

LINEARE ALGEBRA

SARAH ZERBES

Varia. Kurs Webseite: <https://metaphor.ethz.ch/x/2024/hs/401-1151-00L/>

Empfohlene Buecher und Skripte: siehe Webseite

Fragen und Korrekturen zum Skript: <https://forum.math.ethz.ch/c/autumn-24/lineare-algebra-mavt-matl/189>

Haupt-Assistent fuer LA1: Tim Gehringer (tim.gehringer@math.ethz.ch)

Diese Skript basiert zum Teil auf dem Skript von Menny Akka Genosar (siehe <https://metaphor.ethz.ch/x/2022/hs/401-1151-00L/>)

Das Skript wird nach jeder Vorlesung auf die Webseite hochgeladen.

CONTENTS

Varia	1
1. Einfuehrung	3
1.1. Motivation: Fibonacci Folgen	3
1.2. Mengenlehre	6
1.3. Funktionen	8
2. Matrizen und lineare Gleichungssysteme	10
2.1. Körper	10
2.2. Matrizen	12
2.3. Elementare Zeilenoperationen	15
2.4. Lineare Gleichungssysteme	18
3. Vektorraeume	21
3.1. Definition und Beispiele	21
3.2. Unterräume	23
3.3. Basen von Vektorraeumen	26
3.4. Basen von Unterraemen	32
3.5. Zusammenfassung der Rechenmethoden	35
3.6. Zeilen und Spaltenraeume	36
4. Lineare Abbildungen	38
4.1. Definition und Beispiele	38
4.2. Kernel and Image	41
4.3. Lineare Abbildungen als Matrizen	45
4.4. Matrizen als Lineare Abbildungen	47
4.5. Basiswechsel	48
4.6. Eine Bemerkung zu Koordinaten	51
4.7. Zeilenrang gleich Spaltenrang	52
4.8. Zurueck zu linearen Gleichungssystemen	55
5. Intermezzo: Gruppen und Ringe	58
5.1. Gruppen	58
5.2. Ringe	60
6. Vektorraeume linearer Abbildungen	61
6.1. Definition und erste Eigenschaften	61
6.2. Der duale Vektorraum	63
6.3. Die duale Abbildung	66
6.4. Annulatoren	68
6.5. Reflexivitaet	71
7. Quotientenraeume	73

7.1.	Definition und erste Eigenschaften	73
7.2.	Die Isomorphismensätze	75
7.3.	Weitere Anwendungen	77
8.	Determinanten	78
8.1.	Ein erstes Beispiel	78
8.2.	Permutationen	78
8.3.	Determinantenfunktionen	80
8.4.	Erste Eigenschaften	84
8.5.	Determinanten und Invertierbarkeit	86
8.6.	Die Determinante eines Endomorphismus	90
9.	Polynome	91
10.	Eigenwerte und Eigenvektoren	92
10.1.	Definitionen und erste Eigenschaften	92
10.2.	Das charakteristische Polynom	94
10.3.	Diagonalisierung	96
10.4.	Eigenräume	98
10.5.	Algebraische und geometrische Vielfachheit	100
11.	Das minimale Polynom	101
11.1.	Definition und erste Eigenschaften	101
11.2.	Der Satz von Cayley–Hamilton	104
11.3.	Ein alternativer Beweis von Cayley–Hamilton	106
12.	Die Jordansche Normalenform einer Matrix	107
12.1.	Definition und Theorem	107
12.2.	Eigenschaften der Jordanschen Normalenform	108
12.3.	Verallgemeinerte Eigenräume	110
12.4.	Beweis der JNF	112
12.5.	Berechnung der Jordanschen Normalenform	115
13.	Euklidische und Hermitesche Räume	118
13.1.	Normierte Räume	118
13.2.	Innere Produkte	119
13.3.	Konstruktion innerer Produkte	122
13.4.	Orthogonalität	123
13.5.	Gram-Schmidt Orthogonalisierung	125
13.6.	Das orthogonale Komplement	127
13.7.	QR Zerlegung	130
13.8.	Dualräume von Inneren Produkt Räumen	132
13.9.	Die adjungierte Abbildung	133
13.10.	Die Abbildungsmatrix der adjungierten Abbildung	135
14.	Spektraltheorie	136
14.1.	Normale Endomorphismen	136
14.2.	Spektraltheorie über $\mathbb{R}$	139
15.	Isometrien	140
15.1.	Definition und erste Eigenschaften	140
15.2.	Klassifikation der Elemente in $O(2)$ und in $SO(3)$	142
15.3.	Singulärwertzerlegung	144
16.	Direkte Summen und Tensorprodukte von Vektorräumen	146
16.1.	Die äussere direkte Summe zweier Vektorräume	146
16.2.	Vektorräume über einer freien Menge	146
16.3.	Konstruktion des Tensorprodukts	147
16.4.	Erste Eigenschaften	150
16.5.	Beispiel: Komplexifizierung	151
16.6.	Tensorprodukte linearer Abbildungen	151
16.7.	Tensorprodukte und duale Abbildungen	152
16.8.	Darstellung von Tensoren (für Physiker)	154
16.9.	Das symmetrische und alternierende Produkt	154

1.1. Motivation: Fibonacci Folgen. In diesem Kapitel moechte ich Ihnen eine sehr schoene Anwendung der linearen Algebra zeigen, als Motivation fuer die Strukturen, die wie in diesem Kurs untersucht werden.



Machen Sie sich keine Sorgen, wenn Sie in diesem Kapitel nicht alles verstehen; alle diese Strukturen werden wir in den naechsten Monaten genau untersuchen und dabei immer wieder auf dieses Beispiel zurueckkommen.

Definition 1.1.1. Die Fibonacci Folge $(a_n)_{n \geq 0}$ ist definiert durch die Rekursion

$$\begin{aligned} a_0 &= 0, & a_1 &= 1, \\ a_n &= a_{n-1} + a_{n-2} & \text{for } n \geq 2 \end{aligned}$$

Beachte 1.1.2. Die ersten Elemente der Folge sind

$$0, 1, 1, 2, 3, 5, 8, 13, \dots$$

Frage: Koennen wir a_n bestimmen, ohne vorher alle Elemente bis a_{n-1} zu bestimmen?

Definition 1.1.3. Seien $a, b \in \mathbb{R}$. Definiere die Folge $\mathcal{F}_{a,b}$ mittels der Rekursion

$$\begin{aligned} a_0 &= a, & a_1 &= b, \\ a_n &= a_{n-1} + a_{n-2} & \text{for } n \geq 2 \end{aligned}$$

Eine Folge $\mathcal{A}$ ist eine Fibonacci Folge wenn $\mathcal{A} = \mathcal{F}_{a,b}$ für $a, b \in \mathbb{R}$. Es sei V die Menge aller solcher Folgen, d.h.

$$V = \{\mathcal{F}_{a,b} : