle $1 \leq i \leq n$. Damit sind die Vektoren $\{a_i\}_{1 \leq i \leq n}$ linear unabhängig, A ist also eine obere Schranke der $(A_i)_{i \in I}$.

Das Zornsche Lemma impliziert nun, dass es ein maximales $B \in \mathcal{U}$ gibt. Nach Theorem 3.2.6 folgt $\langle B \rangle = V$. Damit ist B ein maximales linear unabhängiges Erzeugendensystem von V und somit eine Basis von V . \square

Definition 3.2.9. Sei S ein Erzeugendensystem von V . Dann heißt S minimal, falls es keine echte Teilmenge von S gibt, die ebenfalls V erzeugt.

Theorem 3.2.10. Eine Familie S von Vektoren ist genau dann eine Basis von V , wenn S ein minimales Erzeugendensystem von V ist.

Beweis.

- (i) Ist S nicht minimal, so gibt es $b \in S$, so dass $S \setminus \{b\}$ ebenfalls ein Erzeugendensystem von V ist. Insbesondere gilt dann aber auch

$$b = \sum_{i=1}^n \lambda^i a_i$$

für spezielle $\lambda^1, \dots, \lambda^n \in F$, $a_1, \dots, a_n \in S \setminus \{b\}$, $n \in \mathbb{N}$. Bringen wir b auf die andere Seite, so widerspricht dies der linearen Unabhängigkeit von S . Somit ist eine Basis minimal.

- (ii) Sei S ein minimales Erzeugendensystem. Ist S nicht linear unabhängig, so gibt es $a_1, \dots, a_n \in S$ und $\lambda^1, \dots, \lambda^n \in F$, $n \in \mathbb{N}$, mit

$$\sum_{i=1}^n \lambda^i a_i = 0$$

und $(\lambda^1, \dots, \lambda^n) \neq (0, \dots, 0)$. Nehme ohne Einschränkung an, dass $\lambda_1 \neq 0$ ist. Dann lässt sich a_1 als Linearkombination von a_2, \dots, a_n schreiben. Nach Theorem 2.4.4 gilt $\langle S \rangle = \langle S \setminus \{a_1\} \rangle$. Somit ist S nicht minimal. Widerspruch. Die Behauptung folgt. □

Korollar 3.2.11. Sei V Vektorraum, $S \subset V$. Nach den Theoremen 3.2.6 und 3.2.10 sind die folgenden Aussagen äquivalent:

- (i) S ist eine Basis von V ,
- (ii) S ist eine maximale linear unabhängige Teilmenge von V ,
- (iii) S ist ein minimales Erzeugendensystem von V .

Im endlich erzeugten Fall können wir auch absteigende Folgen von Erzeugendensystemen betrachten um eine Basis zu bekommen.

Theorem 3.2.12. Sei S ein endliches Erzeugendensystem von V . Dann gibt es eine Basis $S' \subset S$ von V .

Beweis. Ist S keine Basis, so ist S nach Korollar 3.2.11 nicht minimal. Somit gibt es $S_1 \subsetneq S$, das ebenfalls ein Erzeugendensystem ist. Ist S_1 keine Basis, so wiederholen wir diesen Schritt und erhalten $S_2 \subsetneq S_1, \dots$. Da S endlich ist, bricht die strikt absteigende Folge

$$S \supsetneq S_1 \supsetneq S_2 \supsetneq \dots \supsetneq S_k$$

bei einem $k \in \mathbb{N}$ ab. S_k ist minimal und somit nach Korollar 3.2.11 eine Basis. □

In den Übungen werden wir sehen, dass dieses Vorgehen ohne die Endlichkeit des Erzeugendensystems nicht funktioniert.

Es gilt aber auch allgemein

Theorem 3.2.13. Sei S ein Erzeugendensystem von V . Dann gibt es $S' \subset S$, so dass S' eine Basis von V ist.

Beweis. Argumentiere wie im Beweis von Theorem 3.2.8, benutze aber linear unabhängige Teilmengen von S statt von V . □

Theorem 3.2.14. Sei $S \subset V$ linear unabhängig. Dann besitzt V eine Basis B mit $S \subset B$.

Beweis. Gehe wie beim Beweis von Theorem 3.2.8 vor, benutze aber linear unabhängige Teilmengen, die S enthalten statt linearen Teilmengen von V . \square

3.3. Dimension.

Definition 3.3.1. Sei V ein Vektorraum mit Basis S . Ist S endlich und besitzt n Elemente, $n \in \mathbb{N}$, so hat V die Dimension n , $\dim V = n$.

Zunächst einmal ist nicht klar, ob die Dimension wohldefiniert ist, da es verschiedene Basen mit unterschiedlich vielen Elementen geben könnte. Wir werden aber in Theorem 3.3.3 noch zeigen, dass dies nicht der Fall ist.

Lemma 3.3.2 (Austauschsatz von Steinitz). *Sei V ein Vektorraum. Ist die Familie $\{a_1, \dots, a_p\}$ linear unabhängig und sei $\{b_1, \dots, b_n\}$ ein Erzeugendensystem von V . Dann gilt $p \leq n$ und nach einer Umnummerierung der b_i 's ist auch die Familie $\{a_1, \dots, a_p, b_{p+1}, \dots, b_n\}$ ein Erzeugendensystem von V .*

Beweis. Wir beweisen per Induktion nach k , dass $\{a_1, \dots, a_k, b_{k+1}, \dots, b_n\}$ für $0 \leq k \leq n$ nach Umnummerierung der b_i 's ebenfalls ein Erzeugendensystem von V ist. (Wäre $p > n$, so folgte aus diesem Beweis, dass insbesondere $\{a_1, \dots, a_n\}$ ein Erzeugendensystem von V ist. Dann wäre aber a_{n+1} aus a_1, \dots, a_n linear kombinierbar. Widerspruch.) Für $k = 0$ ist die Aussage klar. Sei also $\{a_1, \dots, a_k, b_{k+1}, \dots, b_n\}$ ein Erzeugendensystem von V . Es gibt also eine Linearkombination davon, die a_{k+1} darstellt

$$a_{k+1} = \sum_{i=1}^k \lambda^i a_i + \sum_{j=k+1}^n \mu^j b_j.$$

Da $a_{k+1} \neq 0$ ist und da die a_i 's linear unabhängig sind, ist ein $\mu^j \neq 0$, ohne Einschränkung $\mu^{k+1} \neq 0$. Daher ist $b_{k+1} \in \langle a_1, \dots, a_{k+1}, b_{k+2}, \dots, b_n \rangle$. Setze $S := \{a_1, \dots, a_k, b_{k+2}, \dots, b_n\}$. Nach Theorem 2.4.4 erhalten wir

$$\langle S \cup \{b_{k+1}\} \rangle = \langle S \cup \{a_{k+1}, b_{k+1}\} \rangle = \langle S \cup \{a_{k+1}\} \rangle.$$

Die Behauptung folgt nun per Induktion. \square

Aufgrund des folgenden Theorems ist die Dimension wohldefiniert.

Theorem 3.3.3. *Sei V ein Vektorraum mit Basen S und T . Habe S endlich viele Elemente n , $n \in \mathbb{N}$. Dann besitzt T ebenfalls n Elemente.*

Beweis. Seien $|T|$ und $|S|$ die Anzahlen der Elemente von T bzw. S . Dann gilt nach Lemma 3.3.2 $|T| \leq |S|$ und $|S| \leq |T|$. Die Behauptung folgt. \square

Korollar 3.3.4. *Sei V ein n -dimensionaler Vektorraum. Dann gelten folgende Aussagen.*

- (i) *Jedes Erzeugendensystem von V besitzt wenigstens n Vektoren.*
- (ii) *Jede Familie von strikt mehr als n Vektoren ist linear abhängig.*
- (iii) *Jedes Erzeugendensystem aus n Vektoren ist eine Basis von V .*
- (iv) *Jede Familie von n linear unabhängigen Vektoren ist eine Basis von V .*

Beweis.

- (i) Nach Theorem 3.2.13 gäbe es zu einem Erzeugendensystem mit weniger als n Elementen auch eine Basis mit weniger als n Elementen. Widerspruch zu Theorem 3.3.3.
- (ii) Nach Theorem 3.2.14 könnte man sonst diese Menge zu einer Basis mit mehr als n Elementen ergänzen. Dies widerspricht wiederum Theorem 3.3.3.
- (iii) Ein minimales Erzeugendensystem ist eine Basis.
- (iv) Eine maximale linear unabhängige Familie ist eine Basis. \square

Korollar 3.3.5. Sei W ein Unterraum eines Vektorraumes V mit $\dim V = n$. Dann gilt $\dim W \leq \dim V$ und Gleichheit gilt genau dann, wenn $V = W$ ist.

Beweis. Sei S eine Basis von W und T mit $S \subset T$ eine Basis von V . Dann gilt $|S| \leq |T|$. Bei Gleichheit gilt $S = T$ und die Erzeugnisse stimmen überein. Die Rückrichtung ist trivial. \square

Theorem 3.3.6 (Dimensionsformel). Seien $U, V \subset W$ endlichdimensionale Unterräume eines Vektorraumes W . Dann gilt

$$\dim U + V = \dim U + \dim V - \dim(U \cap V).$$

Beweis. Sei $w_1, \dots, w_{\dim U \cap V}$ eine Basis von $U \cap V$. Seien $u_1, \dots, u_{\dim U - \dim U \cap V}$ so gewählt, dass sie zusammen mit den w_i 's eine Basis von U bilden und seien $v_1, \dots, v_{\dim V - \dim U \cap V}$ so gewählt, dass sie zusammen mit den w_i 's eine Basis von V bilden. Wir behaupten nun, dass alle diese Vektoren zusammen eine Basis von $U + V$ bilden. Es ist klar, dass sie $U + V$ erzeugen. Zum Nachweis der linearen Unabhängigkeit sei

$$\underbrace{\sum_i c^i w_i}_{=: w \in U \cap V} + \underbrace{\sum_j a^j u_j}_{=: u \in U} + \underbrace{\sum_k b^k v_k}_{=: v \in V} = 0.$$

Dann ist $v \in V$ und $v = -w - u \in U$, also $v \in U \cap V$. Die w_i 's bilden zusammen mit den v_i 's eine Basis von V , Vektoren in $U \cap V$ werden aber bereits eindeutig als Linearkombination von w_i 's dargestellt. Die obige Darstellung von v enthält keine w_i 's. Somit ist $v = 0$. Ebenso ist $u = 0$. Die lineare Unabhängigkeit folgt. \square

3.3.1. Koordinaten.

Definition 3.3.7 (Koordinaten). Sei V ein F -Vektorraum mit einer (geordneten) Basis $\{a_1, \dots, a_n\}$. Wir schreiben auch (a_1, \dots, a_n) . Dann lässt sich nach Korollar 3.2.3 jedes $a \in V$ in eindeutiger Weise aus den a_i 's linear kombinieren

$$a = \sum_{i=1}^n \lambda^i a_i,$$

$\lambda^i \in F$. Die Koeffizienten $\lambda^1, \dots, \lambda^n$ heißen Koordinaten von a bezüglich der Basis $\{a_1, \dots, a_n\}$.

3.4. Lineare Abbildungen.

Definition 3.4.1 (Lineare Abbildung). Seien V, W zwei Vektorräume über einem Körper F . Eine Abbildung $f: V \rightarrow W$ heißt linear, falls

$$(3.1) \quad f(a + b) = f(a) + f(b) \quad \text{für alle } a, b \in V \text{ und}$$

$$(3.2) \quad f(\lambda a) = \lambda f(a) \quad \text{für alle } a \in V, \lambda \in F \text{ gelten.}$$

Lemma 3.4.2. Sei $f: V \rightarrow W$ linear. Dann gelten die folgenden Aussagen

$$(1) \quad f(0) = 0,$$

$$(2) \quad f\left(\sum_{i=1}^n \lambda^i a_i\right) = \sum_{i=1}^n \lambda^i f(a_i) \text{ für alle } \lambda^i \in F, a_i \in V, \text{ insbesondere ist}$$

$$f(\lambda a + \mu b) = \lambda f(a) + \mu f(b)$$

für alle $\lambda, \mu \in F$ und $a, b \in V$. Diese letzte Gleichheit ist äquivalent zu (3.1) und (3.2), auch wenn wir $\mu = 1$ wählen.

Beweis.

(i) Benutze (3.2) mit $\lambda = 0$.

(ii) Per Induktion erhalten wir

$$\begin{aligned}
 f\left(\sum_{i=1}^n \lambda^i a_i\right) &= f\left(\lambda^1 a_1 + \sum_{i=2}^n \lambda^i a_i\right) \\
 &\stackrel{(3.1)}{=} f(\lambda^1 a_1) + f\left(\sum_{i=2}^n \lambda^i a_i\right) \\
 &\stackrel{(3.2)}{=} \lambda^1 f(a_1) + \sum_{i=2}^n \lambda^i f(a_i) \quad \text{nach Induktionsannahme} \\
 &= \sum_{i=1}^n \lambda^i f(a_i).
 \end{aligned}$$

Die Gleichung (3.1) erhält man mit $\lambda = \mu = 1$, die Gleichung (3.2) mit $b = 0$. \square

Beispiele 3.4.3.

- (i) Die Identität $\mathbf{1} = \mathbf{1}_V: V \rightarrow V$, $\mathbf{1}(a) := a$ ist linear.
(ii) Seien $f_j^i \in F$, $1 \leq i \leq m$, $1 \leq j \leq n$. Dann ist $f: F^n \rightarrow F^m$ mit

$$f(\xi) \equiv f(\xi^1, \dots, \xi^n) \equiv (f^1(\xi), \dots, f^m(\xi)), \quad f^i(\xi) := \sum_{j=1}^n a_j^i \xi^j$$

linear.

- (iii) Seien $f: V \rightarrow W$ und $g: W \rightarrow X$ linear. Dann ist $g \circ f: V \rightarrow X$ ebenfalls linear.
(iv) Sei $P \in F[X]$. Dann ist die Abbildung D (Differenzieren)

$$D: P[X] = \sum_{i=0}^n a_i X^i \mapsto \sum_{i=1}^n i a_i X^{i-1}$$

eine lineare Abbildung.

- (v) Die Abbildung

$$C^0([0, 1]) \ni f \mapsto \left(x \mapsto \int_0^x f(t) dt \right) \in C^0([0, 1])$$

ist linear.

Definition 3.4.4. Eine lineare Abbildung $f: V \rightarrow W$ heißt Isomorphismus, wenn es eine Abbildung $g: W \rightarrow V$ mit $g \circ f = \mathbf{1}_V$ und $f \circ g = \mathbf{1}_W$ gibt. g heißt die zu f inverse Abbildung. Wir schreiben $g = f^{-1}$. (Wir werden erst in Lemma 3.4.5 sehen, dass g wohldefiniert ist.)

Zwei Vektorräume V und W heißen isomorph, wenn es einen Isomorphismus $f: V \rightarrow W$ gibt. Isomorph zu sein ist eine Äquivalenzrelation auf einer Kollektion von F -Vektorräumen.

Lemma 3.4.5. Sei $f: V \rightarrow W$ ein Isomorphismus. Dann ist die Inverse $g: W \rightarrow V$ von f eindeutig bestimmt und ebenfalls linear.

Beweis. Sei $h: W \rightarrow V$ eine weitere Abbildung mit $h \circ f = \mathbf{1}_V$ und $f \circ h = \mathbf{1}_W$. Dann gilt für alle $b \in W$

$$g(b) = \mathbf{1}_V \circ g(b) = (h \circ f) \circ g(b) = h \circ (f \circ g)(b) = h \circ \mathbf{1}_W(b) = h(b).$$

Somit gilt $h = g$.

g ist linear, denn es gilt

$$\begin{aligned} g(a) + g(b) &= \mathbf{1}_V(g(a) + g(b)) = g \circ f(g(a) + g(b)) \\ &= g(f \circ g(a) + f \circ g(b)) && \text{da } f \text{ linear ist} \\ &= g(a + b); \\ \lambda g(a) &= g \circ f(\lambda g(a)) = g(\lambda f \circ g(a)) && \text{da } f \text{ linear ist} \\ &= g(\lambda a). \end{aligned} \quad \square$$

Lemma 3.4.6 (Koordinatenabbildung). *Sei V ein F -Vektorraum mit geordneter Basis (a_1, \dots, a_n) . Dann ist die Abbildung $f: V \rightarrow F^n$,*

$$V \ni a = \sum_{i=1}^n \lambda^i a_i \mapsto (\lambda^1, \dots, \lambda^n) \in F^n,$$

ein Isomorphismus.

Beweis. Da die Darstellung von a eindeutig ist, ist f wohldefiniert. Die Linearität von f ist klar. Die Umkehrabbildung $g: F^n \rightarrow V$,

$$g(\lambda^1, \dots, \lambda^n) := \sum_{i=1}^n \lambda^i a_i,$$

ist ebenfalls offensichtlich linear und zu f invers. □

Korollar 3.4.7. *Sei V ein n -dimensionaler F -Vektorraum. Dann ist V isomorph zu F^n .*

Theorem 3.4.8. *Sei $f: V \rightarrow W$ linear. Sei $\{a_1, \dots, a_n\} \subset V$ linear abhängig. Dann ist $\{f(a_1), \dots, f(a_n)\} \subset W$ ebenfalls linear abhängig.*

Beweis. Aus $\sum_{i=1}^n \lambda^i a_i = 0$ folgt mit Hilfe der Linearität $\sum_{i=1}^n \lambda^i f(a_i) = 0$. □

Korollar 3.4.9. *Sei $f: V \rightarrow W$ ein Isomorphismus. Dann ist $\{a_1, \dots, a_n\}$ genau dann linear abhängig, wenn $\{f(a_1), \dots, f(a_n)\}$ linear abhängig ist.*

Theorem 3.4.10. *Seien V, W über F endlich erzeugbare Vektorräume. Dann sind V und W genau dann isomorph, falls $\dim V = \dim W$ gilt.*

Beweis.

- (i) Gelte zunächst $\dim V = \dim W$. Dann gibt es nach Lemma 3.4.6 Isomorphismen $\varphi: V \rightarrow F^n$ und $\psi: W \rightarrow F^n$. Somit ist $\psi^{-1} \circ \varphi: V \rightarrow W$ der gesuchte Isomorphismus.
- (ii) Sei $\{a_1, \dots, a_n\}$ eine Basis von V . Sei $f: V \rightarrow W$ ein Isomorphismus. Dann ist $\{f(a_i)\}$ linear unabhängig. Es ist auch eine Basis, denn sei $b \in W$ beliebig, dann gilt

$$V \ni f^{-1}(b) = \sum_{i=1}^n \lambda^i a_i$$

für geeignete $\lambda_i \in F$. Somit ist $b = f\left(\sum_{i=1}^n \lambda^i a_i\right) = \sum_{i=1}^n \lambda^i f(a_i)$, $\{f(a_i)\}$ ist also auch ein Erzeugendensystem von W und damit eine Basis. Da die Anzahl von Basiselementen in beiden Vektorräumen übereinstimmt, gilt $\dim V = \dim W$. □

3.5. Rechenmethoden. Sei V ein endlichdimensionaler Vektorraum, $S \subset V$ endlich. Wir wollen folgende Fragen konkret behandeln können:

- (i) Ist S linear unabhängig?
- (ii) Wie findet man eine Basis von $\langle S \rangle$, falls S linear abhängig ist?
- (iii) Wie findet man eine Basis von $\langle S \rangle$, die aus Elementen von S besteht?
- (iv) Gilt $a \in \langle S \rangle$, d. h. ist a aus Vektoren in S linear kombinierbar?

Wegen Lemma 3.4.6 und da diese Eigenschaften sich unter Isomorphismen sich nicht ändern, können wir ohne Einschränkung $V = F^n$ betrachten.

Notation 3.5.1. Sei $S = (a_1, \dots, a_m)$, $a_i \in F^n$.

- (i) $\langle S \rangle := \langle \{a_1, \dots, a_m\} \rangle$ (Ebenso werden wir die Anordnung der Vektoren vergessen, wenn es um Fragen wie z. B. lineare Unabhängigkeit geht.)
- (ii) $S_{ij} := (a_1, \dots, a_{i-1}, a_j, a_{i+1}, \dots, a_{j-1}, a_i, a_{j+1}, \dots, a_m)$; S_{ij} entsteht aus S durch Vertauschen des i -ten mit dem j -ten Vektor.
- (iii) $S_i(\lambda) := (a_1, \dots, a_{i-1}, \lambda a_i, a_{i+1}, \dots, a_m)$, $0 \neq \lambda \in F$; $S_i(\lambda)$ entsteht aus S durch Multiplikation des i -ten Vektors mit λ .
- (iv) $S_{ij}(\lambda) := (a_1, \dots, a_{i-1}, a_i + \lambda a_j, a_{i+1}, \dots, a_m)$, $\lambda \in F$, $i \neq j$; $S_{ij}(\lambda)$ entsteht aus S durch Addition des λ -fachen des Vektors a_j zu a_i .

Theorem 3.5.2.

- (i) Es gilt $\langle S \rangle = \langle S_{ij} \rangle = \langle S_i(\lambda) \rangle = \langle S_{ij}(\lambda) \rangle$ für alle λ , i, j wie oben.
- (ii) Ist eine der Familien S , S_{ij} , $S_i(\lambda)$, $S_{ij}(\lambda)$ linear unabhängig, so gilt dies auch für alle anderen dieser Familien.

Beweis.

- (i) (a) $\langle S \rangle = \langle S_{ij} \rangle$ ist klar.
- (b) Sei $a = \sum_{k=1}^m \mu^k a_k \in \langle S \rangle$. Dann ist $\sum_{k \neq i} \mu^k a_k + (\frac{1}{\lambda} \mu^i) (\lambda a_i) = a \in \langle S_i(\lambda) \rangle$.
 $\langle S_i(\lambda) \rangle \subset \langle S \rangle$ folgt analog.
- (c) Benutze nun, dass sich die Linearkombination einer Familie von Vektoren nicht ändert, wenn man einen linear kombinierbaren Vektor hinzufügt, vergleiche Theorem 2.4.4.

$$\begin{aligned} \langle S \rangle &= \langle a_1, \dots, a_i, \dots, a_m \rangle \\ &= \langle a_1, \dots, a_i, \dots, a_m, a_i + \lambda a_j \rangle \\ &= \langle a_1, \dots, a_i + \lambda a_j, \dots, a_m \rangle, \quad \text{da } a_i = (a_i + \lambda a_j) - \lambda a_j \text{ gilt,} \\ &= \langle S_{ij}(\lambda) \rangle. \end{aligned}$$

- (ii) Sei S linear abhängig. Man überlegt sich leicht, dass dann auch die anderen drei Familien von Vektoren linear abhängig sind. Insbesondere folgt aus $\mu^i a_i + \mu^j a_j = 0$ auch $\mu^i (a_i + \lambda a_j) + (\mu^j - \mu^i \lambda) a_j = 0$.

Sei umgekehrt eine der Familien S_{ij} , $S_i(\lambda)$, $S_{ij}(\lambda)$ linear abhängig. Die Operationen, die von S zu dieser Familie geführt haben, sind mit Hilfe derselben Operationen invertierbar: Ist $T = S_{ij}$, so ist $T_{ij} = S$; ist $T = S_i(\lambda)$, so ist $T_i(\frac{1}{\lambda}) = S$; ist $T = S_{ij}(\lambda)$, so ist $T_{ij}(-\lambda) = S$. Somit ist in diesem Fall auch S linear abhängig. Die Behauptung folgt. \square

Definition 3.5.3. Seien $a^i = (a_1^i, \dots, a_n^i)$, $1 \leq i \leq m$, Zeilenvektoren und

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^m & a_2^m & \dots & a_n^m \end{pmatrix}$$

die aus diesen Zeilenvektoren bestehende Matrix. Dann heißt der von den Zeilen (a_1^i, \dots, a_n^i) in F^n erzeugte Unterraum Zeilenraum der Matrix A . (Analog heißt der von den Spalten in F^m erzeugte Unterraum Spaltenraum der Matrix A .)

In Analogie zu den Operatoren auf Erzeugendensystemen definieren wir die folgenden elementaren Zeilenoperationen:

- (i) Vertausche die Zeilen i und j .
- (ii) Ersetze die i -te Zeile durch das λ -fache der i -ten Zeile für $0 \neq \lambda \in F$.
- (iii) Ersetze die i -te Zeile a^i durch $a^i + \lambda a^j$ für $j \neq i$ und $\lambda \in F$.

Wie bei linearen Gleichungssystemen sagen wir, dass eine Matrix in Zeilenstufenform ist, wenn die Anzahl der links stehenden Nullen von Zeile zu Zeile strikt wächst oder eine Zeile und alle folgenden Zeilen nur Nullen enthalten.

Theorem 3.5.4. *Eine $(m \times n)$ -Matrix A lässt sich auf Zeilenstufenform bringen.*

Beweis. Verfahre so wie beim Beweis, dass sich ein lineares Gleichungssystem auf Zeilenstufenform bringen lässt. \square

Lemma 3.5.5. *Sei A eine $(m \times n)$ -Matrix. Bringt man A auf Zeilenstufenform, so bilden die Nicht-Null-Zeilen eine Basis des Zeilenraumes von A .*

Beweis. Nach Theorem 3.5.2 ändern elementare Zeilenumformungen den Zeilenraum einer Matrix nicht. Da in Zeilenstufenform die Nicht-Null-Zeilen eine Basis des Zeilenraumes bilden, folgt die Behauptung. \square

Beispiel 3.5.6. Mit Hilfe elementarer Zeilenumformungen wollen wir den Zeilenraum bestimmen. Wir erhalten

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

3.6. Anwendungen auf lineare Gleichungssysteme. Gleichungen der Form

$$a_1x^1 + a_2x^2 + \dots + a_nx^n = b$$

bilden unter den Verknüpfungen

$$\begin{aligned} &(a_1x^1 + \dots + a_nx^n = b) + \\ &(a'_1x^1 + \dots + a'_nx^n = b') := ((a_1 + a'_1)x^1 + \dots + (a_n + a'_n)x^n = (b + b')), \\ &\lambda(a_1x^1 + \dots + a_nx^n = b) := ((\lambda a_1)x^1 + \dots + (\lambda a_n)x^n = (\lambda b)) \end{aligned}$$

einen Vektorraum V . $f: V \rightarrow F^{n+1}$ mit

$$(a_1x^1 + \dots + a_nx^n = b) \mapsto (a_1, \dots, a_n, b)$$

ist ein Isomorphismus.

Sei nun ein lineares Gleichungssystem gegeben. Eine Lösung des Gleichungssystems löst auch alle Gleichungen im von den linearen Gleichungen erzeugten Unterraum U (und eine Lösung aller Gleichungen des Unterraumes ist auch eine Lösung des ursprünglichen linearen Gleichungssystems). Wenden wir dies zweimal an, so sehen wir: Sei S eine Basis von U . Dann ist x genau dann eine Lösung des ursprünglichen Gleichungssystems, wenn x alle Gleichungen in S erfüllt.

Technisches Vorgehen: Um

$$\sum_{k=1}^n a_k^i b^k = b^i$$

für alle $1 \leq i \leq m$ zu lösen können wir alle Lösungen der Matrix bestimmen, die wir erhalten, wenn wir

$$\begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 & b^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 & b^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^m & a_2^m & \dots & a_n^m & b^m \end{pmatrix}$$

auf Zeilenstufenform gebracht haben. Dies ist einfach, wenn wir die Gleichungen von unten nach oben betrachten und das bisherige Ergebnis jeweils in die weiter oben stehenden Gleichungen einsetzen.

3.7. Der Rang.

Definition 3.7.1. Sei A eine $(m \times n)$ -Matrix

$$A = (a_{ij}^i)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^m & a_2^m & \dots & a_n^m \end{pmatrix}$$

über F .

- (i) Dann ist der Zeilenrang die Dimension des von den Zeilen b^1, \dots, b^m , $b^i = (a_1^i, \dots, a_n^i)$, in F^n erzeugten Vektorraumes.
- (ii) Der Spaltenrang ist die Dimension des von den Spalten d_1, \dots, d_n ,

$$d_j = (a_{ij}^i)_{1 \leq i \leq m} = \begin{pmatrix} a_j^1 \\ \vdots \\ a_j^m \end{pmatrix}$$

erzeugten Vektorraumes in F^m . Der von den Spalten erzeugte Vektorraum heißt Spaltenraum von A .

- (iii) Der Rang einer Matrix A ist der Zeilenrang von A , $\text{rang } A$.

Die Begriffsbildung bei der letzten Definition ist sinnvoll, denn es gilt

Theorem 3.7.2. Sei A eine Matrix. Dann ist der Zeilenrang von A gleich dem Spaltenrang von A .

Beweis. Wir wollen die Matrix auf Zeilenstufenform bringen. Für Matrizen in Zeilenstufenform gilt die Behauptung offensichtlich. Zeige also noch, dass sich weder Zeilenrang noch Spaltenrang unter elementaren Zeilenumformungen ändern.

Da sich sogar der Zeilenraum einer Matrix unter elementaren Zeilenumformungen nicht ändert, bleibt auch der Zeilenrang konstant.

Unter elementaren Zeilenumformungen kann sich der Spaltenraum ändern. Sei

$$\sum_{i=1}^n \lambda^i d_i = 0$$

eine lineare Abhängigkeit der Spalten. Dann bleibt diese unter elementaren Zeilenumformungen erhalten. Der Spaltenrang wird unter elementaren Zeilenumformungen also höchstens kleiner. Da sich elementare Zeilenumformungen aber invertieren lassen, bleibt er gleich.

Die Behauptung folgt. \square

Definition 3.7.3. Sei A eine $(n \times n)$ -Matrix. Dann heißt A

- (i) regulär, falls $\text{rang } A = n$ ist,
- (ii) sonst heißt sie singulär.

Theorem 3.7.4. *Sei A eine $(n \times n)$ -Matrix. Dann sind die folgenden Aussagen äquivalent:*

- (i) A ist regulär,
- (ii) $\text{rang } A = n$,
- (iii) die Zeilen sind linear unabhängig,
- (iv) die Spalten sind linear unabhängig,
- (v) das lineare Gleichungssystem $Ax = 0$ hat nur die triviale Lösung.

Beweis. Nach Definition und Theorem 3.7.2 ist nur noch die Äquivalenz zur ausschließlich trivialen Lösbarkeit zu zeigen.

$Ax = 0$ besitzt genau dann eine nichttriviale Lösung, wenn die Spalten von A linear abhängig sind. \square

Theorem 3.7.5. *Sei $Ax = 0$ ein lineares homogenes Gleichungssystem mit n Unbekannten. Dann hat der Raum der Lösungen die Dimension $n - \text{rang } A$.*

Beweis. Bringe das Gleichungssystem auf Zeilenstufenform. Dann kann man die Aussage direkt am Gleichungssystem ablesen. \square

Theorem 3.7.6. *Sei $Ax = b$ ein inhomogenes lineares Gleichungssystem. Sei $(A|b)$ die augmentierte Matrix. Dann ist das Gleichungssystem genau dann lösbar, wenn $\text{rang } A = \text{rang}(A|b)$ gilt.*

Beweis. Ist $Ax = b$ lösbar, so ist b eine Linearkombination der Spalten von A . Somit ändert sich der Spaltenrang durch Hinzunahme der Spalte b nicht.

Der Spaltenraum von A ist im Spaltenraum von $(A|b)$ enthalten. Da die Dimensionen nach Voraussetzung übereinstimmen, ist b im Spaltenraum von A enthalten. Die Linearkombination von b aus den Spalten von A liefert gerade die Lösung x . \square

3.8. Basiswechsel.

Herleitung 3.8.1. Sei V ein n -dimensionaler Vektorraum über F mit geordneten Basen $S = (a_1, \dots, a_n)$ und $T = (b_1, \dots, b_n)$. Sei $\xi \in V$ und gelte

$$\xi = \sum_{i=1}^n \lambda^i a_i = \sum_{k=1}^n \mu^k b_k.$$

Aufgrund der Basiseigenschaft gibt es Zahlen $(d_j^i)_{1 \leq i, j \leq n}$ mit

$$a_i = \sum_{k=1}^n d_i^k b_k.$$

Wir wollen nun herleiten, wie sich die Koeffizienten (λ^i) aus den Koeffizienten (μ^k) bestimmen lassen. Es gilt

$$\xi = \sum_{i=1}^n \lambda^i a_i = \sum_{i,k=1}^n \lambda^i d_i^k b_k = \sum_{k=1}^n \left(\sum_{i=1}^n \lambda^i d_i^k \right) b_k = \sum_{k=1}^n \mu^k b_k.$$

Aufgrund der eindeutigen Darstellbarkeit eines Vektors mit Hilfe einer Basis folgt

$$\mu^k = \sum_{i=1}^n \lambda^i d_i^k.$$

Herleitung 3.8.2. Seien (α_i) , (β_i) und (γ_i) geordnete Basen eines n -dimensionalen F -Vektorraumes V . Gelte

$$\beta_j = \sum_{k=1}^n a_j^k \alpha_k,$$

$$\gamma_i = \sum_{j=1}^n b_i^j \beta_j$$

und

$$\gamma_i = \sum_{k=1}^n c_i^k \alpha_k.$$

Dann folgt

$$\gamma_i = \sum_{j=1}^n b_i^j \beta_j = \sum_{j,k=1}^n b_i^j a_j^k \alpha_k.$$

Aufgrund der eindeutigen Darstellbarkeit in einer Basis erhalten wir

$$c_i^k = \sum_{j=1}^n b_i^j a_j^k.$$

In Matrizenform mit $A = (a_j^k)_{1 \leq k, j \leq n}$, $B = (b_i^j)$ und $C = (c_i^k)$ erhalten wir

$$C = AB.$$

Sei nun insbesondere $\alpha_i = \gamma_i$ für alle $1 \leq i \leq n$. Dann gilt $c_j^i = \delta_j^i$ und

$$\delta_i^k = \sum_{j=1}^n b_i^j a_j^k \quad \text{oder} \quad \mathbf{1} = AB.$$

Definition 3.8.3. Sei A eine $(n \times n)$ -Matrix über F . Eine $(n \times n)$ -Matrix B mit $AB = BA = \mathbf{1}$ heißt zu A inverse Matrix, $B = A^{-1}$.

Bemerkung 3.8.4. Besitzt A eine Inverse, so ist diese eindeutig bestimmt.

Beweis. Seien B und C Inverse. Aus $AC = I$ folgt $B = BI = B(AC) = (BA)C = IC = C$. Die Behauptung folgt. \square

Theorem 3.8.5. Sei A eine $(n \times n)$ -Matrix über F . Dann sind die folgenden Aussagen äquivalent:

- (i) A ist regulär
- (ii) A besitzt eine eindeutig bestimmte Inverse, d. h. eine $(n \times n)$ -Matrix mit $AB = BA = \mathbf{1}$.
- (iii) A ist die Matrix eines Basiswechsels in einem n -dimensionalen F -Vektorraum.

Beweis. Sei A die Matrix eines Basiswechsels. Da auch ein Basiswechsel in der anderen Richtung möglich ist, gibt es nach Herleitung 3.8.2 eine Matrix B mit $I = BA$. Vertauschen wir die Rollen der beiden Basen finden wir eine Matrix C mit $I = AC$. Es gilt $B = BI = B(AC) = (BA)C = IC = C$. B ist nach Bemerkung 3.8.4 eindeutig bestimmt.

Eine invertierbare Matrix ist regulär: Nach Theorem 3.7.4 ist A genau dann invertierbar, wenn das lineare Gleichungssystem $\sum_{i=1}^n a_i^j x^i = 0$, $1 \leq j \leq n$, nur die

triviale Lösung besitzt. Sei B die Inverse zu A . Somit gilt $\sum_{j=1}^n b_j^k a_i^j = \delta_i^k$ für alle $1 \leq i, k \leq n$. Wir erhalten

$$0 = \sum_{i,j=1}^n b_j^k a_i^j x^i = \sum_{i=1}^n \delta_i^k x^i = x^k$$

für $1 \leq k \leq n$. Somit ist jede Lösung trivial.

Sei schließlich A regulär. (e_1, \dots, e_n) ist eine geordnete Basis von F^n . Setze

$$a_k := \sum_{i=1}^k a_k^i e_i.$$

Die a_k 's sind also gerade die Spalten der Matrix A und sind damit linear unabhängig und somit eine Basis von F^n . Also ist A eine Basiswechselmatrix. \square

3.9. Gruppen. Reguläre Matrizen sind Beispiele für die algebraische Struktur einer Gruppe:

Definition 3.9.1 (Gruppe). Eine Gruppe ist eine Menge G zusammen mit

- (i) einer Verknüpfung „ \cdot “: $G \times G \rightarrow G$,
- (ii) einer Abbildung $G \rightarrow G$, die jedem Element $a \in G$ ein $a^{-1} \in G$ zuordnet und
- (iii) einem ausgezeichneten Element $e \in G$,

so dass die folgenden Axiome erfüllt sind:

- (G1) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ für alle $a, b, c \in G$ (Assoziativgesetz)
- (G2) $a \cdot e = a = e \cdot a$ für alle $a \in G$ (Neutrales Element, Einselement)
- (G3) $a \cdot a^{-1} = a^{-1} \cdot a = e$ für alle $a \in G$, d. h. a^{-1} ist die Inverse von a .
- (G4) Gilt $a \cdot b = b \cdot a$ für alle $a, b \in G$, so heißt die Gruppe G kommutativ oder abelsch.

Beispiele 3.9.2.

- (i) Die regulären $(n \times n)$ -Matrizen über einem Körper F mit der Matrizenmultiplikation bilden eine Gruppe, $GL_n(F)$ oder $GL(n, F)$.
- (ii) Die komplexen Zahlen mit Norm 1, also $z \cdot \bar{z} = 1$, bilden mit der Multiplikation eine abelsche Gruppe.
- (iii) Sei F ein Körper. Dann bilden die von Null verschiedenen Elemente mit der Multiplikation eine abelsche Gruppe, $F^\times = F \setminus \{0\}$.
- (iv) Sei F ein Körper. Dann ist $(F, +)$ eine additive Gruppe. (Die Bezeichnung additiv betont, dass die übliche additive Verknüpfung hier die Gruppenoperation ist.) Diese Gruppe ist abelsch.
- (v) $(\mathbb{Z}, +)$ ist eine additive abelsche Gruppe.

Definition 3.9.3 (Untergruppe). Sei G eine Gruppe. Dann ist $U \subset G$ eine Untergruppe von G , falls U nicht leer ist und mit $a, b \in U$ auch $a \cdot b \in U$ und $a^{-1} \in U$ folgen.

Beispiele 3.9.4.

- (i) $\{1, i, -1, -i\}$ bilden eine bezüglich der Multiplikation eine Untergruppe der multiplikativen Gruppe $\mathbb{C} \setminus \{0\}$ der komplexen Zahlen.
- (ii) Die geraden Zahlen bilden eine additive Untergruppe der ganzen Zahlen \mathbb{Z} .

3.10. Komplemente und direkte Summen.

Definition 3.10.1. Sei U ein Unterraum des Vektorraumes V . Dann ist W ein Komplement von U , wenn

- (i) $U \cap W = \{0\}$ und
- (ii) $U + W = V$ gelten.

(Die Summe ist hier wie in Definition 2.4.8 erklärt. Im allgemeinen gibt es mehrere Komplemente eines Unterraumes.)

Ist W ein Komplement von U in V , so heißt V die direkte Summe der Unterräume U und W , $V = U \oplus W$.

Beispiele 3.10.2.

- (i) Sei $V = \mathbb{R}^n$. Sei $\{a_1, \dots, a_n\}$ eine Basis von V . Dann ist $\langle a_{k+1}, \dots, a_n \rangle$ ein Komplement von $\langle a_1, \dots, a_k \rangle$. Ein anderes ist $\langle a_{k+1} + a_1, a_{k+2}, \dots, a_n \rangle$. Also gilt

$$V = \langle a_1, \dots, a_k \rangle \oplus \langle a_{k+1}, \dots, a_n \rangle.$$

- (ii) Sei $V = F[X]$ und $U = \{P(X) \in V : P(0) = 0\}$. Dann ist der Unterraum W , der alle konstanten Polynome enthält, ein Komplement von U in V .

Theorem 3.10.3. *Seien U, W Unterräume von V . Dann ist W genau dann ein Komplement von U in V , wenn jedes $v \in V$ sich in eindeutiger Weise als $u + w$ mit $u \in U$ und $w \in W$ schreiben lässt.*

Beweis. Gelte $U \oplus W = V$. Sei $v \in V$ beliebig. Dann gibt es wegen $U + W = V$ Vektoren $u \in U$ und $w \in W$ mit $u + w = v$. Seien $u' \in U$ und $w' \in W$ weitere Vektoren mit $u' + w' = v$. Dann folgt $u - u' = w' - w \in U \cap W = \{0\}$. Es folgen $u = u'$ und $w = w'$.

Aus der Darstellbarkeit folgt $U + W = V$. Wäre $U \cap W$ strikt größer als $\{0\}$, $0 \neq v \in U \cap W$, so könnten wir 0 nichttrivial als $v - v \in U + W$ kombinieren. Widerspruch. \square

Theorem 3.10.4. *Sei V ein Vektorraum und U ein Unterraum. Dann besitzt U ein Komplement W , $U \oplus W = V$.*

Beweis. Sei $\{a_i\}_{i \in I}$ eine Basis von U . Dann können wir diese Basis durch eine Familie $\{b_j\}_{j \in J}$ zu einer Basis von V ergänzen. Setze $\langle \{b_j\}_{j \in J} \rangle =: W$. Dann gilt $U + W = V$.

Zu $U \cap W = \{0\}$: Gelte nach einer Umbenennung der Vektoren a_i und b_i ohne Einschränkung

$$v = \sum_{i=1}^n \lambda^i a_i = \sum_{j=1}^m \mu^j b_j.$$

Dann folgt $0 = \sum_{i=1}^n \lambda^i a_i - \sum_{j=1}^m \mu^j b_j$. Da $\{a_i, b_j\}$ eine Basis von V ist, folgt $0 = \lambda^i = \mu^j$ für alle $1 \leq i \leq n$ und $1 \leq j \leq m$. Somit ist $v = 0$. \square

4. LINEARE ABBILDUNGEN

4.1. Dimensionsformel. Erinnerung: $f: V \rightarrow W$ heißt linear, falls

$$f(\lambda a + \mu b) = \lambda f(a) + \mu f(b)$$

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

Beispiel 4.1.1. Sei $V = U \oplus W$. Definiere $p_U: V \rightarrow V$, die Projektion auf U , wie folgt: Sei $v \in V$. Dann gibt es $u \in U$ und $w \in W$ mit $v = u + w$. Setze $p_U(v) := u$. p_U ist eine lineare Abbildung.

Definition 4.1.2. Sei $f: V \rightarrow W$ linear. Dann definieren wir

- (i) den Kern von f als

$$\ker f := \{v \in V : f(v) = 0\}$$

- (ii) und das Bild von f als

$$\operatorname{im} f := \{f(v) : v \in V\} = f(V) = \{w \in W : (\exists v \in V : f(v) = w)\}.$$

Theorem 4.1.3. Sei $f: V \rightarrow W$ linear. Dann ist $\ker f$ ein Unterraum von V und $\operatorname{im} f$ ein Unterraum von W .

Beweis.

- (i) Seien $v, w \in \ker f$, $\lambda, \mu \in F$. Dann gilt

$$f(\lambda v + \mu w) = \lambda f(v) + \mu f(w) = \lambda \cdot 0 + \mu \cdot 0 = 0.$$

Wegen $f(0) = 0$ ist $0 \in \ker f$ und $\ker f \neq \emptyset$.

- (ii) Seien $a, b \in \operatorname{im} f$ und $u, v \in V$ mit $f(u) = a$ sowie $f(v) = b$. Dann folgt für $\lambda, \mu \in F$

$$\lambda a + \mu b = \lambda f(u) + \mu f(v) = f(\lambda u + \mu v).$$

Schließlich ist wiederum wegen $f(0) = 0$ auch $0 \in \operatorname{im} f$. \square

Definition 4.1.4. Sei $f: V \rightarrow W$ eine (lineare) Abbildung. Dann heißt f

- (i) injektiv, falls aus $f(a) = f(b)$ stets $a = b$ folgt;
(ii) surjektiv, falls $\operatorname{im} f = W$ gilt (d. h. für jedes $w \in W$ existiert ein $a \in V$ mit $f(a) = w$).

Theorem 4.1.5. Sei $f: V \rightarrow W$ linear. Dann ist f genau dann injektiv, wenn $\ker f = \{0\}$ gilt.

Beweis.

- (i) Ist f injektiv, so folgt aus $f(a) = 0 = f(0)$, dass $a = 0$ gilt. Somit ist $\ker f = \{0\}$.
(ii) Ist f nicht injektiv, so finden wir $a \neq b \in V$ mit $f(a) = f(b)$. Somit ist $f(a - b) = 0$ und $0 \neq a - b \in \ker f$. \square

Theorem 4.1.6. Sei $f: V \rightarrow W$ linear. Dann ist f genau dann ein Isomorphismus, wenn f injektiv und surjektiv ist.

Beweis.

- (i) Sei f ein Isomorphismus. Dann existiert eine lineare Abbildung $g: W \rightarrow V$ mit $g \circ f = \mathbf{1}_V$ und $f \circ g = \mathbf{1}_W$. Sei $f(a) = f(b)$. Nach beidseitiger Anwendung von g erhalten wir

$$a = g \circ f(a) = g \circ f(b) = b.$$

Somit ist f injektiv. Sei $w \in W$ beliebig, $a := g(w)$. Dann folgt $f(a) = f \circ g(w) = w$. Somit ist f surjektiv.

- (ii) Sei f injektiv und surjektiv, so gibt es (Surjektivität) zu jedem $w \in W$ ein eindeutig (Injektivität) bestimmtes $v \in V$ mit $f(v) = w$. Definiere $g(w) := v$. $g: W \rightarrow V$ ist eine Funktion, d. h. auf ganz W eindeutig definiert. Dann gelten nach Definition $g \circ f = \mathbf{1}_V$ und $f \circ g = \mathbf{1}_W$. (Die Linearität von g hatten wir in der Definition 3.4.4 nicht gefordert, sie folgt aber aus Lemma 3.4.5.) \square

Lemma 4.1.7. Sei $\{a_i\}_{i \in I}$ eine Basis von V und seien $b_i \in W$, $i \in I$, beliebig. Dann gibt es genau eine lineare Abbildung $f: V \rightarrow W$ mit $f(a_i) = b_i$ für alle $i \in I$.

Beweis. Sei $J \subset I$ eine beliebige endliche Teilmenge. Sei $a = \sum_{i \in J} \lambda^i a_i$ eine beliebige Linearkombination der Vektoren a_i , also ein beliebiger Vektor in V . Definiere $f(a) := \sum_{i \in J} \lambda^i f(a_i) = \sum_{i \in J} \lambda^i b_i$.

Diese Abbildung f ist linear: Seien nämlich $a, b \in V$ und $\lambda, \mu \in F$ beliebig. Ohne Einschränkung wollen wir $a = \sum_{i=1}^n \alpha^i a_i$ und $b = \sum_{i=1}^n \beta^i a_i$ annehmen, dass sich also

beide Vektoren aus denselben Basisvektoren linear kombinieren lassen und dass diese die angegebenen Indices tragen. Es folgt

$$\begin{aligned} f(\lambda a + \mu b) &= f\left(\lambda \sum_{i=1}^n \alpha^i a_i + \mu \sum_{i=1}^n \beta^i a_i\right) = f\left(\sum_{i=1}^n (\lambda \alpha^i + \mu \beta^i) a_i\right) \\ &= \sum_{i=1}^n (\lambda \alpha^i + \mu \beta^i) b_i = \lambda \sum_{i=1}^n \alpha^i b_i + \mu \sum_{i=1}^n \beta^i b_i = \lambda f(a) + \mu f(b). \end{aligned}$$

Zur Eindeutigkeit: Sei $g: V \rightarrow W$ eine weitere solche Abbildung. Dann folgt für einen beliebigen Vektor $a = \sum_{j \in J} \lambda^j a_j$, $J \subset I$ endlich, ohne Einschränkung $J = \{1, \dots, n\}$,

$$g\left(\sum_{i=1}^n \lambda^i a_i\right) = \sum_{i=1}^n \lambda^i g(a_i) = \sum_{i=1}^n \lambda^i f(a_i) = f\left(\sum_{i=1}^n \lambda^i a_i\right). \quad \square$$

Korollar 4.1.8. Sei $\{a_i\}_{i \in I}$ eine linear unabhängige Teilmenge von V und seien $b_i \in W$, $i \in I$, beliebig. Dann gibt es eine lineare Abbildung $f: V \rightarrow W$ mit $f(a_i) = b_i$ für alle $i \in I$.

Beweis. Ergänze $\{a_i\}_{i \in I}$ zu einer Basis von V , wähle die zugehörigen b_i 's beliebig und wende Lemma 4.1.7 an. \square

Direkt aus dem Beweis von Lemma 4.1.7 folgt auch

Korollar 4.1.9. Sei $\{a_i\}_{i \in I}$ eine Basis von V und seien $b_i \in W$, $i \in I$, beliebig. Sei f die eindeutig bestimmte lineare Abbildung $f: V \rightarrow W$ mit $f(a_i) = b_i$. Dann gilt $\text{im } f = \langle \{b_i\}_{i \in I} \rangle$.

Theorem 4.1.10 (Dimensionsformel). Sei $f: V \rightarrow W$ eine lineare Abbildung. Sei $\dim V = n$. Dann gilt

$$\dim(\ker f) + \dim(\text{im } f) = n = \dim V.$$

Beweis. Da $\ker f$ ein Unterraum von V ist, gibt es eine Basis $\{a_1, \dots, a_p\}$ von $\ker f$. Ergänze diese zu einer Basis $\{a_1, \dots, a_n\}$ von V . Nach Korollar 4.1.9 ist $\text{im } f$ von

$$f(a_1) = 0, \dots, f(a_p) = 0, f(a_{p+1}), \dots, f(a_n)$$

erzeugt. Daher genügt es zu zeigen, dass die Vektoren $f(a_{p+1}), \dots, f(a_n)$ linear unabhängig sind. Sonst hätten wir λ^i 's, nicht alle gleich Null, mit $\sum_{i=p+1}^n \lambda^i f(a_i) = 0$,

also $0 = f\left(\sum_{i=p+1}^n \lambda^i a_i\right)$. Widerspruch, denn der Nicht-Nullvektor $\sum_{i=p+1}^n \lambda^i a_i$ läge dann ebenfalls in $\ker f$. Somit gilt

$$\dim \ker f + \dim \text{im } f = p + (n - p) = n = \dim V. \quad \square$$

Bemerkung 4.1.11. Definiert man im Beweis der Dimensionsformel den Unterraum U durch $U := \langle a_{p+1}, \dots, a_n \rangle$, so ist $V = \ker f \oplus U$. Dann ist $f|_U: U \rightarrow W$, die Einschränkung von f auf den Unterraum $U \subset V$ ebenfalls linear und der obige Beweis zeigt, dass $f|_U$ injektiv ist.

Definition 4.1.12. Sei G eine Gruppe, $U \subset G$ eine Untergruppe. Wir definieren auf G eine Relation \sim_U durch

$$a \sim_U b \iff b^{-1}a \in U.$$

Lemma 4.1.13. Die Relation \sim_U ist eine Äquivalenzrelation.

Beweis.

- (i) \sim_U ist reflexiv, denn es gilt $a^{-1}a = e \in U$.
- (ii) \sim_U ist symmetrisch, denn aus $a \sim_U b$ folgt $b^{-1}a \in U$, also auch $(b^{-1}a)^{-1} = a^{-1}b \in U$ und somit ist $b \sim_U a$.
- (iii) \sim_U ist transitiv, denn aus $a \sim_U b$ und $b \sim_U c$ folgen $b^{-1}a, c^{-1}b \in U$, also auch $c^{-1}bb^{-1}a = c^{-1}a \in U$ und somit $a \sim_U c$. \square

Die folgende Definition benötigen wir später nicht, wenn wir uns nur für Quotientenräume von abelschen Gruppen oder Vektorräumen interessieren.

Definition 4.1.14 (Normalteiler). Ein Normalteiler N einer Gruppe G ist eine Untergruppe von G , so dass für alle $g \in G$ die Relation $gN = Ng$ gilt, wobei $gN = \{ga : a \in N\}$ ist (Linksnebenklasse) und Ng analog definiert ist.

Lemma 4.1.15. N ist genau dann ein Normalteiler von G , wenn für alle $x \in N$ und für alle $g \in G$ auch $g^{-1}xg \in N$ gilt.

Beweis. $gN = Ng$ ist äquivalent zu $gNg^{-1} = N$ oder zu den beiden Aussagen

$$\forall x \in N \forall g \in G \exists y \in N : gxg^{-1} = y$$

und

$$\forall y \in N \forall g \in G \exists x \in N : gxg^{-1} = y.$$

Die erste Aussage ist gerade die Behauptung und die zweite ist äquivalent dazu (man multipliziert mit g von rechts und mit g^{-1} von links und setze $h = g^{-1}$). \square

Theorem 4.1.16. Sei G eine Gruppe mit Normalteiler N . Dann ist

$$G/N := \{gN : g \in G\}$$

mit vertreterweise definierter Verknüpfung eine Gruppe.

Beweis.

- (i) Setze $aN \cdot bN := (a \cdot b)N$. Dies ist wohldefiniert, denn seien $a' = an$ und $b' = bm$ mit $n, m \in N$, so folgt $a'N \cdot b'N = anbmN = anbN = anNb = aNb = abN$, wobei wir $nN = N$ verwendet haben. Zeige dies: Es gilt nach Definition einer Untergruppe $nN \subset N$ und Multiplikation mit n^{-1} von links liefert $N \subset n^{-1}N$; da $n \in N$ beliebig ist, folgt $nN = N$.
- (ii) Setze $(aN)^{-1} := a^{-1}N$.
- (iii) Das neutrale Element ist eN mit dem neutralen Element $e \in G$.
- (iv) Die Assoziativität $(aN \cdot bN) \cdot cN = aN \cdot (bN \cdot cN)$ folgt nach Definition der Multiplikation aus der Assoziativität von G .
- (v) Ebenso folgen $aN \cdot eN = aN = eN \cdot aN$
- (vi) und $aN \cdot a^{-1}N = eN = a^{-1}N \cdot aN$. \square

Theorem 4.1.17. Sei V ein F -Vektorraum. Sei $U \subset V$ ein Unterraum. Definiere \sim_U durch $a \sim_U b$ genau dann, wenn $a - b \in U$. Dann ist \sim_U eine Äquivalenzrelation. Der Faktorraum

$$V/U := \{v + U : v \in V\}$$

mit repräsentantenweise definierter Verknüpfung ist wiederum ein Vektorraum.

Beweis. $(V, +)$ ist eine abelsche Gruppe mit Untergruppe (und damit auch Normalteiler) U . Daher übertragen sich die Kommutativität der Addition (V1), die Assoziativität der Addition (V2), die Eigenschaft des Nullelementes (V3) und des Inversen (V4) direkt.

Die Skalarmultiplikation ist durch $\lambda \cdot (a + U) := \lambda \cdot a + U$ definiert. Sie ist wohldefiniert, denn aus $a + U = b + U$, also $a - b \in U$ folgt $\lambda a - \lambda b = \lambda(a - b) \in U$.

Die vertreterweise definierten Distributiv- und Assoziativgesetze der Skalarmultiplikation ((V5), (V6) und (V7)) übertragen sich direkt, ebenso gilt $1 \cdot (a + U) = (1 \cdot a) + U = a + U$ für $1 \in F$. \square

Bemerkung 4.1.18. Ein Diagramm aus Räumen und Pfeilen, die Abbildungen zwischen diesen Räumen entsprechen, heißt kommutativ, wenn folgendes gilt: Starte in einem Raum und folge den Pfeilen. Dann ist das Ergebnis in einem anderen Raum unabhängig von der speziellen Wahl der Pfeile. Beispiel:

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \varphi \downarrow & & \downarrow \psi \\ C & \xrightarrow{g} & D \end{array}$$

ist genau dann kommutativ, wenn $\psi(f(a)) = g(\varphi(a))$ für alle $a \in A$ gilt. Wir schreiben auch

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \varphi \downarrow & \text{//} & \downarrow \psi \\ C & \xrightarrow{g} & D \end{array}$$

um die Kommutativität anzudeuten.

Weiterhin verwenden wir die Pfeile $A \xrightarrow{i} B$ für eine injektive und $A \xrightarrow{p} B$ für eine surjektive Abbildung. „ \cong “ an einem Pfeil deutet an, dass es sich um einen Isomorphismus handelt.

Versionen des folgenden Theorems gibt es auch für andere Homomorphismen.

Theorem 4.1.19 (Homomorphiesatz). *Sei $f: V \rightarrow W$ eine lineare Abbildung. Dann gibt es eine Zerlegung von f*

$$\begin{array}{ccc} V & \xrightarrow{f} & W, \\ \pi \searrow & & \nearrow i \\ & V/\ker f \xrightarrow[\cong]{\bar{f}} \text{im } f & \end{array}$$

wobei $\pi: V \rightarrow V/\ker f$, $v \mapsto v + \ker f \equiv [v]$, die Projektionsabbildung, $i: \text{im } f \rightarrow W$ die Inklusionsabbildung $w \mapsto w$ und $\bar{f}: V/\ker f \rightarrow \text{im } f$ die von f induzierte Abbildung $\bar{f}([a]) := f(a)$ ist. Das obige Diagramm ist kommutativ. \bar{f} ist ein Isomorphismus.

Beweis. Sei $a \in V$. Dann gilt $\pi(a) = a + \ker f$, $\bar{f}(\pi(a)) = f(a)$ und $i(\bar{f}(\pi(a))) = f(a)$. Die Abbildung \bar{f} ist wohldefiniert, da für $[a] = [b]$ ein $c \in \ker f$ existiert, so dass $a = b + c$ ist. Somit folgt $f(a) = f(b) + 0$.

\bar{f} ist surjektiv, da $\text{im } f$ aus Punkten der Form $f(a)$ besteht und $\bar{f}([a]) = f(a)$ ist. \bar{f} ist injektiv, da aus $\bar{f}([a]) = \bar{f}([b])$ auch $f(a) = f(b)$ oder $f(a - b) = 0$ folgt. Somit ist $a - b \in \ker f$ und daher gilt $[a] = [b]$. \square

4.2. Matrizen. Wir wollen lineare Abbildungen mit Hilfe von Matrizen darstellen.

Sei $f: V \rightarrow W$ eine lineare Abbildung. Seien $\{a_1, \dots, a_n\}$ und $\{b_1, \dots, b_m\}$ Basen von V bzw. W . Nach Lemma 4.1.7 definieren die Bilder $f(a_j)$, $1 \leq j \leq n$, die Abbildung f eindeutig. Es gibt $(a_j^i)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$, so dass

$$f(a_j) = \sum_{i=1}^m a_j^i b_i$$

für alle $1 \leq j \leq n$ gilt. Sei $a = \sum_{j=1}^n x^j a_j$ ein beliebiger Vektor aus V mit Koordinaten (x^1, \dots, x^n) . Mit Hilfe von a_j^i wollen wir angeben, welche Koordinaten (y^1, \dots, y^m) sein Bild $f(a) = \sum_{i=1}^m y^i b_i$ hat. Es gilt

$$(4.1) \quad f(a) = f\left(\sum_{j=1}^n x^j a_j\right) = \sum_{j=1}^n x^j f(a_j) = \sum_{j=1}^n \sum_{i=1}^m x^j a_j^i b_i = \sum_{i=1}^m \left(\sum_{j=1}^n x^j a_j^i\right) b_i.$$

Es folgt

$$y^i = \sum_{j=1}^n x^j a_j^i.$$

Wir erhalten

Theorem 4.2.1. *Seien V, W endlichdimensionale F -Vektorräume. Sei (a_1, \dots, a_n) eine Basis von V und sei (b_1, \dots, b_m) eine Basis von W . Dann wird eine lineare Abbildung $f: V \rightarrow W$ vollständig durch eine $(m \times n)$ -Matrix A beschrieben. Es ist*

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^m & a_2^m & \dots & a_n^m \end{pmatrix}.$$

In der k -ten Spalte von A stehen dabei gerade die Koordinaten des Vektors $f(a_k)$ bezüglich der Basis (b_1, \dots, b_m) .

Umgekehrt definiert eine $(m \times n)$ -Matrix A vermöge (4.1) eine eindeutig bestimmte lineare Abbildung $f: V \rightarrow W$, so dass die k -te Spalte von A gerade die Koordinaten von $f(a_k)$ bezüglich der Basis (b_1, \dots, b_m) enthält.

A heißt die zur linearen Abbildung f gehörige Matrix bezüglich der beiden Basen (a_1, \dots, a_n) und (b_1, \dots, b_m) .

Weiterhin folgt aus den obigen Überlegungen

Theorem 4.2.2. *Sei f wie oben und seien die Basen wiederum fixiert. Dann erhalten wir aus f eine Abbildung für die Koordinatenvektoren*

$$x = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{pmatrix}.$$

Ist x der Koordinatenvektor von $a \in V$, so erfüllt der Koordinatenvektor y von $f(a) \in W$

$$y = Ax.$$

Theorem 4.2.3. *Seien $f: U \rightarrow V$ und $g: V \rightarrow W$ lineare Abbildungen mit zugehörigen Matrizen A bzw. B . Dann ist $D = BA$ die zu $g \circ f$ gehörige Matrix.*

Beweis. Seien $(\alpha_i)_{1 \leq i \leq l}$, $(\beta_j)_{1 \leq j \leq m}$ und $(\gamma_k)_{1 \leq k \leq n}$ Basen von U, V bzw. W . Die zu f gehörige Matrix sei $(a_j^i)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq l}}$ und die zu g gehörige Matrix sei $(b_j^i)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$. Es ist also zu zeigen, dass die zu $g \circ f: U \rightarrow W$ gehörige Matrix durch $(d_j^i)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq l}}$ mit

$$d_j^i = \sum_{k=1}^m b_k^i a_j^k \text{ gegeben ist.}$$

Nach Voraussetzung gelten

$$f(\alpha_j) = \sum_{i=1}^m a_j^i \beta_i \quad \text{und} \quad g(\beta_i) = \sum_{k=1}^n b_i^k \gamma_k.$$

Somit ist

$$g \circ f(\alpha_j) = \sum_{i=1}^m \sum_{k=1}^n a_j^i b_i^k \gamma_k$$

und die Behauptung folgt. \square

Bemerkung 4.2.4. Da die Komposition von Abbildungen assoziativ ist, ist auch die Matrizenmultiplikation assoziativ.

4.3. Rang einer linearen Abbildung.

Definition 4.3.1. Sei $f: V \rightarrow W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen. Dann ist der Rang von f , $\text{rang } f$, durch

$$\text{rang } f := \dim \text{im } f$$

definiert.

Theorem 4.3.2. Sei $f: V \rightarrow W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen. Zu fixierten Basen $S = \{a_i\}$ von V und $T = \{b_j\}$ von W sei A die zu f gehörige Matrix. Dann gilt

$$\text{rang } f = \text{rang } A.$$

Beweis. Der Raum $\text{im } f$ ist durch die Vektoren $f(a_i)$ erzeugt. Bezüglich der Basis T sind diese Vektoren aber gerade die Spalten von A . Daher ist $\text{rang } f$ gleich dem Spaltenrang von A . Die Behauptung folgt. \square

Korollar 4.3.3. Seien A und B zwei Matrizen, die bezüglich verschiedener Basen zu einer linearen Abbildung $f: V \rightarrow W$ gehören. Dann gilt $\text{rang } A = \text{rang } B$.

Theorem 4.3.4. Sei $f: V \rightarrow W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen. Dann gilt

$$\text{rang } f + \dim \ker f = \dim V.$$

Beweis. Mit Definition 4.3.1 folgt dies aus der Dimensionsformel, Theorem 4.1.10. \square

Bemerkung 4.3.5. Wir können die Matrix $A = (a_j^i)$ eines linearen Gleichungssystems

$$\sum_{k=1}^n a_k^i x^k = b^i, \quad 1 \leq i \leq m,$$

als Matrix einer linearen Abbildung $f: F^n \rightarrow F^m$ bezüglich der Standardbasen betrachten. Dann erhalten wir:

- (i) Das homogene lineare Gleichungssystem hat genau $\ker f$ als Lösungsraum und die Dimension des Lösungsraumes ist nach Theorem 4.3.4 gleich $n - \text{rang } A$.
- (ii) Das lineare Gleichungssystem ist genau dann lösbar, wenn $(b^1, \dots, b^m) \in \text{im } f$. $\text{im } f$ und der Spaltenraum von A stimmen überein.
- (iii) Im Falle $\ker f = \{0\}$ hat das lineare Gleichungssystem höchstens eine Lösung. Nach Theorem 4.3.4 gilt $\ker f = \{0\}$ genau dann, wenn $\text{rang } A = n$ gilt.

Theorem 4.3.6. Sei $f: V \rightarrow W$ eine lineare Abbildung. Sei $\dim V = \dim W = n$. Dann sind die folgenden Aussagen äquivalent:

- (i) f ist ein Isomorphismus,
- (ii) f ist injektiv,

(iii) f ist surjektiv.

Beweis. Dies folgt aus Theorem 4.1.6 und der Dimensionsformel

$$\dim \ker f + \dim \operatorname{im} f = \dim V. \quad \square$$

Somit korrespondieren Isomorphismen genau zu regulären Matrizen.

4.4. Basiswechsel. Sei $f: V \rightarrow W$ eine lineare Abbildung. Sei

$$\{a_1, \dots, a_p, a_{p+1}, \dots, a_n\}$$

eine Basis von V , so dass $\{a_{p+1}, \dots, a_n\}$ eine Basis von $\ker f$ ist. Setze $f(a_i) =: b_i$ für $1 \leq i \leq p$. Dann ist die Familie $\{b_1, \dots, b_p\}$ linear unabhängig. Ergänzen wir sie zu einer Basis $\{b_1, \dots, b_p, b_{p+1}, \dots, b_m\}$ so hat die zu f gehörige Matrix die Form

$$(4.2) \quad \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Diese Matrix hat p Einsen, m Zeilen und n Spalten.

Komplizierter wird es, wenn man eine Basis finden will, so dass $f: V \rightarrow V$ eine einfache Gestalt hat.

Bemerkung 4.4.1 (Transformationsregeln). Sei $f: V \rightarrow W$ linear. Wir verwenden die folgenden Bezeichnungen

	V	W
Alte Basen	$S = \{a_1, \dots, a_n\}$	$T = \{b_1, \dots, b_m\}$
Alte Koordinaten	x^1, \dots, x^n	y^1, \dots, y^m
Transformationsmatrizen	$D \in F^{n \times n}$	$E \in F^{m \times m}$
Neue Basen	$S' = \{a'_1, \dots, a'_n\}$	$T' = \{b'_1, \dots, b'_m\}$
Neue Koordinaten	$(x')^1, \dots, (x')^n$	$(y')^1, \dots, (y')^m$

Nach Wahl von D und E gelten

$$x' = Dx \quad \text{und} \quad y' = Ey.$$

Sei f bezüglich S, T durch A und bezüglich S', T' durch A' dargestellt. Es gelten daher

$$y = Ax \quad \text{und} \quad y' = A'x'.$$

Daraus erhalten wir $Ey = A'Dx$ und daraus $y = E^{-1}A'Dx$ sowie durch Vergleich mit $y = Ax$ die Relation

$$A = E^{-1}A'D \quad \text{sowie} \quad A' = EAD^{-1}.$$

Korollar 4.4.2. Seien S und S' zwei geordnete Basen von V . Sei D die Transformationsmatrix, die die Vektoren von S als die Linearkombinationen von Vektoren aus S' darstellt. Sei $f: V \rightarrow V$ bezüglich S durch A und bezüglich S' durch A' dargestellt. Dann gilt

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

Bei Abbildungen $f: V \rightarrow V$ wollen wir stets dieselbe Basis für den Definitionsbereich und Zielbereich verwenden.

Kombinieren wir nun die spezielle Darstellung einer linearen Abbildung aus (4.2) mit den Transformationsregeln, so erhalten wir

Theorem 4.4.3. Sei A eine beliebige $(m \times n)$ -Matrix. Dann gibt es reguläre Matrizen $D \in F^{n \times n}$ und $E \in F^{m \times m}$, so dass die Matrix EAD^{-1} von der Form (4.2) ist.

4.5. Der Vektorraum $\text{Hom}(V, W)$. Seien V, W Vektorräume über F . Seien die Abbildungen $f, g: V \rightarrow W$ linear. Seien $\lambda, \mu \in F$. Definiere

$$(\lambda f + \mu g)(a) := \lambda f(a) + \mu g(a).$$

(Man rechnet nach, dass auch $\lambda f + \mu g$ wieder linear ist.) Wir erhalten

Theorem 4.5.1. Seien V, W Vektorräume über F . Die Menge der linearen Abbildungen $V \rightarrow W$ mit der (wie oben) punktweise erklärten Addition und Skalarmultiplikation ist ein F -Vektorraum: $\text{Hom}(V, W)$.

Beweis. Übung. □

Theorem 4.5.2. Seien V und W Vektorräume der Dimension n bzw. m über F . Die Zuordnung

$$\text{Hom}(V, W) \ni f \mapsto A \in F^{m \times n}$$

aus Theorem 4.2.1 ist ein Isomorphismus.

Beweis. Die Linearität ist klar. Die Zuordnung und ihre Umkehrung sind vor und in Theorem 4.2.1 beschrieben. □

Korollar 4.5.3. Seien V, W endlichdimensionale Vektorräume mit Basen (a_i) bzw. (b_i) . Dann bilden die durch $f_i^j(a_k) := b_i \delta_k^j$ definierten Abbildungen eine Basis von $\text{Hom}(V, W)$. Insbesondere ist $\dim \text{Hom}(V, W) = \dim V \cdot \dim W$.

Die Aussage über die Basen gilt auch für beliebige Vektorräume.

Beweis. Diese Abbildungen entsprechen gerade den Matrizen, die in Zeile i und Spalte j eine 1 und sonst Nullen haben. □

Definition 4.5.4. Abbildungen $f \in \text{Hom}(V, F)$ heißen (lineare) Funktionale auf V .

Wir schreiben $V^* := \text{Hom}(V, F)$. V^* heißt der zu V duale Raum.

Bemerkung 4.5.5.

- (i) Nach Wahl einer Basis (a_1, \dots, a_n) von V und bezüglich der Basis 1 von F hat ein lineares Funktional die Gestalt

$$f\left(\sum_{i=1}^n x^i a_i\right) = \sum_{i=1}^n x^i f(a_i).$$

Wir bezeichnen die Abbildung $F^n \ni (x^1, \dots, x^n) \mapsto \sum_{i=1}^n x^i f(a_i)$ auch als Linearform.

- (ii) Für $0 \neq f \in V^* = \text{Hom}(V, F)$ mit $\dim V = n$ gilt $\dim \ker f = n - 1$, da $\dim \text{im } f = 1$ und $\dim \text{im } f + \dim \ker f = n$ gelten.
 (iii) Habe V die Basis (a_1, \dots, a_n) . Dann bilden die Vektoren $f^i \in V^*$ mit $f^i(a_k) = \delta_k^i$ eine Basis von V^* . Wir schreiben auch $(a^*)^i = f^i$. Für $b = \sum_{i=1}^n x^i a_i$ gilt

$$(a^*)^j(b) = \sum_{i=1}^n x^i (a^*)^j(a_i) = \sum_{i=1}^n x^i \delta_i^j = x^j.$$

Definition 4.5.6 (Duale Abbildung). Sei $f: V \rightarrow W$ linear. Dann definieren wir die duale Abbildung $f^*: W^* \rightarrow V^*$ zu f durch $f^*(\varphi) := \psi$ mit $\psi(\xi) = \varphi(f(\xi))$ für $\varphi \in W^*$ und $\xi \in V$.

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ & \searrow f^*(\varphi) & \downarrow \varphi \\ & & F. \end{array}$$

(Es folgt aus der Definition, dass $f^*(\varphi) \in V^*$ gilt und dass f^* linear ist.)

Lemma 4.5.7. Sei $f: V \rightarrow W$ durch die $(m \times n)$ -Matrix $A = (a_j^i)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ bezüglich Basen $(a_i)_{1 \leq i \leq n}$ und $(b_j)_{1 \leq j \leq m}$ von V bzw. W dargestellt. Dann ist $f^*: W^* \rightarrow V^*$ durch die $(m \times n)$ -Matrix $(b_j^i)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ mit $b_j^i = a_j^i$ bezüglich der Basen $((b^*)^j)_{1 \leq j \leq m}$ und $((a^*)^i)_{1 \leq i \leq n}$ in dem Sinne dargestellt, dass $f^*((b^*)^j) = \sum_{i=1}^n b_j^i (a^*)^i$ für alle $1 \leq j \leq m$ gilt.

Für die Koordinaten erhalten wir die Abbildung

$$(\xi_1, \dots, \xi_m) \mapsto \left(\sum_{j=1}^m \xi_j b_j^i \right)_{1 \leq i \leq n}.$$

Beweis. Nach Definition gilt

$$\begin{aligned} \left(f^*((b^*)^j) \right) (a_k) &= ((b^*)^j \circ f) (a_k) = (b^*)^j \left(\sum_{i=1}^m a_k^i b_i \right) \\ &= \sum_{i=1}^m a_k^i (b^*)^j b_i = \sum_{i=1}^m a_k^i \delta_i^j = a_k^j. \end{aligned}$$

Aus $f^*((b^*)^j) = \sum_{l=1}^n b_l^j (a^*)^l$ folgt andererseits

$$\left(f^*((b^*)^j) \right) (a_k) = \sum_{l=1}^n b_l^j (a^*)^l (a_k) = \sum_{l=1}^n b_l^j \delta_k^l = b_k^j.$$

Die Behauptung für die Koordinaten ergibt sich direkt aus der Linearität. \square

Bemerkung 4.5.8. Achtung, in der Darstellung als Matrix haben wir soeben über andere Indices als für Abbildungen $f: V \rightarrow W$ summiert. Die Koordinaten verändern sich folglich nach der Regel

$$(\xi_1, \dots, \xi_m) \mapsto (\xi_1, \dots, \xi_m) \begin{pmatrix} b_1^1 & \dots & b_n^1 \\ \vdots & & \vdots \\ b_1^m & \dots & b_n^m \end{pmatrix}.$$

Beachte dazu, dass wir die Basisvektoren von V^* anders als die von V oben indiziert haben. Die Koordinaten haben wir folglich unten indiziert. Somit handelt es sich um Zeilenvektoren.

In der Literatur weit verbreitet ist aber die (aus Kovarianzgründen nicht so saubere) Variante, auch diese Koordinaten als Spaltenvektoren zu schreiben, Indices generell unten anzubringen und den oberen Index einer Matrix an die erste Stelle zu senken, also a_{ij} statt a_j^i zu verwenden. Dann gilt für die Komponenten unter f^*

die Abbildungsregel

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix} \mapsto \begin{pmatrix} b_1^1 & \dots & b_1^m \\ \vdots & & \vdots \\ b_n^1 & \dots & b_n^m \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix}.$$

Wie man sich leicht überzeugen kann, sind die Einträge in so erhaltenen Spaltenvektor dieselben wie beim Zeilenvektor. Hier ist nun die Transformationsmatrix die Matrix A^T , die transponierte Matrix, die für $A = (a_j^i)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ durch $A^T := (b_i^j)_{\substack{1 \leq j \leq n \\ 1 \leq i \leq m}}$ mit

$b_i^j := a_j^i$ definiert ist. Graphisch erhält man

$$A^T = \begin{pmatrix} a_1^1 & a_1^2 & a_1^3 & \dots & a_1^m \\ a_2^1 & a_2^2 & a_2^3 & \dots & a_2^m \\ a_3^1 & a_3^2 & a_3^3 & \dots & a_3^m \\ \vdots & \vdots & \vdots & & \vdots \\ a_n^1 & a_n^2 & a_n^3 & \dots & a_n^m \end{pmatrix} \quad \text{aus} \quad A = \begin{pmatrix} a_1^1 & a_2^1 & a_3^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & a_3^2 & \dots & a_n^2 \\ a_1^3 & a_2^3 & a_3^3 & \dots & a_n^3 \\ \vdots & \vdots & \vdots & & \vdots \\ a_1^m & a_2^m & a_3^m & \dots & a_n^m \end{pmatrix}.$$

4.6. Endomorphismen.

Definition 4.6.1 (Ring). Ein Ring R ist eine additive abelsche Gruppe, d. h. wir schreiben die Verknüpfung als „+“, mit einer Multiplikation „·“, so dass folgende Axiome für alle $a, b, c \in R$ gelten

- (R1) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (Assoziativität)
 (R2) $a \cdot (b + c) = a \cdot b + a \cdot c$ (Distributivität)
 (R3) $(a + b) \cdot c = a \cdot c + b \cdot c$ (Distributivität)

Hierbei verwenden wir eine Klammersetzung nach der Regel „Punkt vor Strich“.

- (R4) Ein Ring mit Einselement, also einem Element $1 = 1_R \in R$, das $a \cdot 1 = a = 1 \cdot a$ für alle $a \in R$ erfüllt, heißt Ring mit 1. Häufig wird die Existenz eines Einselementes bereits in der Definition eines Ringes verlangt.
 (R5) Ein Ring heißt kommutativ, falls $a \cdot b = b \cdot a$ für alle $a, b \in R$ gilt.

Beispiele 4.6.2.

- (i) \mathbb{Z} ist ein kommutativer Ring mit Eins.
 (ii) Die geraden Zahlen sind ein Ring ohne Eins.
 (iii) Ein Körper ist ein kommutativer Ring mit Eins.
 (iv) $F^{n \times n}$ ist ein Ring. Für $n \geq 2$ ist er nicht kommutativ.

Definition 4.6.3 (Ringhomomorphismus). Seien R, S Ringe. Dann heißt eine Abbildung $\Phi: R \rightarrow S$ Ringhomomorphismus, wenn $\Phi(a + b) = \Phi(a) + \Phi(b)$ und $\Phi(a \cdot b) = \Phi(a) \cdot \Phi(b)$ für alle $a, b \in R$ gelten. Bei Ringen mit Eins verlangt man zusätzlich $\Phi(1) = 1$.

Der Deutlichkeit halber könnte man auch $\Phi(a +_R b) = \Phi(a) +_S \Phi(b)$ oder $\Phi(1_R) = 1_S$ schreiben.

Definition 4.6.4 (Algebra). Eine Algebra A über einem Körper F ist ein Vektorraum mit einer Multiplikation „·“, so dass folgende Axiome für alle $a, b, c \in A$ und $\lambda \in F$ erfüllt sind:

- (A1) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (Assoziativität)
 (A2) $a \cdot (b + c) = a \cdot b + a \cdot c$ (Distributivität)
 (A3) $(a + b) \cdot c = a \cdot c + b \cdot c$ (Distributivität)
 (A4) $(\lambda a) \cdot b = \lambda(a \cdot b) = a \cdot (\lambda b)$
 (A5) Es existiert ein Einselement $1 = 1_A$ mit $a \cdot 1 = 1 \cdot a = a$ für alle $a \in A$.

Beispiele 4.6.5.

- (i) $F^{n \times n}$ mit der Matrixmultiplikation ist eine Algebra.

- (ii) Sei V ein F -Vektorraum. Dann ist $\text{Hom}(V, V)$ eine Algebra. Lineare Selbstabbildungen eines Vektorraumes heißen auch Endomorphismen, $\text{End}(V)$.
- (iii) $F[X]$ ist eine Algebra über F und ein Ring, der Polynomring in der Variablen X über F .
- (iv) \mathbb{C} ist eine Algebra über \mathbb{R} .

Bemerkung 4.6.6. Sei V ein Vektorraum. Nach Wahl einer Basis ist jedem $f \in \text{Hom}(V, V)$ eine Matrix A zugeordnet. Nach Theorem 4.5.2 ist diese Abbildung

$$\Phi: \text{Hom}(V, V) \rightarrow F^{n \times n}$$

ein Isomorphismus von Vektorräumen. Dann gelten

$$\Phi(f \circ g) = \Phi(f) \cdot \Phi(g), \quad \Phi(f + g) = \Phi(f) + \Phi(g) \quad \text{sowie} \quad \Phi(\text{id}) = \mathbf{1},$$

siehe insbesondere Theorem 4.2.3. Entsprechendes gilt für die Umkehrabbildung von Φ . Wir sagen daher, dass die Algebren $\text{Hom}(V, V)$ und $F^{n \times n}$ isomorph sind.

Korollar 4.4.2 motiviert die folgende

Definition 4.6.7. Seien $A, A' \in F^{n \times n}$. Dann heißen A und A' ähnlich, wenn es eine reguläre Matrix D mit $A' = DAD^{-1}$ gibt.

Bemerkung 4.6.8. Äquivalent dazu hätte man auch zwei Matrizen ähnlich nennen können, wenn es eine lineare Abbildung und zugehörige Basen gibt, so dass diese Abbildung bezüglich der Basen durch diese Matrizen beschrieben wird.

Ziel: Im weiteren Verlauf der Vorlesung wollen wir uns intensiver mit der Frage beschäftigen, wie wir es erreichen können, dass die einer Abbildung $f: V \rightarrow V$ vermöge einer Basis zugeordnete Matrix möglichst einfach aussieht.

Als Vorbereitung definieren wir dazu

Definition 4.6.9. Ein Unterraum U von V heißt bezüglich einer linearen Abbildung $f: V \rightarrow V$ invariant, falls $f(U) \subset U$ gilt.

Beispiele 4.6.10.

- (i) Betrachte $F[X]$ mit der linearen Abbildung $f(p(X)) := p'(X)$, der Ableitung von p ,

$$f\left(\sum_{i=0}^n a_i X^i\right) := \sum_{i=1}^n i a_i X^{i-1}.$$

Die Unterräume der Polynome vom Grad $\leq m$ sind beim Ableiten invariant. $\{p(X) : p(0) = 0\}$ ist unter f nicht invariant.

- (ii) Sei $f: V \rightarrow V$ eine beliebige lineare Abbildung. Dann sind $\ker f$ und im f invariante Unterräume.

Bemerkung 4.6.11.

- (i) Sei U unter $f: V \rightarrow V$ invariant, $\dim V < \infty$. Sei $\{a_1, \dots, a_p\}$ eine Basis von U . Ergänze diese zu einer Basis $\{a_1, \dots, a_p, a_{p+1}, \dots, a_n\}$ von V . Dann ist f bezüglich dieser Basis durch

$$\begin{pmatrix} a_1^1 & \dots & a_p^1 & a_{p+1}^1 & \dots & a_n^1 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ a_1^p & \dots & a_p^p & a_{p+1}^p & \dots & a_n^p \\ 0 & \dots & 0 & a_{p+1}^{p+1} & \dots & a_n^{p+1} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & a_{p+1}^n & \dots & a_n^n \end{pmatrix} \equiv \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

dargestellt, wobei A für einen $p \times p$ -Block, B für einen $p \times (n-p)$ -Block, 0 für einen $(n-p) \times p$ -Block und C für einen $(n-p) \times (n-p)$ -Block steht; hierbei

ist ein Block eine rechteckige Anordnung von Zahlen wie aus der Gleichung ersichtlich.

- (ii) Kann man gegebenenfalls auch noch $B = 0$ erreichen? Gelte $V = U_1 \oplus \dots \oplus U_k$ für f -invariante Unterräume U_i . Dann hat die f darstellende Matrix bezüglich einer Basis, die aus den Basen der Unterräume U_i (in geeigneter Anordnung) besteht, die Blockgestalt

$$\begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & A_k \end{pmatrix}.$$

Dabei ist A_i die Matrix, die $f|_{U_i}$ bezüglich einer Basis von U_i darstellt, $1 \leq i \leq k$.

Da eine solche Darstellung besonders einfach wird, wenn $\dim U_i = 1$ ist, definiert man Eigenvektoren und Eigenräume:

4.7. Eigenvektoren und Eigenräume.

Definition 4.7.1. Ein Vektor $a \in V$ mit $a \neq 0$ heißt Eigenvektor von $f: V \rightarrow V$ zum Eigenwert $\lambda \in F$, falls

$$f(a) = \lambda a$$

gilt. Der Unterraum

$$E_\lambda := \{b \in V : f(b) = \lambda b\}$$

heißt der zu λ gehörige Eigenraum. Die Dimension $\dim E_\lambda$ heißt geometrische Vielfachheit von λ .

Ebenso spricht man bei einer Matrix $A \in F^{n \times n}$ vom Eigenvektor x zum Eigenwert λ , falls

$$Ax = \lambda x$$

gilt. Eigenräume und geometrische Vielfachheit definiert man entsprechend.

Ein $\lambda \in F$ heißt Eigenwert von f bzw. A , wenn es einen zugehörigen Eigenvektor gibt.

Bemerkung 4.7.2.

- (i) λ ist genau dann ein Eigenwert von f , wenn λ ein Eigenwert der f zugeordneten Matrix A ist.
- (ii) $\ker f$ ist der Eigenraum zum Eigenwert 0. Ebenso gilt $\ker(f - \lambda \mathbb{1}) = E_\lambda$.
- (iii) λ ist genau dann ein Eigenwert von f , wenn $f - \lambda \mathbb{1}$ singulär ist.
- (iv) Die Eigenräume E_λ von f sind unter f invariant.
- (v) Für $a \in E_\lambda$ gilt $f(a) = \lambda a$. Daher ist $f|_{E_\lambda}: E_\lambda \rightarrow E_\lambda$ bezüglich einer beliebigen Basis von der Form

$$\begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda \end{pmatrix}.$$

Theorem 4.7.3. Seien λ_i , $1 \leq i \leq m$, paarweise verschiedene Eigenwerte von f mit zugehörigen Eigenvektoren a_i , so sind die Vektoren a_i linear unabhängig.

Beweis. Falls nicht, so gibt es eine nichttriviale Darstellung der Null. Wähle eine kürzeste nichttriviale Darstellung der Null, also eine solche, die möglichst wenige der

Vektoren a_i verwendet. Ohne Einschränkung verwendet die kürzeste Darstellung a_1, \dots, a_n . Es gilt

$$0 = \sum_{i=1}^n \mu^i a_i$$

für geeignete $\mu^i \in F$. Wir erhalten einerseits

$$0 = f\left(\sum_{i=1}^n \mu^i a_i\right) = \sum_{i=1}^n \mu^i f(a_i) = \sum_{i=1}^n \mu^i \lambda_i a_i$$

und andererseits

$$0 = \lambda_1 \sum_{i=1}^n \mu^i a_i = \sum_{i=1}^n \lambda_1 \mu^i a_i.$$

Als Differenz erhalten wir

$$0 = \sum_{i=2}^n (\lambda_i - \lambda_1) \mu^i a_i.$$

Da dies eine kürzere Darstellung der Null als die ursprüngliche ist, muss sie trivial sein. Also gilt $(\lambda_i - \lambda_1) \mu^i a_i = 0$ für alle $2 \leq i \leq n$. Da die Eigenwerte paarweise verschieden sind, folgt bereits $\mu^i a_i = 0$ und somit auch $\mu^i = 0$ für $2 \leq i \leq n$. Also ist auch $\mu^1 a_1 = 0$. Wegen $a_1 \neq 0$, da es sich um einen Eigenvektor handelt, ist $\mu^1 = 0$, die Linearkombination der Null also doch trivial. Widerspruch. \square

Korollar 4.7.4. *Sei V ein n -dimensionaler Vektorraum. Sei $f: V \rightarrow V$ ein Endomorphismus. Dann besitzt f höchstens n verschiedene Eigenwerte. Besitzt f genau n verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n$, so bildet eine Auswahl zugehöriger Eigenvektoren a_1, \dots, a_n eine Basis von V . Bezüglich dieser Basis hat f die Gestalt*

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Bemerkung 4.7.5.

- (i) Eine Matrix eines Endomorphismusses heißt diagonalisierbar, wenn es eine Basis gibt, so dass $A = (a_j^i)_{1 \leq i, j \leq n}$ diagonal ist, also $a_j^i = 0$ für $i \neq j$ erfüllt. Eine solche Matrix ist genau dann diagonalisierbar, wenn es eine Basis aus Eigenvektoren gibt.
- (ii) Die Matrix $\begin{pmatrix} 1 & \\ 0 & 1 \end{pmatrix}$ ist nicht diagonalisierbar, da $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ (bis auf skalare Vielfache) der einzige Eigenvektor ist.

5.1. Polynome.

Bemerkung 5.1.1. Seien $p, q \in F[X]$. Dann gelten

- (i) $\deg(p + q) \leq \max\{\deg p, \deg q\}$
- (ii) $\deg(p \cdot q) = \deg p + \deg q$

$a \in F$ heißt Nullstelle von p , wenn $p(a) = 0$ gilt, wobei $p(a)$ der Ausdruck ist, den wir erhalten, wenn wir in $p(X)$ überall die Variable X durch a ersetzen.

Theorem 5.1.2 (Polynomdivision). *Seien $p, q \in F[X]$, $\deg q > 0$. Dann gibt es eindeutig bestimmte Polynome $s, r \in F[X]$ mit*

$$p(X) = s(X)q(X) + r(X)$$

und $\deg r < \deg q$.

Beweis. Zur Eindeutigkeit: Seien s, r wie behauptet und $s', r' \in F[X]$, so dass

$$p(X) = s'(X)q(X) + r'(X)$$

und $\deg r' < \deg q$ gelten. Wir erhalten

$$0 = (s(X) - s'(X))q(X) + (r(X) - r'(X)).$$

Ist $s \neq s'$, so ist $s(X) - s'(X)$ ein Polynom vom Grad ≥ 0 . Also ist $0 + \deg q(X) \leq \deg((s(X) - s'(X))q(X)) = \deg(r'(X) - r(X)) < \deg q(X)$. Widerspruch.

Zur Existenz: Wir beweisen die Aussage per Induktion nach $\deg p$. Für $\deg p < \deg q$ können wir $s = 0$ und $r = p$ wählen. Sei also $\deg p \geq \deg q$. Nehme an, dass

$$p(X) = a_n X^n + \dots + a_0 \quad \text{mit } a_n \neq 0$$

und

$$q(X) = b_m X^m + \dots + b_0 \quad \text{mit } b_m \neq 0 \text{ und } m \leq n$$

gelten. Es folgt

$$\begin{aligned} a_n X^n + \dots + a_0 &= \frac{a_n}{b_m} X^{n-m} (b_m X^m + \dots + b_0) + a_n X^n + a_{n-1} X^{n-1} + \dots + a_0 \\ &\quad - a_n X^n - \frac{a_n b_{m-1}}{b_m} X^{n-1} - \frac{a_n b_{m-2}}{b_m} X^{n-2} - \dots - \frac{a_n b_0}{b_m} X^{n-m} \\ &= \frac{a_n}{b_m} X^{n-m} q(X) + a_{n-1} X^{n-1} + \dots + a_0 \\ &\quad - \frac{a_n b_{m-1}}{b_m} X^{n-1} - \frac{a_n b_{m-2}}{b_m} X^{n-2} - \dots - \frac{a_n b_0}{b_m} X^{n-m} \\ &= \frac{a_n}{b_m} X^{n-m} q(X) + R(X), \end{aligned}$$

wobei $R(X)$ ein Polynom vom Grad $\leq (n-1)$ ist. Nach Induktionsvoraussetzung existieren daher $A, B \in F[X]$ mit $\deg B < \deg q$, so dass

$$R(X) = A(X)q(X) + B(X)$$

gilt. Es folgt

$$p(X) = \left(\frac{a_n}{b_m} X^{n-m} + A(X) \right) q(X) + B(X)$$

wie behauptet. □

Korollar 5.1.3. *Sei $a \in F$ Nullstelle eines Polynomes $p \in F[X]$. Dann gibt es ein eindeutig bestimmtes Polynom $q(X)$, so dass*

$$p(X) = q(X)(X - a)$$

mit $\deg q = (\deg p) - 1$ gilt.

Beweis. Nach Bemerkung 5.1.1 ist die Aussage über den Grad klar.

Theorem 5.1.2 liefert eine Darstellung wie gefordert, $p(X) = q(X)(X - a) + r(X)$, aber möglicherweise mit einem zusätzlichen Polynom $r(X)$, das $\deg r < 1$ erfüllt und somit konstant ist. Auswerten an der Stelle $X = a$ liefert $r = 0$. □

Per Induktion erhält man hieraus

Korollar 5.1.4. Ein Polynom $p(X)$ vom Grad $n \geq 1$ besitzt höchstens n Nullstellen.

Definition 5.1.5. Sei $a \in F$ eine Nullstelle des Polynomes $p \in F[X]$. Dann besitzt p nach Korollar 5.1.3 (gegebenenfalls mehrfach angewandt) eine Darstellung der Form

$$p(X) = (X - a)^k q(X)$$

mit $k \in \mathbb{N}$ und $q \in F[X]$, so dass $q(a) \neq 0$ gilt. k heißt dann die Vielfachheit der Nullstelle a .

Wie das Beispiel $X^2 + 1 \in \mathbb{R}[X]$ zeigt, braucht ein Polynom keine Nullstelle zu besitzen. Es gilt aber

Theorem 5.1.6 (Fundamentalsatz der Algebra). Sei $p \in \mathbb{C}[X]$ mit $\deg p \geq 1$. Dann besitzt p eine Nullstelle.

Beweis. Siehe Vorlesung „Funktionentheorie“. □

Korollar 5.1.7. Sei $p \in \mathbb{C}[X]$ ein Polynom mit $\deg p = n > 0$. Dann gibt es komplexe Zahlen $d, \lambda_1, \dots, \lambda_n$, so dass

$$p(X) = d(X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n)$$

ist. λ_i sind gerade die Nullstellen von p . Sie kommen entsprechend ihrer Vielfachheit ggf. mehrfach vor.

Beweis. Induktion, Übung. □

Lemma 5.1.8. Sei $p \in \mathbb{C}[X]$ ein Polynom mit reellen Koeffizienten. Sei $a + ib$, $a, b \in \mathbb{R}$, eine Nullstelle von p , so ist auch die komplex konjugierte Zahl $\overline{a + ib} := a - ib$ eine Nullstelle von p .

Beweis. Sei $p(x) = a_n x^n + \dots + a_1 x + a_0$ mit $a_i \in \mathbb{R}$. Setze $c := a + ib$. Es gilt

$$0 = \overline{p(c)} = \overline{a_n c^n + \dots + a_1 c + a_0} = a_n \bar{c}^n + \dots + a_1 \bar{c} + a_0 = p(\bar{c}).$$

Somit ist auch \bar{c} eine Nullstelle von p . □

5.2. Permutationen.

Definition 5.2.1. Eine Permutation der Zahlen $1, \dots, n$ ist eine bijektive Abbildung $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$.

Schematisch können wir eine Permutation durch

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

oder kürzer durch

$$(\sigma(1) \ \sigma(2) \ \sigma(3) \ \dots \ \sigma(n))$$

darstellen, im zweiten Fall auch mit Kommata getrennt. (Aufgrund der Bijektivität von σ kommt jede der Zahlen $1, 2, \dots, n$ in jeder dieser Zeilen genau einmal vor.)

Bemerkung 5.2.2.

- (i) Es gibt genau $n! := 1 \cdot 2 \cdots n$ Permutationen von $1, \dots, n$.
- (ii) Die Permutationen mit $(\sigma \circ \tau)(i) := \sigma(\tau(i))$ bilden eine Gruppe: S_n .
- (iii) S_n ist genau für $n = 1, 2$ abelsch.

Beweis. Übung. □

Definition 5.2.3. Sei σ eine Permutation der Zahlen $1, 2, \dots, n$. Sei $s = s(\sigma)$ die Anzahl der Paare (i, j) mit $1 \leq i < j \leq n$ und $\sigma(i) > \sigma(j)$. Dann definieren wir das Signum der Permutation σ durch

$$\text{sign } \sigma := (-1)^{s(\sigma)}.$$

Eine Permutation σ mit $\text{sign}(\sigma) = 1$ heißt gerade, eine Permutation σ mit $\text{sign}(\sigma) = -1$ heißt ungerade.

Definition 5.2.4 (Transposition). Setze für $1 \leq i, j \leq n$ mit $i \neq j$

$$\tau_{ij}(k) := \begin{cases} j & k = i, \\ i & k = j, \\ k & \text{sonst.} \end{cases}$$

Die Permutation vertauscht gerade die Zahlen i und j . Es gilt $\tau^{-1} = \tau$.

Theorem 5.2.5. Sei σ eine Permutation der Zahlen $1, 2, \dots, n$. Sei τ eine Transposition. Dann gilt

$$\text{sign}(\tau \circ \sigma) = -\text{sign}(\sigma).$$

Beweis. Seien σ und σ' zwei Permutationen,

$$(\sigma(1), \sigma(2), \dots, \sigma(n)) \quad \text{und} \quad (\sigma'(1), \sigma'(2), \dots, \sigma'(n)),$$

die sich nur dadurch unterscheiden, dass in dieser Darstellung zwei benachbarte Einträge vertauscht wurden. Dann gilt nach Definition

$$\text{sign } \sigma = -\text{sign } \sigma'.$$

Sei $(\tau\sigma(1), \dots, \tau\sigma(n))$ aus $(\sigma(1), \dots, \sigma(n))$ durch Vertauschen der Zahlen i und j entstanden, die k weitere Zahlen in dieser Darstellung zwischen sich haben. Durch $2k + 1$ Vertauschungen nebeneinanderliegender Elemente kann man diese beiden Permutationen ineinander überführen. Daher gilt

$$\text{sign}(\tau \circ \sigma) = (-1)^{2k+1} \text{sign}(\sigma) = -\text{sign}(\sigma). \quad \square$$

Theorem 5.2.6. Jede Permutation der Zahlen $1, 2, \dots, n$ lässt sich als Produkt von Transpositionen schreiben.

Beweis. Für die Identität und eine beliebige Permutation τ gilt $\tau \circ \tau = \mathbf{1}$. Sonst kann man durch Permutationen sukzessive erreichen, dass $(1, 2, \dots, n)$ in eine Permutation überführt wird, bei der in dieser Darstellung das Anfangsstück, das mit $(\sigma(1), \sigma(2), \dots, \sigma(n))$ übereinstimmt um mindestens eine Position länger wird. Die Aussage folgt nun per Induktion. \square

Theorem 5.2.7. Sei σ eine Permutation, die sich als Produkt von r Transpositionen schreiben lässt. Dann gilt $\text{sign } \sigma = (-1)^r$.

Beweis. Benutze Theorem 5.2.5. \square

Da $\text{sign } \sigma$ unabhängig von Transpositionen definiert ist, gilt

Korollar 5.2.8. Lässt sich eine Permutation als Produkt von k und l Transpositionen schreiben, dann gilt

$$k \equiv l \pmod{2},$$

d. h. k und l sind entweder beide gerade oder beide ungerade.

Theorem 5.2.9. Sei $n \geq 2$. Dann gibt es genau $\frac{1}{2}n!$ gerade und $\frac{1}{2}n!$ ungerade Permutationen der Zahlen $1, 2, \dots, n$.

Beweis. Sei τ eine beliebige Transposition. Da die Permutationen eine Gruppe bilden, ist $\sigma \mapsto \tau\sigma$ bijektiv. Die geraden Permutationen werden dabei gerade auf die ungeraden und umgekehrt abgebildet. Folglich gibt es von jeder Sorte gleich viele. \square

Bemerkung 5.2.10. Eine Untergruppe der symmetrischen Gruppe S_n ist die alternierende Gruppe. Sie besteht aus den Elementen σ der symmetrischen Gruppe mit $\text{sign } \sigma = 1$.

Definition 5.2.11. Seien G, H Gruppen. Dann ist $\varphi: G \rightarrow H$ ein Gruppenhomomorphismus, wenn für alle $a, b \in G$

$$\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$$

gilt.

Beispiele 5.2.12.

- (i) Lineare Abbildungen sind Gruppenhomomorphismen, wenn wir einen Vektorraum als additive Gruppe betrachten.
- (ii) Die Koordinatenabbildung eines Vektorraumes ist ein Gruppenhomomorphismus bezüglich der additiven Gruppe eines Vektorraumes.
- (iii) Die Abbildung $\Phi: \text{Hom}(V, V) \rightarrow F^{n \times n}$ wie in Bemerkung 4.6.6 ist ein Gruppenhomomorphismus.

Ein weiteres Beispiel für einen Gruppenhomomorphismus ergibt sich aus

Theorem 5.2.13. Seien σ, τ zwei Permutationen. Dann gilt

$$\text{sign}(\sigma \circ \tau) = \text{sign}(\sigma) \cdot \text{sign}(\tau).$$

Beweis. Stelle die Permutationen als Verkettung von Transpositionen dar und zähle deren Anzahl. \square

Korollar 5.2.14. Sei σ eine Permutation. Dann gilt $\text{sign}(\sigma^{-1}) = \text{sign } \sigma$.

Beweis. Es gilt $\text{sign}(\sigma) \cdot \text{sign}(\sigma^{-1}) = \text{sign}(\sigma \circ \sigma^{-1}) = \text{sign } \mathbf{1} = 1$. \square

5.3. Determinanten. In diesem Kapitel werden wir häufiger zwischen einer Matrix $A = (a_j^i)_{1 \leq i, j \leq n}$ und deren Spalten $a_j = (a_j^i)_{1 \leq i \leq n} = \sum_{i=1}^n a_j^i e_i$ mit $e_i = (0, \dots, 0, 1, 0, \dots, 0)^T$ hin- und herwechseln. Ebenso werden wir zwischen $\det A$ und $\det(a_1, \dots, a_n)$ hin- und herwechseln.

Definition 5.3.1 (Determinante). Eine Funktion $\det: F^{n \times n} \rightarrow F$, die n Vektoren in F^n ein Element in F zuordnet, $\det(a_1, \dots, a_n)$, heißt n -dimensionale Determinantenfunktion, wenn sie folgende Axiome erfüllt

- (i) $\det(a_1, \dots, a_{i-1}, \lambda a_i, a_{i+1}, \dots, a_n) = \lambda \det(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n)$
für alle $\lambda \in F$ und $1 \leq i \leq n$
- (ii) $\det(a_1, \dots, a_{i-1}, a_i + b_i, a_{i+1}, \dots, a_n) = \det(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n) + \det(a_1, \dots, a_{i-1}, b_i, a_{i+1}, \dots, a_n)$ für alle $1 \leq i \leq n$ (Linearität)
- (iii) $\det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = 0$, falls $a_i = a_j$ für $i \neq j$
- (iv) $\det(e_1, \dots, e_n) = 1$ (Normierung)

Da eine Determinantenfunktion in jedem Argument linear ist, heißt sie multilinear.

Bemerkung 5.3.2. Die Determinante ist ein Polynom in den Einträgen. Wir werden später auch Matrizen statt Körperelementen einsetzen. Daher empfehlen wir, sich beim zweiten Lesen zu überlegen, dass sich die Resultate über Determinanten auf diesen Fall anpassen lassen.

Wir werden erst später zeigen, dass es eine eindeutig bestimmte Determinantenfunktion gibt.

Lemma 5.3.3. *Sei \det eine Determinantenfunktion. Dann gilt*

$$\begin{aligned}
 (i) \quad & \det\left(a_1, \dots, a_{i-1}, \sum_{k=1}^m \lambda^k b_k, a_{i+1}, \dots, a_n\right) \\
 &= \sum_{k=1}^m \lambda^k \det(a_1, \dots, a_{i-1}, b_k, a_{i+1}, \dots, a_n) \\
 (ii) \quad & \det\left(a_1, \dots, a_{i-1}, a_i + \sum_{\substack{k=1 \\ k \neq i}}^n \lambda^k a_k, a_{i+1}, \dots, a_n\right) \\
 &= \det(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n) \\
 (iii) \quad & \det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = -\det(a_1, \dots, a_j, \dots, a_i, \dots, a_n) \text{ für } i \neq j \\
 & \hspace{15em} (\text{Antisymmetrie})
 \end{aligned}$$

Beweis.

- (i) Dies folgt direkt aus der Multilinearität.
- (ii) Ist eine Folgerung aus der Multilinearität und aus (iii) in der Definition der Determinante.
- (iii) Es gilt

$$\begin{aligned}
 \det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) &= \det(a_1, \dots, a_i + a_j, \dots, a_j, \dots, a_n) \\
 &= \det(a_1, \dots, a_i + a_j, \dots, a_j - a_i - a_j, \dots, a_n) \\
 &= -\det(a_1, \dots, a_i + a_j, \dots, a_i, \dots, a_n) \\
 &= -\det(a_1, \dots, a_j, \dots, a_i, \dots, a_n). \quad \square
 \end{aligned}$$

Theorem 5.3.4. *Sei \det eine Determinantenfunktion. Ist die Familie der Vektoren $\{a_1, \dots, a_n\}$ linear abhängig, so gilt*

$$\det(a_1, \dots, a_n) = 0.$$

Beweis. Sei $a_i = \sum_{\substack{k=1 \\ k \neq i}}^n \lambda^k a_k$. Dann folgt

$$\det(a_1, \dots, a_n) = \sum_{\substack{k=1 \\ k \neq i}}^n \lambda^k \det(a_1, \dots, a_{i-1}, a_k, a_{i+1}, \dots, a_n) = 0,$$

da in den Determinanten auf der rechten Seite stets zwei Vektoren doppelt vorkommen. \square

Der Eindeigkeitssatz für die Determinantenfunktion lautet

Theorem 5.3.5. *Sei $n \geq 1$. Dann gibt es höchstens eine n -dimensionale Determinantenfunktion.*

Beweis. Seien d, D zwei Determinantenfunktionen. Dann gilt

$$1 = d(e_1, \dots, e_n) = D(e_1, \dots, e_n).$$

Sei (i_1, \dots, i_n) eine Permutation von $1, \dots, n$. Da wir eine Permutation als Verkettung von Transpositionen schreiben können erhalten wir aus der Antisymmetrie

$$\begin{aligned}
 d(e_{i_1}, \dots, e_{i_n}) &= \text{sign}(i_1, \dots, i_n) d(e_1, \dots, e_n) \\
 &= \text{sign}(i_1, \dots, i_n) D(e_1, \dots, e_n) \\
 &= D(e_{i_1}, \dots, e_{i_n}).
 \end{aligned}$$

Sei nun $a_j = \sum_{i=1}^n a_j^i e_i$. Wir erhalten

$$\begin{aligned} d(a_1, \dots, a_n) &= \sum_{i_1=1}^n a_1^{i_1} \sum_{i_2=1}^n a_2^{i_2} \dots \sum_{i_n=1}^n a_n^{i_n} d(e_{i_1}, \dots, e_{i_n}) \\ &= \sum_{i_1=1}^n a_1^{i_1} \sum_{i_2=1}^n a_2^{i_2} \dots \sum_{i_n=1}^n a_n^{i_n} D(e_{i_1}, \dots, e_{i_n}) \\ &= D(a_1, \dots, a_n). \end{aligned}$$

Dabei haben wir das mittlere Gleichheitszeichen für den Fall, dass (i_1, \dots, i_n) eine Permutation von $(1, \dots, n)$ ist, bereits nachgerechnet. Sonst stimmen jedoch mindestens zwei Indices überein und beide Terme verschwinden. Die Behauptung folgt. \square

Der Existenzsatz für die Determinantenfunktion ist

Theorem 5.3.6. *Für $n \geq 1$ existiert eine Determinantenfunktion. Sie ist durch*

$$\det(a_1, \dots, a_n) \equiv \det A := \sum_{\sigma \in S_n} \text{sign } \sigma \cdot a_1^{\sigma(1)} \cdot a_2^{\sigma(2)} \dots a_n^{\sigma(n)}$$

gegeben. (Leibnizformel)

Beweis. Bereits aus dem Beweis des Eindeutigkeitsatzes folgt, dass es höchstens diese Determinantenfunktion geben kann.

Bei der angegebenen Funktion handelt es sich auch um eine Determinantenfunktion, denn sie ist offensichtlich multilineare und normiert und wir erhalten für $a_j^i = a_k^i$ für alle i und $j \neq k$

$$\begin{aligned} \det A &= \det(a_1, \dots, a_j, \dots, \underbrace{a_k}_{=a_j}, \dots, a_n) \\ &= \sum_{\sigma \in S_n} \text{sign } \sigma \cdot a_1^{\sigma(1)} \dots a_j^{\sigma(j)} \dots a_k^{\sigma(k)} \dots a_n^{\sigma(n)} \\ &= \sum_{\substack{\sigma \in S_n \\ \sigma(j) < \sigma(k)}} \text{sign } \sigma \cdot a_1^{\sigma(1)} \dots a_j^{\sigma(j)} \dots a_j^{\sigma(k)} \dots a_n^{\sigma(n)} \\ &\quad + \sum_{\substack{\sigma \in S_n \\ \sigma(j) > \sigma(k)}} \text{sign } \sigma \cdot a_1^{\sigma(1)} \dots a_j^{\sigma(j)} \dots a_j^{\sigma(k)} \dots a_n^{\sigma(n)} = 0, \end{aligned}$$

denn für jeden Term in der obigen Summe gibt es einen Term in der unteren Summe, der zur gleichen Permutation gehört, nur dass $\sigma(j)$ und $\sigma(k)$ vertauscht sind. Da diese Permutationen sich also um genau eine Transposition unterscheiden, heben sich diese Terme gerade paarweise auf und die Behauptung folgt. \square

Korollar 5.3.7. *Für quadratische Matrizen $A \in F^{n \times n}$, $C \in F^{m \times m}$ und $B \in F^{n \times m}$ gilt*

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det A \cdot \det C.$$

Beweis. Sei $(a_j^i)_{1 \leq i, j \leq n+m} = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$. In der Leibnizformel

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \sum_{\sigma \in S_{n+m}} \text{sign } \sigma \cdot a_1^{\sigma(1)} \cdot a_2^{\sigma(2)} \dots a_{n+m}^{\sigma(n+m)}$$

verschwinden alle Terme, für die $\sigma(i) > n$ für mindestens ein i mit $1 \leq i \leq n$ gilt. Zu $\sigma \in S_{n+m}$ mit $\sigma(i) \leq n$ für $1 \leq i \leq n$ gibt es Permutationen $\tau \in S_n$ und $\rho \in S_m$,

so dass $\sigma(i) = \tau(i)$ für alle $1 \leq i \leq n$ und $\sigma(n+i) = n + \rho(i)$ für alle $1 \leq i \leq m$ gilt. Stellt man τ und ρ als Verkettungen von Transpositionen dar, so kann man diese auch auf $1, \dots, n$ bzw. $n+1, \dots, n+m$ wirken lassen und erhält $\text{sign } \sigma = \text{sign } \tau \cdot \text{sign } \rho$. Es folgt somit

$$\begin{aligned} \det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} &= \sum_{\tau \in S_n} \sum_{\rho \in S_m} \text{sign } \tau \cdot \text{sign } \rho \cdot a_1^{\tau(1)} \cdots a_n^{\tau(n)} \cdot a_{n+1}^{n+\rho(1)} \cdots a_{n+m}^{n+\rho(m)} \\ &= \sum_{\tau \in S_n} \text{sign } \tau \cdot a_1^{\tau(1)} \cdots a_n^{\tau(n)} \cdot \sum_{\rho \in S_m} \text{sign } \rho \cdot a_{n+1}^{n+\rho(1)} \cdots a_{n+m}^{n+\rho(m)} \\ &= \det A \cdot \det C. \end{aligned} \quad \square$$

Per Induktion folgt

Korollar 5.3.8. Sei $A = (a_j^i)_{1 \leq i, j \leq n}$ eine obere Dreiecksmatrix, d. h. gelte $a_j^i = 0$ für $i < j$. Dann gilt

$$\det A = a_1^1 \cdots a_n^n = \prod_{i=1}^n a_i^i.$$

Theorem 5.3.9. Sei A eine $(n \times n)$ -Matrix. Dann gilt

$$\det A = \det A^T.$$

Beweis. Sei $A = (a_j^i)$. Es folgt

$$\begin{aligned} \det A &= \sum_{\sigma \in S_n} \text{sign } \sigma \cdot a_1^{\sigma(1)} a_2^{\sigma(2)} \cdots a_n^{\sigma(n)} \\ &= \sum_{\sigma \in S_n} \underbrace{\text{sign } \sigma}_{=\text{sign}(\sigma^{-1})} \cdot a_{\sigma^{-1}(1)}^1 a_{\sigma^{-1}(2)}^2 \cdots a_{\sigma^{-1}(n)}^n \end{aligned}$$

(nach Umsortieren der Terme nach dem oberen Index)

$$\begin{aligned} &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{\sigma(1)}^1 a_{\sigma(2)}^2 \cdots a_{\sigma(n)}^n \\ &= \det A^T. \end{aligned} \quad \square$$

Korollar 5.3.10. Die Axiome, die wir für die Spalten einer Matrix in der Definition einer Determinantenfunktion gefordert haben, gelten somit auch für die Zeilen $a^i \equiv (a_j^i)_{1 \leq j \leq n}$, also etwa

$$\det A = - \det \begin{pmatrix} a^2 \\ a^1 \\ a^3 \\ \vdots \\ a^n \end{pmatrix}.$$

Wir können die Berechnung von Determinanten von $(n \times n)$ -Matrizen auf die von $((n-1) \times (n-1))$ -Matrizen zurückführen.

Theorem 5.3.11 (Laplace). *Sei A eine $(n \times n)$ -Matrix. Sei A_j^i die $((n-1) \times (n-1))$ -Matrix, die entsteht, wenn wir die i -te Zeile und die j -te Spalte streichen.*

$$\begin{pmatrix} a_1^1 & \cdots & a_{j-1}^1 & a_{j+1}^1 & \cdots & a_n^1 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_1^{i-1} & \cdots & a_{j-1}^{i-1} & a_{j+1}^{i-1} & \cdots & a_n^{i-1} \\ a_1^{i+1} & \cdots & a_{j-1}^{i+1} & a_{j+1}^{i+1} & \cdots & a_n^{i+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_1^n & \cdots & a_{j-1}^n & a_{j+1}^n & \cdots & a_n^n \end{pmatrix}$$

Dann gilt für festes j , $1 \leq j \leq n$,

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_j^i \det A_j^i$$

und für festes i , $1 \leq i \leq n$

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_j^i \det A_j^i.$$

Beweis. Wegen $\det A = \det A^T$ genügt es, die erste Gleichheit nachzuweisen.

Beide Seiten sind in der j -ten Spalte linear. Daher genügt es, die Aussage für Matrizen der Form

$$A = \begin{pmatrix} a_1^1 & \cdots & a_{j-1}^1 & 0 & a_{j+1}^1 & \cdots & a_n^1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_1^{i-1} & \cdots & a_{j-1}^{i-1} & 0 & a_{j+1}^{i-1} & \cdots & a_n^{i-1} \\ a_1^i & \cdots & a_{j-1}^i & a_j^i & a_{j+1}^i & \cdots & a_n^i \\ a_1^{i+1} & \cdots & a_{j-1}^{i+1} & 0 & a_{j+1}^{i+1} & \cdots & a_n^{i+1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_1^n & \cdots & a_{j-1}^n & 0 & a_{j+1}^n & \cdots & a_n^n \end{pmatrix}$$

nachzuweisen. Ohne Einschränkung können wir $a_j^i \neq 0$ annehmen und daher genügt es aufgrund der Eigenschaften der Determinante (die rechte Seite bleibt ohnehin unverändert), anzunehmen, dass A von der Form

$$A = \begin{pmatrix} a_1^1 & \cdots & a_{j-1}^1 & 0 & a_{j+1}^1 & \cdots & a_n^1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_1^{i-1} & \cdots & a_{j-1}^{i-1} & 0 & a_{j+1}^{i-1} & \cdots & a_n^{i-1} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_1^{i+1} & \cdots & a_{j-1}^{i+1} & 0 & a_{j+1}^{i+1} & \cdots & a_n^{i+1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_1^n & \cdots & a_{j-1}^n & 0 & a_{j+1}^n & \cdots & a_n^n \end{pmatrix}$$

ist. Weiterhin dürfen wir die ausgezeichnete Spalte mit der rechts davon vertauschen. Auf diese Weise bringen wir die ausgezeichnete Spalte ganz nach rechts. Ebenso dürfen wir die ausgezeichnete Zeile sukzessive mit der darunterliegenden Zeile vertauschen. Bei den Vertauschungen ändert sich jeweils das Vorzeichen auf beiden Seiten. Wir dürfen also ohne Einschränkung annehmen, dass

$$A = \begin{pmatrix} a_1^1 & \cdots & a_{n-1}^1 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ a_1^{n-1} & \cdots & a_{n-1}^{n-1} & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

gilt. Nun ist

$$\begin{aligned}
 \det A &= \sum_{\sigma \in S_n} \text{sign } \sigma \cdot a_1^{\sigma(1)} \cdots a_{n-1}^{\sigma(n-1)} \cdot a_n^{\sigma(n)} \\
 &= \sum_{\substack{\sigma \in S_n \\ \sigma(n)=n}} \text{sign } \sigma \cdot a_1^{\sigma(1)} \cdots a_{n-1}^{\sigma(n-1)} \cdot a_n^{\sigma(n)} \\
 &= \sum_{\sigma \in S_{n-1}} (-1)^{n+n} \text{sign } \sigma \cdot a_1^{\sigma(1)} \cdots a_{n-1}^{\sigma(n-1)} \cdot 1 \\
 &= (-1)^{n+n} \det A_n^n \\
 &= \sum_{i=1}^n (-1)^{i+n} a_n^i \det A_n^i.
 \end{aligned}$$

Die Behauptung folgt. \square

Die Determinante ist ein Gruppenhomomorphismus:

Theorem 5.3.12 (Determinantenmultiplikationssatz). *Seien $A, B \in F^{n \times n}$. Dann gilt*

$$\det(AB) = \det A \cdot \det B.$$

Beweis. Seien $A = (a_j^i)$, $B = (b_j^i)$. Dann ist $AB = \left(\sum_{k=1}^n a_k^i b_j^k \right)$. Setzen wir $b^i := (b_j^i)$, den Vektor, der in der i -ten Zeile von B steht, so können wir

$$AB = \begin{pmatrix} a_1^1 & \cdots & a_n^1 \\ \vdots & \ddots & \vdots \\ a_1^n & \cdots & a_n^n \end{pmatrix} \begin{pmatrix} b^1 \\ \vdots \\ b^n \end{pmatrix} = \begin{pmatrix} \sum_{i_1=1}^n a_{i_1}^1 b^{i_1} \\ \vdots \\ \sum_{i_n=1}^n a_{i_n}^n b^{i_n} \end{pmatrix}$$

schreiben, wenn wir eine Kollektion aus Zeilenvektoren als Matrix interpretieren. Wir erhalten aufgrund der Multilinearität und der Symmetrieeigenschaften der Determinante

$$\begin{aligned}
 \det(AB) &= \det \begin{pmatrix} \sum_{i_1=1}^n a_{i_1}^1 b^{i_1} \\ \vdots \\ \sum_{i_n=1}^n a_{i_n}^n b^{i_n} \end{pmatrix} \\
 &= \sum_{i_1=1}^n \sum_{i_2=1}^n \cdots \sum_{i_n=1}^n a_{i_1}^1 a_{i_2}^2 \cdots a_{i_n}^n \det \begin{pmatrix} b^{i_1} \\ \vdots \\ b^{i_n} \end{pmatrix} \\
 &= \sum_{\sigma \in S_n} a_{\sigma(1)}^1 a_{\sigma(2)}^2 \cdots a_{\sigma(n)}^n \cdot \det \begin{pmatrix} b^{\sigma(1)} \\ \vdots \\ b^{\sigma(n)} \end{pmatrix} \\
 &= \sum_{\sigma \in S_n} a_{\sigma(1)}^1 a_{\sigma(2)}^2 \cdots a_{\sigma(n)}^n \cdot \text{sign } \sigma \cdot \det \begin{pmatrix} b^1 \\ \vdots \\ b^n \end{pmatrix} \\
 &= \det A \cdot \det B.
 \end{aligned}$$

\square

Theorem 5.3.13. Eine quadratische Matrix $A \in F^{n \times n}$ ist genau dann regulär, wenn $\det A \neq 0$ gilt.

Beweis.

- (i) Sei A regulär. Dann existiert nach Theorem 3.8.5 eine Inverse A^{-1} . Der Determinantenmultiplikationssatz liefert

$$1 = \det \mathbf{1} = \det (AA^{-1}) = \det A \cdot \det A^{-1}.$$

Somit ist $\det A \neq 0$.

- (ii) Sei nun $\det A \neq 0$. Nach Theorem 5.3.4 sind daher die Spalten von A linear unabhängig. Somit ist A regulär. \square

Der Beweis liefert auch

Korollar 5.3.14. Sei A regulär. Dann gilt

$$\det A^{-1} = \frac{1}{\det A}.$$

Korollar 5.3.15. Ähnliche Matrizen haben die gleiche Determinante.

Beweis. Sei $B = DAD^{-1}$. Dann folgt

$$\det B = \det D \cdot \det A \cdot \det D^{-1} = \det D \cdot (\det D)^{-1} \cdot \det A = \det A. \quad \square$$

Da ähnliche Matrizen dieselbe Determinante haben, ist die Determinante eines Endomorphismusses zwischen endlichdimensionalen Vektorräumen wohldefiniert.

Definition 5.3.16. Sei $f: V \rightarrow V$ ein Endomorphismus. Sei A die Matrix zu f bezüglich einer beliebigen Basis. Setze

$$\det f := \det A.$$

Theorem 5.3.17. Sei $A = (a_j^i)$ eine $(n \times n)$ -Matrix. Sei A_j^i die $((n-1) \times (n-1))$ -Matrix, die wir erhalten, wenn wir die i -te Zeile und die j -te Spalte wie in Theorem 5.3.11 streichen. Definiere die $(n \times n)$ -Matrix B durch $b_j^i := (-1)^{i+j} \det A_j^i$. Dann gilt

$$AB = \det A \cdot \mathbf{1}.$$

B heißt die Adjunkte von A , A^{adj} .

Ist A regulär, so gilt

$$A^{-1} = \frac{B}{\det A}.$$

Beweis. Es gilt

$$\sum_{j=1}^n a_j^i b_k^j = \sum_{j=1}^n a_j^i (-1)^{j+k} \det A_j^k = \delta_k^i \det A,$$

denn für $i = k$ steht in der Mitte gerade die Entwicklung (siehe Theorem 5.3.11, Entwicklungssatz von Laplace) von $\det A$ nach der j -ten Spalte und für $i \neq k$ die Entwicklung der Determinante einer Matrix mit gleicher i -ter und k -ter Zeile. \square

5.4. Charakteristisches Polynom.

Theorem 5.4.1. Sei $f: V \rightarrow V$ durch die Matrix A beschrieben. Dann ist $\lambda \in F$ genau dann ein Eigenwert von f , wenn

$$\det(A - \lambda \mathbf{1}) = 0$$

gilt.

Beweis. Ist λ ein Eigenwert, so ist $A - \lambda \mathbf{1}$ singulär und somit $\det(A - \lambda \mathbf{1}) = 0$. Die Rückrichtung funktioniert genauso. \square

Die Determinantenabbildung ist ein Polynom in den Einträgen der Matrix. Daher können wir Determinanten nicht nur für Matrizen mit Einträgen in F , sondern auch mit Einträgen $a_j^i \in F[X]$ oder $F[\lambda]$ nach der Leibnizformel definieren.

Lemma 5.4.2. *Sei $A = (a_j^i)$ eine $(n \times n)$ -Matrix. Dann ist*

$$\chi_A(\lambda) := \det(A - \lambda \mathbf{1})$$

ein Polynom vom Grad n . Wir schreiben

$$\chi_A(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} S_1 \lambda^{n-1} + \dots + (-1)^0 S_n,$$

wobei hier zunächst $S_i \in F$ gilt.

Beweis. Da in jeder Zeile (oder Spalte) nur ein Term mit λ , nämlich $a_j^i - \lambda \delta_j^i$ vorkommt, folgt per Induktion (Leibnizformel), dass $\chi_A(\lambda)$ ein Polynom vom Grad $\leq n$ ist. Durch Entwicklung der Determinante sieht man, dass $\chi_A(\lambda)$ die Summe von

$$(a_1^1 - \lambda)(a_2^2 - \lambda) \cdots (a_n^n - \lambda)$$

und einem Polynom vom Grade $\leq n - 2$ ist. Somit ist $(-1)^n$ der Koeffizient vor λ^n . \square

Definition 5.4.3 (Elementarsymmetrische Funktionen). Seien $\mu_1, \dots, \mu_n \in F$. Definiere $\sigma_k \equiv \sigma_k(\mu) \equiv \sigma_k(\mu_1, \dots, \mu_n)$ durch

$$\begin{aligned} \sigma_0 &:= 1, \\ \sigma_1 &:= \sum_{1 \leq i \leq n} \mu_i, \\ \sigma_2 &:= \sum_{1 \leq i_1 < i_2 \leq n} \mu_{i_1} \mu_{i_2}, \\ \sigma_k &:= \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \mu_{i_1} \mu_{i_2} \cdots \mu_{i_k}, \quad 1 \leq k \leq n, \\ \sigma_n &:= \mu_1 \mu_2 \cdots \mu_n. \end{aligned}$$

Sei $A = (a_j^i) \in F^{n \times n}$. Wir definieren

$$\begin{aligned} S_0(A) &:= 1, \\ S_1(A) &:= \sum_{i=1}^n a_i^i =: \text{tr } A, \quad (\text{Spur von } A) \\ S_k(A) &:= \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \det (a_j^i)_{i,j \in \{i_1, \dots, i_k\}}, \\ S_n(A) &:= \det A. \end{aligned}$$

Häufiger setzt man alle anderen Funktionen gleich Null: $0 = S_{-1} = S_{-2} = \dots = S_{n+1} = S_{n+2} = \dots = \sigma_{-1} = \sigma_{-2} = \dots = \sigma_{n+1} = \sigma_{n+2} = \dots$

Bemerkung 5.4.4. Sei $\sigma \in S_n$, so gilt

$$\sigma_k(\mu_1, \dots, \mu_n) = \sigma_k(\sigma(\mu_1), \dots, \sigma(\mu_n)).$$

Ist $A \in F^{n \times n}$ diagonal, so ist

$$S_k(A) = \sigma_k(a_1^1, \dots, a_n^n).$$

Lemma 5.4.5. *Sei $A \in F^{n \times n}$. Dann gilt*

$$\det(A - \lambda \mathbf{1}) = (-1)^n \lambda^n + (-1)^{n-1} S_1(A) \lambda^{n-1} + \dots + (-1)^{n-k} S_k(A) \lambda^{n-k} + \dots + S_n(A).$$

Ist A diagonal, so gilt

$$\begin{aligned} \det(A - \lambda \mathbf{1}) &= (-1)^n \lambda^n + (-1)^{n-1} \sigma_1(a_1^1, \dots, a_n^n) \lambda^{n-1} + \dots \\ &\quad + (-1)^{n-k} \sigma_k(a_1^1, \dots, a_n^n) \lambda^{n-k} + \dots + \sigma_n(a_1^1, \dots, a_n^n). \end{aligned}$$

Beweis. Den allgemeinen Fall lassen wir als Übungsaufgabe.

Ist A diagonal und $n = 1$, so ist $\det(A - \lambda \mathbf{1}) = a_1^1 - \lambda$ wie behauptet. Gelte die Behauptung bereits für n . Sei innerhalb dieses Beweises σ_k die n -dimensionale k -te elementarsymmetrische Funktion von (a_1^1, \dots, a_n^n) und $\tilde{\sigma}_k$ die $(n+1)$ -dimensionale k -te elementarsymmetrische Funktion von $(a_1^1, \dots, a_{n+1}^{n+1})$. Es gilt für $1 \leq k \leq n$

$$\begin{aligned} \tilde{\sigma}_k &= \sum_{1 \leq i_1 < \dots < i_k \leq n+1} a_{i_1}^{i_1} \cdots a_{i_k}^{i_k} \\ &= \sum_{1 \leq i_1 < \dots < i_{k-1} < n+1} a_{i_1}^{i_1} \cdots a_{i_{k-1}}^{i_{k-1}} \cdot a_{n+1}^{n+1} + \sum_{1 \leq i_1 < \dots < i_k < n+1} a_{i_1}^{i_1} \cdots a_{i_k}^{i_k} \\ &= \sigma_{k-1} \cdot a_{n+1}^{n+1} + \sigma_k. \end{aligned}$$

Dann folgt für die Diagonalmatrix $A \in F^{(n+1) \times (n+1)}$ mit den Diagonaleinträgen $a_1^1, \dots, a_n^n, a_{n+1}^{n+1}$

$$\begin{aligned} \det(A - \lambda \mathbf{1}) &= (a_1^1 - \lambda) \cdot (a_2^2 - \lambda) \cdots (a_{n+1}^{n+1} - \lambda) \\ &= ((-1)^n \lambda^n + (-1)^{n-1} \sigma_1 \lambda^{n-1} + \dots + (-1)^{n-k} \sigma_k \lambda^{n-k} + \dots + \sigma_n) \cdot \\ &\quad \cdot (a_{n+1}^{n+1} - \lambda) \\ &= \left(\sum_{k=0}^n (-1)^{n-k} \sigma_k \lambda^{n-k} \right) \cdot (a_{n+1}^{n+1} - \lambda) \end{aligned}$$

und nach Indexverschiebung in der ersten Summe

$$\begin{aligned} &= \sum_{k=1}^{n+1} (-1)^{n+1-k} \sigma_{k-1} a_{n+1}^{n+1} \lambda^{n+1-k} + \sum_{k=0}^n (-1)^{n+1-k} \sigma_k \lambda^{n+1-k} \\ &= \underbrace{\sigma_n a_{n+1}^{n+1}}_{=\tilde{\sigma}_{n+1}} + \sum_{k=1}^n (-1)^{n+1-k} \underbrace{(\sigma_{k-1} a_{n+1}^{n+1} + \sigma_k)}_{=\tilde{\sigma}_k} \lambda^{n+1-k} + (-1)^{n+1} \sigma_0 \lambda^{n+1} \\ &= \sum_{k=0}^{n+1} (-1)^{n+1-k} \tilde{\sigma}_k \lambda^{n+1-k}. \end{aligned}$$

Die Behauptung für Diagonalmatrizen folgt. \square

Definition 5.4.6. Das Polynom

$$\chi_A(X) := \det(A - X \mathbf{1})$$

heißt charakteristisches Polynom der Matrix A . Sei f ein durch A dargestellter Endomorphismus. Setze

$$\chi_f(X) := \det(A - X \mathbf{1}).$$

Lemma 5.4.7. Sei die Abbildung f ein Endomorphismus eines endlichdimensionalen Vektorraumes. Dann ist $\chi_f(X)$ wohldefiniert.

Beweis. Sei S regulär und $B = SAS^{-1}$. Dann gilt

$$\begin{aligned} \chi_B(X) &= \det(SAS^{-1} - X \mathbf{1}) = \det(SAS^{-1} - SX \mathbf{1} S^{-1}) \\ &= \det(S(A - X \mathbf{1})S^{-1}) = \det S \cdot \det(A - X \mathbf{1}) \cdot \det S^{-1} \\ &= \det(A - X \mathbf{1}) = \chi_A(X). \end{aligned} \quad \square$$

Zusammenfassend erhalten wir

Theorem 5.4.8. *Die Eigenwerte eines linearen Endomorphismusses $f: V \rightarrow V$ zwischen endlichdimensionalen Vektorräumen sind genau die Nullstellen von $\chi_f(X)$ in F .*

Als Korollar zur Wohldefiniertheit von $\chi_f(X)$ erhalten wir

Korollar 5.4.9. *Seien A, A' ähnlich, so gilt $S_k(A) = S_k(A')$ für alle $1 \leq k \leq n$.*

Alternativ könnte man dieses Resultat für $k = 1$ mit dem folgenden Lemma beweisen.

Lemma 5.4.10. *Seien $A, B \in F^{n \times n}$. Dann gilt*

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

Beweis. Seien $A = (a_j^i)_{1 \leq i, j \leq n}$ und $B = (b_j^i)_{1 \leq i, j \leq n}$. Es gilt

$$\operatorname{tr}(AB) = \sum_{i,j=1}^n a_j^i b_i^j = \sum_{i,j=1}^n b_i^j a_j^i = \operatorname{tr}(BA). \quad \square$$

Zur geometrischen Vielfachheit eines Eigenwertes gibt es auch ein algebraisches Analogon

Definition 5.4.11. Sei $\lambda \in F$ ein Eigenwert einer linearen Selbstabbildung $f: V \rightarrow V$ (zwischen endlichdimensionalen Räumen; in Zukunft wollen wir diesen Zusatz immer als automatisch gegeben ansehen, wenn Determinanten oder charakteristische Polynome auftreten). Dann heißt die Vielfachheit der Nullstelle λ von $\chi_f(X)$ die algebraische Vielfachheit von λ .

Theorem 5.4.12. *Sei λ ein Eigenwert von $f: V \rightarrow V$. Dann ist die geometrische Vielfachheit von λ kleiner oder gleich der algebraischen Vielfachheit von λ .*

Im Beispiel $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ist die geometrische Vielfachheit des Eigenwertes 1 gleich eins, die algebraische ist 2.

Beweis. Sei $\dim E_\lambda = k$. Dann gibt es k linear unabhängige Eigenvektoren von f zum Eigenwert λ . Wir ergänzen diese zu einer Basis von V . Dann ist f bezüglich dieser Basis durch eine Matrix A der Form

$$\begin{pmatrix} \lambda & 0 & \dots & 0 & * & \dots & * \\ 0 & \lambda & \dots & 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda & * & \dots & * \\ 0 & 0 & \dots & 0 & b_1^1 & \dots & b_{n-k}^1 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & b_1^{n-k} & \dots & b_{n-k}^{n-k} \end{pmatrix}$$

mit geeigneten $b_j^i \in F$ dargestellt. Dabei bezeichnet $*$ einen beliebigen Eintrag aus F . Wir setzen $B := (b_j^i)_{1 \leq i, j \leq n-k}$. Wir erhalten durch sukzessive Entwicklung nach der ersten, zweiten, \dots , k -ten Spalte $\chi_f(X) = \chi_A(X) = \det(A - X\mathbf{1}_n) = (\lambda - X)^k \det(B - X\mathbf{1}_{n-k})$. Somit ist die algebraische Vielfachheit von λ mindestens gleich k . Die Behauptung folgt. \square

5.5. Diagonalisierbarkeit und Trigonalisierbarkeit.

Definition 5.5.1. Eine lineare Selbstabbildung $f: V \rightarrow V$ eines endlichdimensionalen Vektorraumes V heißt diagonalisierbar, wenn es eine Basis gibt, so dass f bezüglich dieser Basis durch eine Diagonalmatrix dargestellt wird.

Lemma 5.5.2. Sei $f: W \rightarrow W$ linear und sei $\dim W < \infty$. Sei U der Eigenraum zum Eigenwert λ und V der Eigenraum zum Eigenwert $\mu \neq \lambda$. Dann gilt

$$\dim(U + V) = \dim U + \dim V.$$

(Die Verallgemeinerung auf endlich viele Eigenräume beweist man mit einem analogen Beweis.)

Beweis. Sei $a_i, 1 \leq i \leq \dim U$, eine Basis von U und $b_j, 1 \leq j \leq \dim V$ eine Basis von V . Wir behaupten, dass die Vektoren $\{(a_i)_{1 \leq i \leq \dim U}, (b_j)_{1 \leq j \leq \dim V}\}$ linear unabhängig sind. Sonst erhalten wir eine nichttriviale Linearkombination der Null

$$\underbrace{\sum_{i=1}^{\dim U} \lambda^i a_i}_{=:a} + \underbrace{\sum_{j=1}^{\dim V} \mu^j b_j}_{=:b} = 0,$$

wobei a ein Eigenvektor zum Eigenwert λ und b ein Eigenvektor zum Eigenwert μ ist. Nach Theorem 4.7.3 folgt nun $a = b = 0$. \square

Theorem 5.5.3. Eine lineare Abbildung $f: V \rightarrow V$ ist genau dann diagonalisierbar, wenn das charakteristische Polynom $\chi_f(X)$ in Linearfaktoren zerfällt, d. h. wenn es sich in der Form $a \cdot (X - a_1) \cdots (X - a_n)$ für $a, a_i \in F$ schreiben lässt, und die geometrische Vielfachheit jedes Eigenwertes von f mit seiner algebraischen Vielfachheit übereinstimmt.

Beweis.

- (i) Ist f diagonalisierbar, so ist die Behauptung klar.
- (ii) Da $\chi_f(X)$ in Linearfaktoren zerfällt, summieren sich die geometrischen wie algebraischen Vielfachheiten der (einfach aufgeführten) Eigenwerte von f zu n auf. Wähle zu jedem Eigenraum eine Basis aus Eigenvektoren. Nach Lemma 5.5.2 bilden diese eine Basis eines n -dimensionalen Unterraumes, also von V selbst. Also besitzt V eine Basis aus Eigenvektoren von f . f ist somit diagonalisierbar. \square

Definition 5.5.4. Eine lineare Selbstabbildung $f: V \rightarrow V$ eines endlichdimensionalen Raumes V heißt genau dann trigonalisierbar, wenn es eine Basis gibt, so dass f durch eine obere Dreiecksmatrix dargestellt wird, d. h. durch eine Matrix $A = (a_j^i)$ mit $a_j^i = 0$ für $j < i$.

Theorem 5.5.5. Ein linearer Endomorphismus $f: V \rightarrow V$ ist genau dann trigonalisierbar, wenn das charakteristische Polynom $\chi_f(X)$ in Linearfaktoren zerfällt.

Beweis. Ist f trigonalisierbar, so ist die Behauptung klar.

Zerfälle das charakteristische Polynom von f in Linearfaktoren. Wir wollen per Induktion zeigen, dass f trigonalisierbar ist. Sei λ_1 eine Nullstelle von $\chi(f)$ und sei f durch die Matrix A dargestellt. Dann ist $A - \lambda_1 \mathbf{1}$ singulär, besitzt also einen nichttrivialen Kern. Somit gibt es ein $a_1 \in \ker(A - \lambda_1 \mathbf{1})$, also einen Eigenwert von A zum Eigenwert λ_1 . Bezüglich einer Basis (a_1, \dots, a_n) , a_i für $i \geq 2$ noch nicht

fixiert, in der der erste Vektor a_1 ist, hat A die Gestalt

$$\begin{pmatrix} \lambda_1 & a_2^1 & \dots & a_n^1 \\ 0 & a_2^2 & \dots & a_n^2 \\ 0 & a_2^3 & \dots & a_n^3 \\ \vdots & \vdots & & \vdots \\ 0 & a_2^n & \dots & a_n^n \end{pmatrix}.$$

Betrachte nun $g: \langle a_2, \dots, a_n \rangle \rightarrow \langle a_2, \dots, a_n \rangle$, definiert durch

$$\begin{pmatrix} a_2^2 & \dots & a_n^2 \\ a_2^3 & \dots & a_n^3 \\ \vdots & & \vdots \\ a_2^n & \dots & a_n^n \end{pmatrix}.$$

Es gilt $\chi_f(X) = (\lambda_1 - X) \cdot \chi_g(X)$. Da χ_f in Linearfaktoren zerfällt, gilt dies auch für $\chi_g(X)$. Nach Induktionsvoraussetzung gibt es daher eine Basis b_2, \dots, b_n , bezüglich der g durch eine obere Dreiecksmatrix dargestellt ist. Somit ist f bezüglich der Basis (a_1, b_2, \dots, b_n) durch eine obere Dreiecksmatrix dargestellt. \square

6. VEKTORRÄUME MIT SKALARPRODUKT

Sobald wir Skalarprodukte verwenden, wollen wir stets annehmen, dass wir \mathbb{R} - oder \mathbb{C} -Vektorräume betrachten.

6.1. Euklidische Vektorräume. Ein euklidischer Vektorraum ist ein Vektorraum mit einem reellen Skalarprodukt.

Definition 6.1.1. Sei V ein \mathbb{R} -Vektorraum. Ein (reelles) Skalarprodukt auf V ist eine Funktion $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$, die folgendes erfüllt:

- (i) $\langle a, b \rangle = \langle b, a \rangle$ für alle $a, b \in V$ (Symmetrie)
- (ii) $\langle \lambda a + \mu b, c \rangle = \lambda \langle a, c \rangle + \mu \langle b, c \rangle$ für alle $a, b, c \in V$ und $\lambda, \mu \in \mathbb{R}$ (Linearität)
- (iii) $\langle a, a \rangle \geq 0$ und $\langle a, a \rangle = 0 \iff a = 0$ für alle $a \in V$ (positive Definitheit)

Bemerkung 6.1.2.

- (i) Aus der Symmetrie und der Linearität im ersten Argument folgt auch die Linearität im zweiten Argument

$$\langle c, \lambda a + \mu b \rangle = \langle \lambda a + \mu b, c \rangle = \lambda \langle a, c \rangle + \mu \langle b, c \rangle = \lambda \langle c, a \rangle + \mu \langle c, b \rangle.$$

Eine Funktion in zwei Argumenten, die in beiden Argumenten linear ist, heißt bilinear.

- (ii) Per Induktion zeigt man für beliebige Linearkombinationen

$$\left\langle \sum_{i=1}^m \lambda^i a_i, \sum_{j=1}^n \mu^j b_j \right\rangle = \sum_{i=1}^m \sum_{j=1}^n \lambda^i \mu^j \langle a_i, b_j \rangle.$$

Somit ist ein Skalarprodukt durch seine Werte auf einer Basis eindeutig bestimmt.

- (iii) Sei $V = \mathbb{R}^n$. Für Vektoren $\xi = (\xi^1, \dots, \xi^n)$ und $\eta = (\eta^1, \dots, \eta^n)$ definieren wir

$$\langle \xi, \eta \rangle := \sum_{i=1}^n \xi^i \eta^i.$$

Dies ist ein Skalarprodukt auf \mathbb{R}^n (Übung), das Standard-Skalarprodukt.

(iv) Sei $V = \mathbb{R}^2$. Definiere für $\xi = (\xi^1, \xi^2)$ und $\eta = (\eta^1, \eta^2)$

$$\langle \xi, \eta \rangle := \xi^1 \eta^1 + 5\xi^1 \eta^2 + 5\xi^2 \eta^1 + 26\xi^2 \eta^2.$$

Dies ist ein weiteres Skalarprodukt auf \mathbb{R}^2 . Es stimmt nicht mit dem Standardskalarprodukt auf \mathbb{R}^2 überein.

(v) Sei V der Vektorraum der auf $[a, b] \subset \mathbb{R}$ stetigen reellwertigen Funktionen, $V = C^0([a, b])$. Gelte bei diesen Beispielen stets $a < b$. Definiere für $f, g \in V$

$$\langle f, g \rangle = \int_a^b f(x)g(x) dx.$$

Dies ist ein Skalarprodukt auf V (vgl. Analysis-Vorlesung).

Definition 6.1.3. Sei $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$ ein Skalarprodukt und sei a_1, \dots, a_n eine Basis von V . Dann definieren wir

$$c_{ij} := \langle a_i, a_j \rangle \quad \text{für } 1 \leq i, j \leq n.$$

$C = (c_{ij})_{1 \leq i, j \leq n}$ heißt Matrix zum Skalarprodukt $\langle \cdot, \cdot \rangle$.

Beachte, dass wir hier beide Indices unten schreiben. In Matrizenform stellen wir C durch

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

dar. Es ist nicht ideal, Matrizen der Form (c_j^i) und (c_{ij}) graphisch gleich darzustellen, jedoch üblich.

Theorem 6.1.4. Die Matrix $C = (c_{ij})_{1 \leq i, j \leq n}$ eines reellen Skalarproduktes ist symmetrisch, d. h. es gilt $c_{ij} = c_{ji}$ für alle $1 \leq i, j \leq n$ oder $C = C^T$ mit $C^T := (b_{ij})_{1 \leq i, j \leq n}$ mit $b_{ij} = c_{ji}$.

Beweis. Sei a_1, \dots, a_n eine Basis. Dann gilt

$$c_{ij} = \langle a_i, a_j \rangle = \langle a_j, a_i \rangle = c_{ji}. \quad \square$$

Beispiele 6.1.5.

- (i) Bezüglich der Standardbasis ist die Matrix des Standardskalarproduktes auf \mathbb{R}^n gleich $\mathbf{1} = I = (\delta_{ij})_{1 \leq i, j \leq n}$.
- (ii) Das Skalarprodukt mit

$$\langle \xi, \eta \rangle := \xi^1 \eta^1 + 5\xi^1 \eta^2 + 5\xi^2 \eta^1 + 26\xi^2 \eta^2$$

ist bezüglich der Standardbasis des \mathbb{R}^2 durch die Matrix

$$\begin{pmatrix} 1 & 5 \\ 5 & 26 \end{pmatrix}$$

und bezüglich der Basis aus den Vektoren $(1, 0)$ und $(-5, 1)$ durch die Matrix $\mathbf{1}$ dargestellt.

- (iii) Sei V der Vektorraum der Polynome vom Grad ≤ 3 mit Basis $(1, x, x^2, x^3)$. Dann ist das Skalarprodukt

$$\langle p, q \rangle = \int_0^1 p(x)q(x) dx$$

durch die Matrix

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}$$

dargestellt.

Lemma 6.1.6. Sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt, dem bezüglich der Basis a_1, \dots, a_n die Matrix $(c_{ij})_{1 \leq i, j \leq n}$ zugeordnet ist. Seien $\xi = \sum_{i=1}^n \xi^i a_i$ und $\eta = \sum_{j=1}^n \eta^j a_j$ beliebig. Dann gilt

$$\langle \xi, \eta \rangle = \sum_{i, j=1}^n \xi^i c_{ij} \eta^j.$$

Beweis. Benutze die Linearität in beiden Argumenten wie in Bemerkung 6.1.2 (ii). \square

Bemerkung 6.1.7. Wir wollen schließlich noch das Verhalten einer ein Skalarprodukt darstellenden Matrix unter Basistransformationen untersuchen.

Sei V ein reeller Vektorraum mit Basen S und T , $S = (a_1, \dots, a_n)$ und $T = (b_1, \dots, b_n)$. Gelte

$$b_k = \sum_{i=1}^n d_k^i a_i.$$

Das Skalarprodukt $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$ werde bezüglich S durch die Matrix $C = (c_{ij})_{1 \leq i, j \leq n}$ beschrieben. Dann gilt

$$\begin{aligned} m_{kl} := \langle b_k, b_l \rangle &= \left\langle \sum_{i=1}^n d_k^i a_i, \sum_{j=1}^n d_l^j a_j \right\rangle \\ &= \sum_{i, j=1}^n d_k^i d_l^j \langle a_i, a_j \rangle = \sum_{i, j=1}^n d_k^i d_l^j c_{ij}. \end{aligned}$$

Setze $M := (m_{ij})_{1 \leq i, j \leq n}$ und $D := (d_k^i)_{1 \leq i, k \leq n}$. Dann gilt

$$M = D^T C D.$$

6.2. Unitäre Vektorräume. Mit \bar{z} bezeichnen wir die komplex konjugierte Zahl zu z . Ist also $z = x + iy$ mit $x, y \in \mathbb{R}$, so ist $\bar{z} = x - iy$.

Ein Vektorraum mit einem unitären Skalarprodukt heißt unitärer Vektorraum.

Definition 6.2.1. Sei V ein \mathbb{C} -Vektorraum. Ein unitäres Skalarprodukt auf V ist eine Funktion $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{C}$, die die folgenden Eigenschaften erfüllt:

- (i) $\langle a, b \rangle = \overline{\langle b, a \rangle}$ für alle $a, b \in V$ (hermitesch)
- (ii) $\langle \lambda a + \mu b, c \rangle = \lambda \langle a, c \rangle + \mu \langle b, c \rangle$ für alle $a, b, c \in V$ und alle $\lambda, \mu \in \mathbb{C}$ (Linearität im ersten Argument)
- (iii) $\langle a, a \rangle \geq 0$ und $\langle a, a \rangle = 0 \iff a = 0$ für alle $a \in V$ (positiv definit)

Bemerkung 6.2.2. Folgende Eigenschaften und Beispiele sind analog zum reellen Fall

- (i) Seien $a, b, c \in V$ und $\lambda, \mu \in \mathbb{C}$. Dann gilt

$$\langle c, \lambda a + \mu b \rangle = \overline{\langle \lambda a + \mu b, c \rangle} = \overline{\lambda \langle a, c \rangle + \mu \langle b, c \rangle} = \bar{\lambda} \langle c, a \rangle + \bar{\mu} \langle c, b \rangle.$$

Linearität im ersten Argument und dieses Verhalten im zweiten Argument bezeichnet man als Sesquilinearität.

Es gibt auch die umgekehrte Konvention, d. h. man definiert ein unitäres Skalarprodukt so, dass es im zweiten Argument statt im ersten Argument linear ist und dass die übrigen Eigenschaften unverändert gelten. Im ersten Argument werden dann Skalare komplex konjugiert nach außen gezogen.

- (ii) Per Induktion folgt hieraus für beliebige Linearkombinationen

$$\left\langle \sum_{i=1}^m \lambda^i a_i, \sum_{j=1}^n \mu^j b_j \right\rangle = \sum_{i=1}^m \sum_{j=1}^n \lambda^i \overline{\mu^j} \langle a_i, b_j \rangle.$$

Daher ist ein unitäres Skalarprodukt durch seine Werte auf einer Basis bereits eindeutig bestimmt.

- (iii) Ist $V = \mathbb{C}^n$ und seien $x = (x^1, \dots, x^n)$ und $y = (y^1, \dots, y^n)$ Vektoren in V , so ist durch

$$\langle x, y \rangle := \sum_{i=1}^n x^i \overline{y^i}$$

ein unitäres Skalarprodukt auf \mathbb{C}^n definiert.

- (iv) Sei V der komplexe Vektorraum der auf $[a, b]$ komplexwertigen stetigen Funktionen einer reellen Variablen. Dann definiert

$$\langle f, g \rangle := \int_a^b f(t) \overline{g(t)} dt$$

ein unitäres Skalarprodukt auf V . Vergleiche wieder eine Analysis-Vorlesung für die positive Definitheit.

Definition 6.2.3. Sei V ein komplexer Vektorraum und $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{C}$ ein unitäres Skalarprodukt. Sei (a_1, \dots, a_n) eine Basis von V . Dann heißt die Matrix $C = (c_{ij})_{1 \leq i, j \leq n}$ mit

$$c_{ij} = \langle a_i, a_j \rangle, \quad 1 \leq i, j \leq n$$

die Matrix des Skalarproduktes $\langle \cdot, \cdot \rangle$ bezüglich der Basis a_1, \dots, a_n .

Theorem 6.2.4. Die Matrix C eines unitären Skalarproduktes ist hermitesch, d. h. es gilt $C^T = \overline{C}$.

Beweis. Es gilt

$$c_{ij} = \langle a_i, a_j \rangle = \overline{\langle a_j, a_i \rangle} = \overline{c_{ji}}.$$

Es folgt $C^T = \overline{C}$ wie behauptet. \square

Wie im reellen Fall zeigt man:

Lemma 6.2.5. Sei $\langle \cdot, \cdot \rangle$ ein unitäres Skalarprodukt, dem bezüglich einer Basis a_1, \dots, a_n die Matrix $(c_{ij})_{1 \leq i, j \leq n}$ zugeordnet ist. Seien $\xi = \sum_{i=1}^n \xi^i a_i$ und $\eta = \sum_{j=1}^n \eta^j a_j$ beliebig. Dann gilt

$$\langle \xi, \eta \rangle = \sum_{i, j=1}^n \xi^i c_{ij} \overline{\eta^j}.$$

Bemerkung 6.2.6. Wir wollen wiederum das Verhalten eines Skalarproduktes darstellenden Matrix unter Basistransformationen untersuchen.

Sei V ein komplexer Vektorraum mit Basen S und T , $S = (a_1, \dots, a_n)$ und $T = (b_1, \dots, b_n)$. Gelte

$$b_k = \sum_{i=1}^n d_k^i a_i.$$

Das Skalarprodukt $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{C}$ werde bezüglich S durch die Matrix $C = (c_{ij})_{1 \leq i, j \leq n}$ beschrieben. Dann gilt

$$\begin{aligned} m_{kl} &:= \langle b_k, b_l \rangle = \left\langle \sum_{i=1}^n d_k^i a_i, \sum_{j=1}^n d_l^j a_j \right\rangle \\ &= \sum_{i,j=1}^n d_k^i \overline{d_l^j} \langle a_i, a_j \rangle = \sum_{i,j=1}^n d_k^i \overline{d_l^j} c_{ij}. \end{aligned}$$

Setze $M := (m_{ij})_{1 \leq i, j \leq n}$ und $D := (d_k^i)_{1 \leq i, k \leq n}$. Dann gilt

$$M = D^T C D.$$

Bemerkung 6.2.7. Reelles und komplexes Skalarprodukt verhalten sich sehr ähnlich. Daher werden wir häufiger nur den komplexen Sachverhalt untersuchen. Ein analoges reelles Resultat folgt dann analog.

6.3. Norm.

Theorem 6.3.1 (Cauchy-Schwarzsche Ungleichung). *Sei V ein Vektorraum mit Skalarprodukt. Dann gilt*

$$|\langle a, b \rangle|^2 \leq \langle a, a \rangle \cdot \langle b, b \rangle$$

für alle $a, b \in V$. Gleichheit gilt genau dann, wenn a und b linear abhängig sind.

Beweis. Der Fall $b = 0$ ist einfach (Übung).

Sei $\lambda \in \mathbb{C}$ beliebig. Dann gilt

$$0 \leq \langle a - \lambda b, a - \lambda b \rangle = \langle a, a \rangle - \lambda \overline{\langle a, b \rangle} - \bar{\lambda} \langle a, b \rangle + \lambda \bar{\lambda} \langle b, b \rangle.$$

Setzen wir speziell $\lambda = \frac{\langle a, b \rangle}{\langle b, b \rangle}$, so folgt nach Multiplikation mit $\langle b, b \rangle$

$$0 \leq \langle a, a \rangle \cdot \langle b, b \rangle - 2|\langle a, b \rangle|^2 + |\langle a, b \rangle|^2.$$

Die behauptete Ungleichung folgt.

Gilt Gleichheit, so gilt insbesondere auch in der ersten Ungleichung Gleichheit, also $0 = a - \lambda b$ aufgrund der positiven Definitheit. \square

Korollar 6.3.2.

(i) *Seien a_1, \dots, a_n und b_1, \dots, b_n reelle Zahlen. Dann gilt*

$$\sum_{i=1}^n a_i b_i \leq \left(\sum_{i=1}^n a_i^2 \right)^{1/2} \cdot \left(\sum_{j=1}^n b_j^2 \right)^{1/2}.$$

(ii) *Seien $f, g: [a, b] \rightarrow \mathbb{C}$ stetig. Dann gilt*

$$\left| \int_a^b f(t)g(t) dt \right|^2 \leq \int_a^b |f(t)|^2 dt \cdot \int_a^b |g(t)|^2 dt.$$

Beweis. Dies folgt direkt aus der Cauchy-Schwarzschen Ungleichung für den Vektorraum

(i) \mathbb{R}^n mit Standardskalarprodukt.

(ii) der stetigen Funktionen auf $[a, b]$ mit $\int fg$ als Skalarprodukt. \square

Definition 6.3.3. Sei V ein reeller oder komplexer Vektorraum. Eine Funktion $\| \cdot \|: V \rightarrow \mathbb{R}$ heißt Norm auf V , wenn sie die folgenden Bedingungen erfüllt

(i) $\| \lambda a \| = |\lambda| \cdot \| a \|$ für alle $\lambda \in F$ und $a \in V$,

(ii) $\| a + b \| \leq \| a \| + \| b \|$ für alle $a, b \in V$,

(Dreiecksungleichung)

(iii) $\| a \| = 0 \iff a = 0$.

Theorem 6.3.4. Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Dann ist die Funktion $\| \cdot \|: V \rightarrow \mathbb{R}$, durch $\|a\| := \sqrt{\langle a, a \rangle}$ definiert, eine Norm.

Beweis.

$$(i) \quad \|\lambda a\| = \sqrt{\langle \lambda a, \lambda a \rangle} = \sqrt{\lambda \bar{\lambda} \langle a, a \rangle} = |\lambda| \sqrt{\langle a, a \rangle} = |\lambda| \cdot \|a\|.$$

$$(ii) \quad \begin{aligned} \|a + b\|^2 &= \langle a + b, a + b \rangle = \langle a, a \rangle + \langle a, b \rangle + \langle b, a \rangle + \langle b, b \rangle \\ &\leq \langle a, a \rangle + 2|\langle a, b \rangle| + \langle b, b \rangle \\ &\leq \|a\|^2 + 2 \cdot \|a\| \cdot \|b\| + \|b\|^2 \quad (\text{Cauchy-Schwarz}) \\ &= (\|a\| + \|b\|)^2. \end{aligned}$$

(iii) Die positive Definitheit der Norm folgt aus der positiven Definitheit des Skalarproduktes. \square

In einem Skalarproduktraum (=Vektorraum mit Skalarprodukt) werden wir unter einer Norm immer $\|a\| := \sqrt{\langle a, a \rangle}$ verstehen.

6.4. Orthonormalbasen.

Definition 6.4.1. Sei V ein Skalarproduktraum. Dann heißt $a \in V$ senkrecht zu $b \in V$, wenn $\langle a, b \rangle = 0$ gilt. Eine Teilmenge $S \subset V$ heißt orthogonal zu einer Teilmenge $T \subset V$, wenn $\langle s, t \rangle = 0$ für alle $(s, t) \in S \times T$ gilt.

Bemerkung 6.4.2.

- (i) Ist a orthogonal zu b , so ist b orthogonal zu a .
- (ii) Der Nullvektor ist zu allen anderen Vektoren orthogonal, da aus

$$\langle 0, a \rangle = \langle 0 + 0, a \rangle = \langle 0, a \rangle + \langle 0, a \rangle$$

$\langle 0, a \rangle = 0$ folgt. Dies ist auch der einzige Vektor mit dieser Eigenschaft, da aus $\langle a, a \rangle = 0$ bereits $a = 0$ folgt.

- (iii) Zwei Teilmengen S und T sind genau dann orthogonal, wenn $\langle S \rangle$ und $\langle T \rangle$ orthogonal sind (Übung).

Definition 6.4.3.

- (i) Eine nichtleere Familie/Teilmenge S von V heißt orthogonal, wenn je zwei verschiedene Elemente aus S orthogonal zueinander sind.
- (ii) Eine orthogonale Familie S heißt orthonormiert, wenn $\langle a, a \rangle = 1$ für alle $a \in S$ gilt.
- (iii) Eine orthonormierte Familie, die zugleich Basis von V ist, heißt orthonormierte Basis oder Orthonormalbasis von V .

Bemerkung 6.4.4.

- (i) Eine orthogonale Familie S von Vektoren mit $0 \notin S$ kann man zu einer orthonormalen Familie machen, indem man jeden Vektor $a \in S$ durch $\frac{a}{\|a\|}$ ersetzt. Die neue Familie ist dann durch Normieren aus der alten hervorgegangen (Übung).
- (ii) Sei V der Vektorraum der auf $[-\pi, \pi]$ stetigen reellwertigen Funktionen mit Skalarprodukt

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x) dx.$$

Dann ist

$$\{1, \sin x, \cos x, \sin 2x, \cos 2x, \sin 3x, \cos 3x, \dots\}$$

eine orthogonale Familie (Übung). Dies spielt bei Fourierreihen eine Rolle.

Theorem 6.4.5. Sei S eine orthogonale Familie mit $0 \notin S$. Dann ist S linear unabhängig.

Beweis. Gelte

$$\sum_{i=1}^n \lambda^i a_i = 0$$

mit $\lambda^i \in F$ und $a_i \in S$ für $1 \leq i \leq n$ und ein $n \in \mathbb{N}^+$. Dann folgt für $1 \leq j \leq n$

$$0 = \langle 0, a_j \rangle = \left\langle \sum_{i=1}^n \lambda^i a_i, a_j \right\rangle = \sum_{i=1}^n \lambda^i \langle a_i, a_j \rangle = \lambda^j \langle a_j, a_j \rangle.$$

Wegen $a_j \neq 0$ folgt also $\lambda^j = 0$. Somit ist S linear unabhängig. \square

Theorem 6.4.6. Sei V ein endlichdimensionaler Skalarproduktraum. Sei a_1, \dots, a_n eine Orthonormalbasis. Dann gilt für jedes $b \in V$

$$b = \sum_{i=1}^n \langle b, a_i \rangle a_i.$$

Beweis. Wir wissen bereits, dass sich b als Linearkombination der Form

$$b = \sum_{i=1}^n \lambda^i a_i$$

darstellen lässt. Hieraus folgt

$$\langle b, a_j \rangle = \left\langle \sum_{i=1}^n \lambda^i a_i, a_j \right\rangle = \sum_{i=1}^n \lambda^i \langle a_i, a_j \rangle = \lambda^j.$$

Wir erhalten die Behauptung. \square

Der folgende Algorithmus erlaubt es, aus einer Basis eine Orthogonalbasis zu gewinnen.

Theorem 6.4.7 (Gram-Schmidtsches Orthogonalisierungsverfahren). Jeder endlichdimensionale Skalarproduktraum besitzt eine Orthonormalbasis.

Beweis. Sei a_1, \dots, a_n eine beliebige Basis. Daraus konstruieren wir induktiv eine orthogonale Basis vermöge

$$b_{k+1} := a_{k+1} - \sum_{j=1}^k \frac{\langle b_j, a_{k+1} \rangle}{\langle b_j, b_j \rangle} b_j.$$

Nach Definition handelt es sich weiterhin um eine Basis. Insbesondere gilt also $b_j \neq 0$ für alle j . Die Familie der b_j ist orthogonal, da wir für $i \in \{1, \dots, k\}$ induktiv

$$\begin{aligned} \langle b_i, b_{k+1} \rangle &= \langle b_i, a_{k+1} \rangle - \sum_{j=1}^k \frac{\langle b_j, a_{k+1} \rangle}{\langle b_j, b_j \rangle} \langle b_i, b_j \rangle \\ &= \langle b_i, a_{k+1} \rangle - \langle b_i, a_{k+1} \rangle \\ &= 0 \end{aligned}$$

erhalten. Wir normieren nun die Vektoren b_i . Nach Theorem 6.4.5 sind sie linear unabhängig. Aufgrund ihrer Anzahl handelt es sich somit um eine Orthonormalbasis. \square

Bemerkung 6.4.8. Um die Fälle $F = \mathbb{R}$ und $F = \mathbb{C}$ gleichzeitig behandeln zu können, setzen wir $A^* := \bar{A}^T$ für $A \in F^{n \times n}$. Im Reellen gilt $A^* = A^T$.

Theorem 6.4.9. Sei V ein Skalarproduktraum, $\dim V < \infty$. Sei $\{b_1, \dots, b_n\}$ eine Orthonormalbasis. Definiere Vektoren d_k durch

$$d_k = \sum_{i=1}^n a_k^i b_i.$$

Setze $A := (a_j^i)_{1 \leq i, j \leq n}$. Dann ist $\{d_1, \dots, d_n\}$ genau dann eine Orthonormalbasis von V , wenn

$$A^* A = I$$

gilt.

Beweis. Wie angekündigt zeigen wir nur den unitären Fall. Es gilt

$$\begin{aligned} \langle d_k, d_l \rangle &= \left\langle \sum_{i=1}^n a_k^i b_i, \sum_{j=1}^n a_l^j b_j \right\rangle = \sum_{i,j=1}^n a_k^i \overline{a_l^j} \langle b_i, b_j \rangle \\ &= \sum_{i,j=1}^n a_k^i \overline{a_l^j} \delta_{ij} = \sum_{i=1}^n a_k^i \overline{a_l^i}. \end{aligned}$$

Sei also $\{d_1, \dots, d_n\}$ orthonormal. Dann folgt

$$\delta_{kl} = \sum_{i=1}^n a_k^i \overline{a_l^i} \quad \text{oder} \quad A^T \bar{A} = I.$$

Durch komplexes Konjugieren erhalten wir $A^* A = I$.

Gelte umgekehrt $A^* A = I$. Mit der obigen Rechnung erhalten wir daher $\langle d_k, d_l \rangle = \delta_{kl}$. Somit ist $\{d_1, \dots, d_n\}$ orthonormiert und daher auch linear unabhängig; es handelt sich somit um eine Orthonormalbasis von V . \square

Definition 6.4.10. Sei A eine $(n \times n)$ -Matrix mit $A^* = A^{-1}$. Ist $A \in \mathbb{R}^{n \times n}$, so heißt A orthogonal; ist $A \in \mathbb{C}^{n \times n}$, so heißt A unitär.

Theorem 6.4.11. Sei $A \in \mathbb{R}^{n \times n}$. Dann sind die folgenden Aussagen äquivalent:

- (i) A ist orthogonal
- (ii) $A^T A = I$
- (iii) $AA^T = I$
- (iv) A vermittelt eine Basistransformation zwischen orthonormierten Basen eines n -dimensionalen euklidischen Vektorraumes.
- (v) Die Spalten der Matrix A bilden eine orthonormierte Basis des Vektorraumes \mathbb{R}^n mit Standardskalarprodukt.
- (vi) Die Zeilen der Matrix A bilden eine orthonormierte Basis des Vektorraumes \mathbb{R}^n mit Standardskalarprodukt.

Beweis. Nach Definition sind die Aussagen (i), (ii) und (iii) äquivalent. Theorem 6.4.9 impliziert, dass (ii) und (iv) äquivalent sind.

Sei $A = (a_j^i)_{1 \leq i, j \leq n}$. Dann folgt aus (ii)

$$\sum_{i=1}^n a_k^i a_l^i = \delta_{kl} \quad \text{für } 1 \leq k, l \leq n.$$

(v) folgt. Die Umkehrung „(v) \implies (ii)“ folgt ebenso aus dieser Gleichung.

Genauso wie man die Äquivalenz zwischen (ii) und (v) zeigt, erhält man auch die Äquivalenz von (iii) und (vi). \square

Theorem 6.4.12. Sei $A \in \mathbb{C}^{n \times n}$. Dann sind die folgenden Aussagen äquivalent:

- (i) A ist unitär
- (ii) $A^* A = I$

- (iii) $AA^* = I$
- (iv) A vermittelt eine Basistransformation zwischen orthonormierten Basen eines n -dimensionalen unitären Vektorraumes.
- (v) Die Spalten der Matrix A bilden eine orthonormierte Basis des Vektorraumes \mathbb{C}^n mit Standardskalarprodukt.
- (vi) Die Zeilen der Matrix A bilden eine orthonormierte Basis des Vektorraumes \mathbb{C}^n mit Standardskalarprodukt.

Beweis. Vollständig analog zum reellen Fall. □

6.5. Orthogonale und unitäre Endomorphismen.

Definition 6.5.1. Sei V ein Skalarproduktraum. Dann heißt ein Endomorphismus $f: V \rightarrow V$ orthogonal (bzw. unitär), falls

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

gilt.

Theorem 6.5.2. Sei V ein endlichdimensionaler Skalarproduktraum. Sei $f: V \rightarrow V$ ein orthogonaler (bzw. unitärer) Endomorphismus. Dann gilt

- (i) $\|f(a)\| = \|a\|$
- (ii) Sind a und b orthogonal, so auch $f(a)$ und $f(b)$.
- (iii) Ist λ ein Eigenwert von f , so gilt $|\lambda| = 1$.
- (iv) f ist eine Isomorphismus.

Beweis.

- (i) Es ist

$$\|f(a)\| = \sqrt{\langle f(a), f(a) \rangle} = \sqrt{\langle a, a \rangle} = \|a\|.$$

- (ii) Dies folgt aus

$$0 = \langle a, b \rangle = \langle f(a), f(b) \rangle.$$

- (iii) Ist a ein Eigenvektor zum Eigenwert λ , so gilt nach nach (i)

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

Wegen $a \neq 0$ folgt die Behauptung.

- (iv) Wegen (iii) ist 0 kein Eigenwert. Also folgt $\ker f = \{0\}$, f ist also injektiv. Da V endlichdimensional ist, ist f auch surjektiv und damit ein Isomorphismus.

(Wir bemerken, dass (iv) für unendlichdimensionale Skalarprodukträume nicht aus unserer Definition folgt: Betrachte $f: l^2 \rightarrow l^2$ mit $\langle (a_i)_{i \in \mathbb{N}}, (b_j)_{j \in \mathbb{N}} \rangle := \sum_{i \in \mathbb{N}} a_i \bar{b}_i$ und $l^2 \equiv l^2(\mathbb{N}, \mathbb{C}) \subset \mathbb{C}^{\mathbb{N}}$ der Teilmenge, auf der die zugehörige Norm endlich ist, sowie

$$f(a_0, a_1, a_2, a_3, \dots) := (0, a_0, a_1, a_2, a_3, \dots).$$

Dann ist f nicht surjektiv. Die anderen Teilaussagen gelten mit demselben Beweis auch für unendlichdimensionale Skalarprodukträume.) □

Theorem 6.5.3. Sei $\{a_1, \dots, a_n\}$ eine Orthonormalbasis von V . Dann ist ein Endomorphismus $f: V \rightarrow V$ genau dann orthogonal (bzw. unitär), falls

$$\langle f(a_i), f(a_j) \rangle = \delta_{ij}$$

für alle $1 \leq i, j \leq n$ gilt.

Beweis. Es ist klar, dass aus der Orthogonalität (bzw. Unitarität) $\langle f(a_i), f(a_j) \rangle = \delta_{ij}$ folgt.

Gelte also $\langle f(a_i), f(a_j) \rangle = \delta_{ij}$. Wir wollen nachweisen, dass f „das Skalarprodukt erhält“: Seien $a = \sum_{i=1}^n \lambda^i a_i$ und $b = \sum_{j=1}^n \mu^j a_j$. Wir erhalten

$$\begin{aligned} \langle f(a), f(b) \rangle &= \left\langle f \left(\sum_{i=1}^n \lambda^i a_i \right), f \left(\sum_{j=1}^n \mu^j a_j \right) \right\rangle \\ &= \sum_{i,j=1}^n \lambda^i \overline{\mu^j} \underbrace{\langle f(a_i), f(a_j) \rangle}_{=\delta_{ij}=\langle a_i, a_j \rangle} \\ &= \left\langle \sum_{i=1}^n \lambda^i a_i, \sum_{j=1}^n \mu^j a_j \right\rangle = \langle a, b \rangle. \end{aligned}$$

Somit ist f orthogonal (bzw. unitär). \square

Theorem 6.5.4. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei S eine Orthonormalbasis. Dann ist ein Endomorphismus $f: V \rightarrow V$ genau dann orthogonal (bzw. unitär), wenn die Matrix A von f bezüglich S orthogonal (bzw. unitär) ist, d. h. wenn $A^*A = I$ gilt.*

Beweis. Sei $S = \{b_1, \dots, b_n\}$ und $A = (a_j^i)_{1 \leq i, j \leq n}$, d. h. es gelte

$$f(b_k) = \sum_{i=1}^n a_k^i b_i.$$

Wir erhalten

$$\langle f(b_k), f(b_l) \rangle = \sum_{i,j=1}^n a_k^i \overline{a_l^j} \underbrace{\langle b_i, b_j \rangle}_{=\delta_{ij}} = \sum_{i=1}^n \overline{a_l^i} a_k^i.$$

Nach Theorem 6.5.3 ist f genau dann orthogonal, wenn $\langle f(b_k), f(b_l) \rangle = \delta_{kl}$ gilt. Aufgrund unserer Rechnung ist dies aber äquivalent zu $\delta_{kl} = \sum_{i=1}^n \overline{a_l^i} a_k^i$ und somit auch zu $A^*A = I$. Die Behauptung folgt. \square

Beispiel 6.5.5. Bezüglich des Standardskalarproduktes des \mathbb{R}^2 sind die Matrizen

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

für beliebige $\varphi \in \mathbb{R}$ orthogonal, insbesondere also auch I und $-I$.

Theorem 6.5.6. *Seien A, B orthogonale (bzw. unitäre) $(n \times n)$ -Matrizen. Dann gelten*

- (i) A^{-1} ist orthogonal (bzw. unitär).
- (ii) AB ist orthogonal (bzw. unitär).
- (iii) $|\det A| = 1$.

Die orthogonalen (bzw. unitären) $(n \times n)$ -Matrizen bilden somit eine Untergruppe der invertierbaren Matrizen $GL_n(F)$,

- die orthogonale Gruppe $O(n)$ bzw.
- die unitäre Gruppe $U(n)$.

Weiterhin definiert man:

- Die spezielle lineare Gruppe $SL(n, F) \equiv SL_n(F)$ besteht aus den Elementen der “general linear group” $GL(n, F) \equiv GL_n(F)$ mit Determinante 1.
- Die spezielle orthogonale Gruppe $SO(n)$ besteht aus den orthogonalen Matrizen A mit $\det A = 1$.

Der Determinantenmultiplikationssatz liefert, dass es sich damit tatsächlich um Untergruppen von $GL_n(F)$ bzw. $O(n)$ handelt.

Beweis.

- (i) Nach Definition ist $A^*A = I$. Weiterhin gilt $AA^* = I$. Also ist A^* die Inverse zu A , $A^{-1} = A^*$. Es folgt aus

$$(A^{-1})^* A^{-1} = (A^*)^* A^* = AA^* = I,$$

dass auch A^{-1} orthogonal ist.

- (ii) Dies folgt aus

$$(AB)^*(AB) = B^*(A^*A)B = B^*IB = B^*B = I.$$

- (iii) Es gilt

$$\begin{aligned} 1 &= \det I = \det(A^*A) = \det A^* \cdot \det A = \det \overline{A}^T \cdot \det A \\ &= \det \overline{A} \cdot \det A = \overline{\det A} \cdot \det A = |\det A|^2. \end{aligned}$$

Die Behauptung folgt. \square

6.6. Orthogonale Komplemente.

Definition 6.6.1. Sei V ein Skalarproduktraum und sei $S \subset V$ beliebig. Dann heißt

$$S^\perp := \{a \in V : \langle a, b \rangle = 0 \text{ für alle } b \in S\}$$

das orthogonale Komplement von S in V .

Theorem 6.6.2. Sei V ein Skalarproduktraum und sei $S \subset V$. Dann gelten

- (i) S^\perp ist ein Unterraum von V ,
- (ii) $\langle S \rangle \cap S^\perp = \{0\}$,
- (iii) $(S^\perp)^\perp \supset \langle S \rangle$ und
- (iv) $S^\perp = \langle S \rangle^\perp$.

Beweis.

- (i) Seien $a, b \in S^\perp$ und $\lambda, \mu \in F$. Dann gilt für alle $c \in S$

$$\langle \lambda a + \mu b, c \rangle = \lambda \langle a, c \rangle + \mu \langle b, c \rangle = 0.$$

Somit ist auch $\lambda a + \mu b \in S^\perp$. Wegen $0 \in S^\perp$ ist S^\perp nicht leer.

- (ii) Sei $a \in S^\perp \cap \langle S \rangle$. Dann gibt es $\lambda^i \in F$ und $a_i \in S$ mit

$$\sum_{i=1}^n \lambda^i a_i = a.$$

Wegen $a \in S^\perp$ folgt hieraus

$$\langle a, a \rangle = \sum_{i=1}^n \overline{\lambda^i} \langle a, a_i \rangle = 0.$$

Somit ist $a = 0$ und wir erhalten die Behauptung.

- (iii) Seien $a \in S$ und $b \in S^\perp$. Dann gilt $\langle a, b \rangle = \overline{\langle b, a \rangle} = 0$. Somit ist $S \subset (S^\perp)^\perp$. Da $(S^\perp)^\perp$ ein Unterraum ist, folgt sogar $\langle S \rangle \subset (S^\perp)^\perp$.

(Ist $\dim V < \infty$, so gilt sogar Gleichheit (s. u.). Im Raum $l^2 \equiv l^2(\mathbb{N}) \equiv l^2(\mathbb{N}, \mathbb{R}) \subset \mathbb{R}^{\mathbb{N}}$ ist $S = \{e_0, e_1, e_2, \dots\}$ mit $S^\perp = \{0\}$ aber $\langle S \rangle \subsetneq l^2$ ein Gegenbeispiel, da $l^2(\mathbb{N}, \mathbb{R})$ auch Folgen mit unendlich vielen von Null verschiedenen Einträgen enthält.)

(iv) Wegen $S \subset \langle S \rangle$ ist klar, dass $\langle S \rangle^\perp \subset S^\perp$ gilt. Sei also $a \in S^\perp$ und $b = \sum_{i=1}^n \lambda^i b_i$ mit $b_i \in S$ ein beliebiges Element in $\langle S \rangle$. Dann gilt

$$\langle a, b \rangle = \sum_{i=1}^n \overline{\lambda^i} \langle a, b_i \rangle = 0.$$

Somit gilt auch $a \in \langle S \rangle^\perp$. \square

Theorem 6.6.3. *Sei V ein endlichdimensionaler Skalarproduktraum. Ist $U \subset V$ ein Unterraum, so gilt $V = U \oplus U^\perp$.*

Beweis. Nach Theorem 6.6.2 gilt $U \cap U^\perp = \{0\}$. Es genügt also nachzuweisen, dass sich jeder Vektor $c \in V$ in der Form $c = a + b$ mit $a \in U$ und $b \in U^\perp$ darstellen lässt. Sei $\{a_1, \dots, a_r\}$ eine Orthonormalbasis von U . Definiere a durch

$$a := \langle c, a_1 \rangle a_1 + \dots + \langle c, a_r \rangle a_r$$

und setze $b := c - a$. Dann ist offensichtlich $c = a + b$. Wir müssen also noch zeigen, dass $b \in U^\perp$ gilt: Es ist

$$\begin{aligned} \langle b, a_i \rangle &= \langle c - a, a_i \rangle \\ &= \langle c, a_i \rangle - \left\langle \sum_{j=1}^r \langle c, a_j \rangle a_j, a_i \right\rangle \\ &= \langle c, a_i \rangle - \sum_{j=1}^r \langle c, a_j \rangle \delta_{ji} = 0. \end{aligned}$$

Somit steht b orthogonal zu einer Basis von U . Wir erhalten $b \in U^\perp$ und die Behauptung folgt. \square

Theorem 6.6.4. *Sei V ein endlichdimensionaler Skalarproduktraum. Ist $U \subset V$ ein Unterraum, so gelten*

- (i) $\dim V = \dim U + \dim U^\perp$.
- (ii) $(U^\perp)^\perp = U$,

Beweis.

- (i) Dies folgt direkt aus Theorem 6.6.3 und der Dimensionsformel, Theorem 3.3.6.
- (ii) Es gilt $V = U \oplus U^\perp$ und $V = (U^\perp)^\perp \oplus \left((U^\perp)^\perp \right)^\perp$. Nach Theorem 6.6.2 gilt aber $U \subset (U^\perp)^\perp$ und ebenso $U^\perp \subset \left((U^\perp)^\perp \right)^\perp$. Dies ist aber nur möglich, wenn bereits $U = (U^\perp)^\perp$ und $U^\perp = \left((U^\perp)^\perp \right)^\perp$ gelten. \square

Beispiel 6.6.5. Sei V der Vektorraum der auf $[-a, a]$, $a > 0$, stetigen reellwertigen Funktionen mit (L^2) -Skalarprodukt

$$\langle f, g \rangle := \int_{-a}^a f(x)g(x) dx.$$

Sei $U := \{f \in V : f(-x) = -f(x)\}$ der Raum der ungeraden Funktionen und $G := \{f \in V : f(-x) = f(x)\}$ der Raum der geraden Funktionen in V . Wir behaupten, dass $U^\perp = G$ gilt.

Beweis. Zunächst einmal ist klar, dass $G \subset U^\perp$ gilt. Sei nun $h \in U^\perp$ beliebig. Wir setzen $h_1(x) := \frac{1}{2}(h(x) + h(-x))$ und $h_2(x) := \frac{1}{2}(h(x) - h(-x))$. Dann gilt

$h = h_1 + h_2$. Es ist $h_1 \in G$ und $h_2 \in U$. Wir sind fertig, wenn wir $h_2 \equiv 0$ zeigen können. Aus $h \in U^\perp$ und $h_2 \in U$ erhalten wir

$$0 = \langle h, h_2 \rangle = \int_{-a}^a h(x)h_2(x) dx = \int_{-a}^a h_1(x)h_2(x) dx + \int_{-a}^a h_2(x)h_2(x) dx.$$

Das erste Integral auf der rechten Seite mit einer geraden und einer ungeraden Funktion verschwindet. Da h und somit auch h_2 stetig ist, folgt also $h_2 \equiv 0$ wie behauptet. \square

6.7. Adjungierte Abbildungen. Zu einem Endomorphismus $f: V \rightarrow V$ wollen wir eine lineare Abbildung $g: V \rightarrow V$ mit

$$\langle f(a), b \rangle = \langle a, g(b) \rangle$$

finden.

Definition 6.7.1. Sei V ein Skalarproduktraum. Zu $a \in V$ definieren wir $a^*: V \rightarrow \mathbb{R}$ (bzw. $a^*: V \rightarrow \mathbb{C}$) durch

$$a^*(b) = \langle b, a \rangle \quad \text{für } b \in V.$$

Bemerkung: Die Abbildung a^* ist linear, denn es gilt

$$a^*(\lambda b + \mu c) = \langle \lambda b + \mu c, a \rangle = \lambda \langle b, a \rangle + \mu \langle c, a \rangle = \lambda a^*(b) + \mu a^*(c).$$

Theorem 6.7.2. Sei V ein endlichdimensionaler Skalarproduktraum. Dann gibt es zu jedem Funktional φ ein eindeutig bestimmtes $a \in V$ mit $\varphi(\xi) = \langle \xi, a \rangle$ für alle $\xi \in V$. Somit ist $\varphi = a^*$.

Beweis. Sei $\{a_1, \dots, a_n\}$ eine orthonormierte Basis von V . Definiere a durch

$$a := \sum_{i=1}^n \overline{\varphi(a_i)} a_i.$$

Wir erhalten

$$\langle a_j, a \rangle = \sum_{i=1}^n \overline{\varphi(a_i)} \langle a_j, a_i \rangle = \varphi(a_j).$$

Also stimmen a^* und φ auf einer Basis von V überein und sind daher gleich.

Es bleibt noch zu zeigen, dass a eindeutig bestimmt ist. Gelte also $\langle \xi, a \rangle = \langle \xi, a' \rangle$ für alle $\xi \in V$. Dies ist äquivalent zu $\langle \xi, a - a' \rangle = 0$ für alle $\xi \in V$ und gilt insbesondere für $\xi = a - a'$. Also folgt $a = a'$. \square

Bemerkung 6.7.3. Aus diesem Beweis folgt insbesondere, dass $\langle \xi, a \rangle = 0$ für alle $\xi \in V$ nur für $a = 0$ erfüllt sein kann.

Bemerkung 6.7.4. Sei $f: V \rightarrow V$ ein Endomorphismus. Ist $a \in V$ fest, so definieren wir durch

$$\varphi(\xi) := \langle f(\xi), a \rangle$$

ein lineares Funktional $\varphi: V \rightarrow F$: Es gilt nämlich

$$\varphi(\lambda b + \mu c) = \langle f(\lambda b + \mu c), a \rangle = \lambda \langle f(b), a \rangle + \mu \langle f(c), a \rangle = \lambda \varphi(b) + \mu \varphi(c).$$

Nach Theorem 6.7.2 gibt es zu φ ein eindeutig bestimmtes $a_0 \in V$ mit $\varphi(\xi) = \langle \xi, a_0 \rangle$ für alle $\xi \in V$. Für fixiertes f ordnen wir auf diese Weise jedem $a \in V$ ein $a_0 \in V$ zu. Wir bezeichnen diese Zuordnung als $g: V \rightarrow V$, $g(a) := a_0$. g ist durch die Gleichung

$$\langle f(\xi), a \rangle = \langle \xi, g(a) \rangle \quad \text{für alle } \xi \in V$$

festgelegt. Wir behaupten, dass g eine lineare Abbildung ist. Aus der definierenden Gleichung erhalten wir

$$\begin{aligned}\langle \xi, g(\lambda a + \mu b) \rangle &= \langle f(\xi), \lambda a + \mu b \rangle = \bar{\lambda} \langle f(\xi), a \rangle + \bar{\mu} \langle f(\xi), b \rangle \\ &= \bar{\lambda} \langle \xi, g(a) \rangle + \bar{\mu} \langle \xi, g(b) \rangle = \langle \xi, \lambda g(a) + \mu g(b) \rangle.\end{aligned}$$

Da dies für alle $\xi \in V$ gilt, erhalten wir $g(\lambda a + \mu b) = \lambda g(a) + \mu g(b)$, also die Linearität von g .

Wir sagen, dass g die zu f adjungierte Abbildung ist und schreiben $g = f^*$.

Definition 6.7.5. Die zum Endomorphismus $f: V \rightarrow V$ adjungierte Abbildung $f^*: V \rightarrow V$ ist durch

$$\langle f(\xi), \eta \rangle = \langle \xi, f^*(\eta) \rangle \quad \text{für alle } \xi, \eta \in V$$

definiert.

Beispiele 6.7.6.

(i) Sei $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ durch

$$f(x^1, x^2, x^3) := (x^1 - x^2, -x^1 + x^2 + 2x^3, x^2 + x^3)$$

gegeben. Nach Definition gilt $\langle f(x), y \rangle = \langle x, g(y) \rangle$. Somit erhalten wir aus

$$\begin{aligned}\langle f(x), y \rangle &= (x^1 - x^2)y^1 + (-x^1 + x^2 + 2x^3)y^2 + (x^2 + x^3)y^3 \\ &= x^1(y^1 - y^2) + x^2(-y^1 + y^2 + y^3) + x^3(2y^2 + y^3) \\ &= \langle x, f^*(y) \rangle\end{aligned}$$

die adjungierte Abbildung

$$f^*(y^1, y^2, y^3) := (y^1 - y^2, -y^1 + y^2 + y^3, 2y^2 + y^3).$$

Die darstellenden Matrizen bezüglich der Standardbasis sind

$$A_f = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{sowie} \quad A_{f^*} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix}.$$

(ii) Im Raum $V = L^2([a, b])$ der auf $[a, b]$ stetigen reellwertigen Funktionen mit L^2 -Skalarprodukt ist

$$\Phi: V \rightarrow V,$$

$$\Phi(f)(x) := x \cdot f(x)$$

selbstadjungiert, d. h. es gilt $\Phi^* = \Phi$, denn es gilt

$$\langle \Phi(f), g \rangle = \langle f, \Phi(g) \rangle.$$

Den Zusammenhang zwischen der f und der f^* darstellenden Matrix haben wir bereits im Beispiel gesehen.

Theorem 6.7.7. Sei V ein endlichdimensionaler Skalarproduktraum. Sei S eine Orthonormalbasis. Wird $f: V \rightarrow V$ bezüglich S durch die Matrix A dargestellt, so wird die adjungierte Abbildung $f^*: V \rightarrow V$ bezüglich S durch die Matrix A^* dargestellt.

Beweis. Seien $S = (d_1, \dots, d_n)$, $A = (a_j^i)_{1 \leq i, j \leq n}$ und $A_{f^*} = (b_j^i)_{1 \leq i, j \leq n}$. Wir erhalten

$$\begin{aligned}f(d_k) &= \sum_{i=1}^n a_k^i d_i, \\ f^*(d_l) &= \sum_{j=1}^n b_l^j d_j,\end{aligned}$$

$$\begin{aligned}\langle f(d_k), d_l \rangle &= \langle d_k, f^*(d_l) \rangle, \\ \langle f(d_k), d_l \rangle &= \left\langle \sum_{i=1}^n a_k^i d_i, d_l \right\rangle = \sum_{i=1}^n a_k^i \langle d_i, d_l \rangle = \sum_{i=1}^n a_k^i \delta_{il} = a_k^l, \\ \langle d_k, f^*(d_l) \rangle &= \left\langle d_k, \sum_{j=1}^n \overline{b_l^j} d_j \right\rangle = \sum_{j=1}^n \overline{b_l^j} \delta_{kj} = \overline{b_l^k}.\end{aligned}$$

(In den letzten beiden Zeilen sind die jeweils letzten Gleichheiten aus Kovarianzgründen unschön geschrieben. Für die zugehörigen reellen bzw. komplexen Zahlen gilt die Gleichheit jedoch.) Also ist $a_k^l = \overline{b_l^k}$ für alle $1 \leq k, l \leq n$. Die Behauptung folgt. \square

Theorem 6.7.8. Sei V ein endlichdimensionaler Skalarproduktraum. Sei $f: V \rightarrow V$ linear. Dann gilt

$$\ker f^* = (\operatorname{im} f)^\perp.$$

Beweis. Es gilt

$$\begin{aligned}\ker f^* &= \{\xi \in V : f^*(\xi) = 0\} \\ &= \{\xi \in V : \langle \eta, f^*(\xi) \rangle = 0 \text{ für alle } \eta \in V\} \\ &= \{\xi \in V : \langle f(\eta), \xi \rangle = 0 \text{ für alle } \eta \in V\} \\ &= (\operatorname{im} f)^\perp.\end{aligned}$$

\square

Beispiel 6.7.9. Sei

$$\sum_{k=1}^n a_k^i x^k = b^i, \quad 1 \leq i \leq n$$

ein reelles lineares quadratisches Gleichungssystem. Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ die (bezüglich der Standardbasis) durch $A = (a_k^i)$ dargestellte Abbildung. Dann ist das lineare Gleichungssystem genau dann lösbar, wenn $b = (b^1, \dots, b^n) \in \operatorname{im} f$ gilt. Auf \mathbb{R}^n führen wir das Standardskalarprodukt ein und erhalten aus Theorem 6.7.8, dass das lineare Gleichungssystem genau dann lösbar ist, wenn $b \in (\ker f^*)^\perp$ gilt. Über \mathbb{R} wird f^* durch die transponierte Matrix dargestellt. Somit folgt: Das lineare Gleichungssystem ist genau dann lösbar, wenn für jede Lösung $y = (y_1, \dots, y_n)$ von

$$\sum_{i=1}^n a_k^i y_i = 0 \quad \text{auch} \quad \sum_{i=1}^n y_i b^i = 0 \quad \text{gilt.}$$

Theorem 6.7.10. Sei V ein endlichdimensionaler Skalarproduktraum. Sei $f: V \rightarrow V$ ein Endomorphismus. Dann haben f und f^* den gleichen Rang.

Beweis. Es gilt

$$\begin{aligned}\operatorname{rang} f &= \dim \operatorname{im} f = \dim V - \dim(\operatorname{im} f)^\perp \\ &= \dim V - \dim(\ker f^*) && \text{(Theorem 6.7.8)} \\ &= \dim(\operatorname{im} f^*) = \operatorname{rang} f^*.\end{aligned}$$

(Alternativ kann man über \mathbb{R} wegen $A^* = A^T$ benutzen, dass Zeilenrang und Spaltenrang übereinstimmen.) \square

Korollar 6.7.11. Sei $A \in \mathbb{C}^{n \times n}$. Dann gilt $\operatorname{rang} A = \operatorname{rang} \overline{A}$.

Beweis. Setze $B := \overline{A}^T$. Da der Rang einer Matrix B mit dem von B^T übereinstimmt, folgt

$$\operatorname{rang} A = \operatorname{rang} B^* = \operatorname{rang} B = \operatorname{rang} B^T = \operatorname{rang} \overline{A}.$$

\square

6.8. Diagonalisierung von selbstadjungierten linearen Endomorphismen.

Wir wollen insbesondere zeigen, dass sich solche Endomorphismen sogar mit Hilfe einer Orthonormalbasis stets diagonalisieren lassen.

Definition 6.8.1. Sei V ein Skalarproduktraum. Der Endomorphismus $f: V \rightarrow V$ heißt selbstadjungiert, wenn

$$\langle f(\xi), \eta \rangle = \langle \xi, f(\eta) \rangle$$

für alle $\xi, \eta \in V$ gilt, also $f = f^*$ ist.

Theorem 6.8.2. Sei V ein endlichdimensionaler Skalarproduktraum mit Orthonormalbasis S . Der Endomorphismus $f: V \rightarrow V$ werde bezüglich S durch die Matrix A beschrieben. Dann ist $f: V \rightarrow V$ genau dann selbstadjungiert, wenn $A = A^*$ gilt.

Beweis. Dies folgt direkt aus Theorem 6.7.7. \square

Theorem 6.8.3. Sei V ein n -dimensionaler Skalarproduktraum. Sei $f: V \rightarrow V$ ein selbstadjungierter Endomorphismus. Dann sind alle Nullstellen des charakteristischen Polynoms $\chi_f(X)$ reell, es gilt also

$$\chi_f(X) = \pm(X - \lambda_1) \cdots (X - \lambda_n)$$

für $\lambda_1, \dots, \lambda_n \in \mathbb{R}$.

Beweis. Wir betrachten zunächst den unitären Fall. Wie jedes Polynom zerfällt das charakteristische Polynom in Linearfaktoren und wir müssen lediglich nachweisen, dass die Eigenwerte von f reell sind. Sei also λ ein Eigenwert von f zum Eigenvektor ξ . Wir erhalten

$$\lambda \langle \xi, \xi \rangle = \langle \lambda \xi, \xi \rangle = \langle f(\xi), \xi \rangle = \langle \xi, f(\xi) \rangle = \langle \xi, \lambda \xi \rangle = \bar{\lambda} \langle \xi, \xi \rangle.$$

Wegen $\xi \neq 0$ bzw. $\langle \xi, \xi \rangle \neq 0$ folgt $\lambda = \bar{\lambda}$. Somit ist λ reell.

Im reellen Fall benutzen wir die sogenannte Komplexifizierung. Die zu $f: V \rightarrow V$ gehörige reelle symmetrische Matrix definiert eine Abbildung $\tilde{f}: \mathbb{C}^n \rightarrow \mathbb{C}^n$. Eine reelle symmetrische Matrix ist hermitesch, wenn wir sie als komplexe Matrix auffassen. Somit ist $\tilde{f}: \mathbb{C} \rightarrow \mathbb{C}$ bezüglich des Standardskalarproduktes auf \mathbb{C}^n selbstadjungiert. Es gilt

$$\chi_f(X) = \chi_A(X) = \chi_{\tilde{f}}(X).$$

Somit stimmen die Eigenwerte von f und \tilde{f} überein und sind aufgrund der obigen Überlegungen reell. \square

Theorem 6.8.4. Sei V ein endlichdimensionaler Skalarproduktraum. Sei $f: V \rightarrow V$ ein selbstadjungierter Endomorphismus. Dann besitzt V eine Orthonormalbasis aus Eigenvektoren von f .

Beweis. Wir wissen, dass das charakteristische Polynom $\chi_f(X)$ in Linearfaktoren zerfällt. (Im Fall $F = \mathbb{C}$ folgt dies aus dem Fundamentalsatz der Algebra und für $F = \mathbb{R}$ (und \mathbb{C}) aus Theorem 6.8.3.)

Benutze Induktion nach $\dim V =: n$. Für $n = 1$ ist die Aussage wahr. Sei also $n \geq 2$. Sei a_1 ein Eigenvektor zum Eigenwert λ mit $\|a_1\| = 1$. Definiere $W := \{\xi \in V: \langle a_1, \xi \rangle = 0\}$. Nach Theorem 6.6.4 folgt $\dim W = n - 1$. W ist ein bezüglich f invarianter Unterraum, es gilt nämlich für $\xi \in W$

$$\langle a_1, f(\xi) \rangle = \langle f(a_1), \xi \rangle = \langle \lambda a_1, \xi \rangle = \lambda \langle a_1, \xi \rangle = 0.$$

Es folgt $f(\xi) \in W$. Nach Induktionsvoraussetzung existiert eine Orthonormalbasis $\{a_2, \dots, a_n\}$ von W aus Eigenvektoren von $f|_W$ und somit auch von f . Also ist $\{a_1, a_2, \dots, a_n\}$ wie behauptet eine Orthonormalbasis. \square

Theorem 6.8.5. *Jeder selbstadjungierte Endomorphismus eines endlichdimensionalen Skalarproduktraumes ist diagonalisierbar.*

Beweis. Benutze eine Orthonormalbasis aus Eigenvektoren. Bezüglich dieser ist die zugehörige Matrix eine Diagonalmatrix mit den Eigenwerten entsprechend ihrer Vielfachheit auf der Diagonalen. \square

Theorem 6.8.6. *Sei A eine reelle symmetrische (bzw. komplexe hermitesche) $(n \times n)$ -Matrix. Dann gibt es eine orthogonale (bzw. unitäre) Matrix D mit*

$$DAD^* = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

mit $\lambda_1, \dots, \lambda_n \in \mathbb{R}$.

Beweis. Dies folgt direkt aus den bisherigen Theoremen dieses Kapitels, da für orthogonale (bzw. unitäre) Matrizen $D^{-1} = D^*$ gilt. \square

6.9. Kästchenform orthogonaler Matrizen. Wir wollen unitäre und orthogonale Endomorphismen bezüglich Orthonormalbasen durch „einfache“ Matrizen darstellen.

Theorem 6.9.1. *Sei $f: V \rightarrow V$ ein unitärer Endomorphismus eines endlichdimensionalen unitären Vektorraumes V . Dann gibt es eine Orthonormalbasis von V , die aus Eigenvektoren von f besteht.*

Beweis. Wir beweisen dies per Induktion nach $n = \dim V$. Der Induktionsanfang ist trivial. Sei $n \geq 2$. Das charakteristische Polynom zerfällt über \mathbb{C} in Linearfaktoren

$$\chi_f(X) = \pm(X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n).$$

Nach Theorem 6.5.2 gilt $|\lambda_i| = 1$. Sei a_1 ein Eigenvektor zum Eigenwert λ_1 mit $\|a_1\| = 1$. Sei W das orthogonale Komplement von a_1 , $V = W \oplus \langle a_1 \rangle$. Für jedes $\xi \in W$ gilt

$$0 = \langle a_1, \xi \rangle = \langle f(a_1), f(\xi) \rangle = \lambda_1 \langle a_1, f(\xi) \rangle.$$

Da $\lambda_1 \neq 0$ gilt, folgt $f(\xi) \in W$. Somit induziert die unitäre Abbildung f eine unitäre Abbildung $f|_W: W \rightarrow W$. Nach Induktionsvoraussetzung gibt es für $f|_W$ eine Orthonormalbasis aus Eigenvektoren. Zusammen mit a_1 erhält man wie gewünscht eine Orthonormalbasis aus Eigenvektoren. \square

Korollar 6.9.2. *Sei $A \in \mathbb{C}^{n \times n}$ unitär. Dann gibt es eine unitäre Matrix D mit*

$$DAD^* = \begin{pmatrix} e^{i\varphi_1} & 0 & \dots & 0 \\ 0 & e^{i\varphi_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{i\varphi_n} \end{pmatrix} \quad \text{und} \quad 0 \leq \varphi_i < 2\pi.$$

Lemma 6.9.3. *Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ orthogonal. Fixiere eine Orthonormalbasis. Dann gibt es $0 \leq \varphi < 2\pi$, so dass f bezüglich dieser Basis durch*

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix}$$

dargestellt wird.

Beweis. Wir benutzen Theorem 6.4.11. Die Spalten und die Zeilen der Matrix bilden also eine Orthonormalbasis. Einen beliebigen Einheitsvektor in \mathbb{R}^2 können wir als

$$\left(\cos \varphi, \pm \sqrt{1 - \cos^2 \varphi} \right) = (\cos \varphi, \pm \sin \varphi)$$

schreiben. Somit ist f durch eine Matrix der Form

$$\begin{pmatrix} \cos \varphi & \pm \sin \varphi \\ \pm \sin \varphi & \pm \cos \varphi \end{pmatrix}$$

dargestellt. Man überzeugt sich leicht, dass die Vorzeichen für $\sin \varphi \neq 0$ bzw. $\cos \varphi \neq 0$ wie angegeben gewählt werden müssen um eine orthogonale Matrix zu erhalten. Ggf. ist φ durch $-\varphi + 2\pi k$, $k \in \mathbb{Z}$, zu ersetzen um Vorzeichen abzuändern. \square

Durch eine spezielle Basiswahl können wir Kästchen der zweiten Form ausschließen.

Theorem 6.9.4. *Sei $f: V \rightarrow V$ eine orthogonale Selbstabbildung eines euklidischen Vektorraumes mit $\dim V < \infty$. Dann gibt es eine Orthonormalbasis von V , in der sich f in der Form*

$$\begin{pmatrix} \mathbf{1}_k & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & -\mathbf{1}_l & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \cos \varphi_1 & -\sin \varphi_1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \sin \varphi_1 & \cos \varphi_1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \cos \varphi_m & -\sin \varphi_m \\ & & & & & & \sin \varphi_m & \cos \varphi_m \end{pmatrix}$$

mit $k + l + 2m = \dim V$ darstellen lässt.

(In „Matrizensprache“ übersetzt erhalten wir: Zu jeder orthogonalen Matrix A gibt es also eine orthogonale Matrix D , so dass DAD^T von der angegebenen Gestalt ist.)

Beweis. Vermöge einer Matrixdarstellung $A \in \mathbb{R}^{\dim V \times \dim V}$ ordnen wir f wiederum eine Abbildung $\tilde{f}: \mathbb{C}^n \rightarrow \mathbb{C}^n$ zu. Wegen $A^*A = A^T A = \mathbf{1}$ ist A als Matrix in $\mathbb{C}^{n \times n}$ unitär. Nach Theorem 6.9.1 gibt es eine Orthonormalbasis aus Eigenvektoren für \tilde{f} zu Eigenwerten λ_i mit $|\lambda_i| = 1$. Zu jedem λ_i mit $\lambda_i = \pm 1$ finden wir einen reellen Eigenvektor zu diesem Eigenwert, den Realteil des komplexen Eigenvektors: Aus $A\xi = \lambda_i \xi$ folgt $A\bar{\xi} = \lambda_i \bar{\xi}$ und $A(\xi + \bar{\xi}) = \lambda_i(\xi + \bar{\xi})$. Daher ist $\operatorname{Re} \xi$ ein Eigenvektor oder der Nullvektor. Ist $\operatorname{Re} \xi$ der Nullvektor, so ist $i\xi$ ein reeller Eigenvektor.

Da das orthogonale Komplement eines Eigenvektors jeweils invariant unter f ist, erhalten wir induktiv die beiden Blöcke $\mathbf{1}_k$ und $\mathbf{1}_l$.

Sei ξ ein Eigenvektor zum Eigenwert $e^{i\varphi} \notin \{-1, 1\}$, $A\xi = e^{i\varphi}\xi$. Durch komplexe Konjugation erhalten wir einen zweiten Eigenvektor: $A\bar{\xi} = e^{-i\varphi}\bar{\xi}$. Setze $u := \xi + \bar{\xi}$ und $v := -i(\xi - \bar{\xi})$. Da ξ und $\bar{\xi}$ wegen $e^{i\varphi} \notin \{-1, 1\}$ Eigenvektoren von A zu unterschiedlichen Eigenwerten sind, gilt $\xi \neq \pm \bar{\xi}$ und somit $u \neq 0 \neq v$. Wir erhalten

$$\begin{aligned} Au &= A\xi + A\bar{\xi} = e^{i\varphi}\xi + e^{-i\varphi}\bar{\xi} = \frac{e^{i\varphi} + e^{-i\varphi}}{2}(\xi + \bar{\xi}) + \frac{e^{i\varphi} - e^{-i\varphi}}{2}(\xi - \bar{\xi}) \\ &= \cos \varphi \cdot u - \sin \varphi \cdot v, \\ -\frac{1}{i}Av &= A\xi - A\bar{\xi} = e^{i\varphi}\xi - e^{-i\varphi}\bar{\xi} = \frac{e^{i\varphi} - e^{-i\varphi}}{2}(\xi + \bar{\xi}) + \frac{e^{i\varphi} + e^{-i\varphi}}{2}(\xi - \bar{\xi}) \\ &= i \sin \varphi \cdot u + \cos \varphi \cdot \frac{v}{-i}, \\ Av &= \sin \varphi \cdot u + \cos \varphi \cdot v. \end{aligned}$$

Wir wählen $\frac{1}{\sqrt{2}}u$ und $\frac{1}{\sqrt{2}}v$ als reellen Teil einer Orthonormalbasis. Beachte dazu, dass aus $\langle \xi, \bar{\xi} \rangle_{\mathbb{C}} = 0$ die Relationen $\|\xi + \bar{\xi}\|^2 = 2$, $\|-i(\xi - \bar{\xi})\|^2 = 2$ und $\langle \xi + \bar{\xi}, -i(\xi - \bar{\xi}) \rangle = 0$ folgen, zunächst mit komplexen Skalarprodukten und Normen. Da die Einträge aber reell sind, gelten dieselben Beziehungen auch über \mathbb{R} und zugehörigen Skalarprodukten und Normen. Dies liefert den ersten Block:

$$\begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} = \begin{pmatrix} \cos \varphi_1 & -\sin \varphi_1 \\ \sin \varphi_1 & \cos \varphi_1 \end{pmatrix}, \quad \varphi_1 = -\varphi.$$

Da beide Zeilen und Spalten, die durch diesen Block verlaufen, bereits „in diesem Block Länge Eins haben“, enthalten sie sonst nur Nullen. Die restlichen Blöcke erhält man per Induktion. \square

ANHANG A. MENGENLEHRE UND LOGIK

Dies ist ein Minimalstüberblick. Wir benutzen [2].

A.1. **Logik.** Wir verwenden die logischen Verknüpfungen wie in der folgenden Wahrheitstabelle angegeben (w = wahr, f = falsch)

a	b	$\neg a$	$a \vee b$	$a \wedge b$	$a \implies b$	$a \iff b$	$a \longleftarrow b$
w	w	f	w	w	w	w	w
w	f	f	w	f	f	f	w
f	w	w	w	f	w	f	f
f	f	w	f	f	w	w	w

wobei

$\neg a$	nicht a
$a \vee b$	a oder b
$a \wedge b$	a und b , sowohl a als auch b
$a \implies b$	a impliziert b
$a \iff b$	a und b sind äquivalent
$a \longleftarrow b$	a folgt aus b

bedeutet und a und b für Aussagen stehen, die wahr oder falsch sein können. a ist eine hinreichende Bedingung für b , wenn $a \implies b$ gilt. a ist eine notwendige Bedingung für b , wenn $a \longleftarrow b$ gilt.

A.2. **Mengenlehre.**

- (i) Es gibt eine Menge, die leere Menge \emptyset , die kein Element enthält.
- (ii) Eine Menge A besteht aus Elementen, etwa x, y und z , $A = \{x, y, z\}$. x ist ein Element von A , $x \in A$; $\{x, y\}$ ist eine Teilmenge von A , $\{x, y\} \subset A$, sogar eine echte Teilmenge, $\{x, y\} \subsetneq A$, da $z \in A$ aber $z \notin \{x, y\}$ gelten. Allgemein definieren wir $(A \subset B) \iff ((x \in A) \implies (x \in B))$.
- (iii) Zwei Mengen sind genau dann gleich, wenn sie dieselben Elemente enthalten. $A = B \iff ((A \subset B) \wedge (B \subset A))$.

Zur Konstruktion neuer Mengen

- (iv) $\{x \in A : f(x)\}$ definiert eine neue Menge, die genau aus den $x \in A$ besteht, für die die Aussage $f(x)$ wahr ist.
- (v) Seien A, B Mengen. Dann gibt es die Menge $\{A, B\}$.
- (vi) Sei \mathcal{M} eine Menge von Mengen. Dann gibt es eine Menge, die genau alle Elemente von \mathcal{M} als Teilmengen enthält, $\bigcup \mathcal{M}$. Wichtige Spezialfälle sind: Sei $(A_i)_{i \in I}$ eine beliebige Familie von Mengen. Dann gibt es eine Menge, die genau die Elemente enthält, die in wenigstens einem A_i , $i \in I$, enthalten sind,

die Vereinigung der Mengen A_i , $\bigcup_{i \in I} A_i$. Für zwei Mengen A und B schreiben wir $A \cup B$.

- (vii) Sei $(A_i)_{i \in I}$ eine beliebige Familie von Mengen. Dann gibt es eine Menge, die genau die Elemente enthält, die in allen Mengen A_i , $i \in I$, enthalten sind, den Durchschnitt der Mengen A_i , $\bigcap_{i \in I} A_i$. Für zwei Mengen A und B schreiben wir $A \cap B$.
- (viii) $\{x \in A : x \notin B\}$ heißt Komplement von B in A , $\complement B$.
- (ix) Zu jeder Menge A gibt es eine Potenzmenge $\mathcal{P}(A)$, die alle Teilmengen von A als Elemente enthält.
- (x) Es gibt die natürlichen Zahlen $\mathbb{N}: 0, 1, 2, \dots$, die kleinste Menge, die 0 und zu jeder Zahl i den Nachfolger i^+ enthält, so dass jede Zahl Nachfolger maximal einer Zahl und Null nicht Nachfolger einer Zahl ist. Dabei verwenden wir die Abkürzungen $0 := \emptyset$, $1 := 0^+ := \{0\} = \{\emptyset\}$, $2 := 1^+ = \{0, 1\} = \{\emptyset, \{\emptyset\}\}$, $3 := 2^+ = \{0, 1, 2\}$, \dots . Diese Definition ermöglicht die vollständige Induktion: Ist eine Aussage für 0 wahr und mit i auch für den Nachfolger $i^+ \equiv i + 1$, so gilt sie für alle natürlichen Zahlen. Weiterhin können wir Funktionen rekursiv definieren: Sei $f: X \rightarrow X$ eine Funktion und $x \in X$. Dann gibt es eine Funktion $u: \mathbb{N} \rightarrow X$ mit $u(0) = x$ und $u(n+1) = f(u(n))$ für alle $n \in \mathbb{N}$.
- (xi) Zu einer Familie von Mengen $(A_i)_{i \in I}$ gibt es ein kartesisches Produkt dieser Mengen, das wir im Fall von zwei Mengen A, B mit $A \times B$ und im Fall, dass $i = \mathbb{N}$ ist, mit $A_0 \times A_1 \times A_2 \times \dots$ bezeichnen. Die Elemente sind geordnete Paare oder Tupel. Wir bezeichnen sie mit (a, b) bzw. (a_0, a_1, a_2, \dots) mit $a \in A$, $b \in B$ sowie $a_i \in A_i$, $i \in \mathbb{N}$.
Vermöge $(a, b) \equiv \{\{1, a\}, \{2, b\}\}, \dots$ kann man die Existenz geordneter Tupel auf Mengen zurückführen.
- (xii) Das Auswahlaxiom besagt, dass das kartesische Produkt einer nichtleeren Familie von nichtleeren Mengen wiederum nicht leer ist. Es ist äquivalent zum Zornschen Lemma, das wir benutzen, um die Existenz von Basen zu zeigen.
- (xiii) Eine Relation R auf X ist eine Teilmenge von $X \times X$. Die Relation xRy ist genau dann erfüllt, wenn $(x, y) \in R$ ist.
- (xiv) Eine Funktion $f: X \rightarrow Y$ ist ein Tripel $(X, Y, \text{graph } f)$. Der Graph von f , $\text{graph } f$, ist eine Teilmenge von $X \times Y$, so dass für jedes $x \in X$ genau ein $y \in Y$ existiert, so dass (x, y) in dieser Teilmenge ist, $f(x) = y$.
- (xv) Wie haben vorher Familien verwendet: Diese führt man wie folgt ein: Zunächst erklärt man das kartesische Produkt für zwei Mengen und kann damit Funktionen erklären. Sei $x: I \rightarrow X$ eine Funktion. Dann schreiben wir auch $(x_i)_{i \in I}$ mit $x_i \equiv x(i)$ und nennen dies eine Familie. Ist $X = \mathcal{P}(A)$ eine Potenzmenge, so spricht man auch von einer Mengenfamilie von Teilmengen von A .
- (xvi) Zu einer Funktion $f: X \rightarrow Y$ gibt es eine Funktion $f^{-1}: \mathcal{P}(Y) \rightarrow \mathcal{P}(X)$ mit

$$f^{-1}(B) = \{x \in X : f(x) \in B\}$$

für $B \subset Y$.

Zur Mächtigkeit/Kardinalität:

- (xvii) Eine Abbildung $f: A \rightarrow B$ heißt injektiv, falls aus $f(x) = f(y)$ bereits $x = y$ folgt. Sie heißt surjektiv, falls es zu jedem $z \in B$ ein $x \in A$ gibt, so dass $f(x) = z$ gilt. Ist f injektiv und surjektiv, so heißt f bijektiv.
- (xviii) Eine Menge A ist weniger mächtig als eine Menge B (oder die Kardinalität von A ist kleiner als die von B), wenn es eine injektive Abbildung $A \rightarrow B$ gibt. (Beachte: Weniger mächtig heißt nicht notwendigerweise strikt weniger mächtig.) Eine Menge A heißt genauso mächtig wie eine Menge B , wenn es eine bijektive Abbildung $A \rightarrow B$ gibt.

- (xix) Der (nichttriviale) Satz von Schröder-Bernstein besagt: Ist A weniger mächtig als B und B weniger mächtig als A , so sind A und B gleich mächtig.
- (xx) Eine Menge heißt endlich, wenn sie gleich mächtig wie eine natürliche Zahl ist. Eine Menge heißt unendlich, wenn sie nicht endlich ist oder, äquivalent dazu, wenn sie eine echte Teilmenge besitzt, die gleich mächtig wie die Menge selbst ist. Eine Menge heißt abzählbar, wenn sie gleich mächtig wie \mathbb{N} ist. Eine Menge heißt höchstens abzählbar, wenn sie gleich mächtig wie eine Teilmenge von \mathbb{N} ist. Eine Menge heißt überabzählbar, wenn sie strikt mächtiger (= mächtiger, aber nicht gleich mächtig) als \mathbb{N} ist.
- (xxi) $\mathbb{N} \times \mathbb{N}$ und \mathbb{N} sind gleich mächtig.
- (xxii) Sei A eine Menge. Dann ist $\mathcal{P}(A)$ strikt mächtiger als A .

LITERATUR

1. Gerd Fischer, *Lineare Algebra*, fifth ed., Grundkurs Mathematik, vol. 17, Friedr. Vieweg & Sohn, Braunschweig, 1979.
2. Paul R. Halmos, *Naive Mengenlehre*, Vandenhoeck & Ruprecht, Göttingen, 1976.
3. Serge Lang, *Linear algebra*, third ed., Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1989.
4. Falko Lorenz, *Lineare Algebra. I*, third ed., Bibliographisches Institut, Mannheim, 1992.
5. Urs Stammbach, *Lineare Algebra*, Teubner Studienskripten, vol. 82, B. G. Teubner, Stuttgart, 1980.

OLIVER C. SCHNÜRER, FACHBEREICH MATHEMATIK UND STATISTIK, UNIVERSITÄT KONSTANZ,
78457 KONSTANZ, GERMANY

E-mail address: `Oliver.Schnuerer@uni-konstanz.de`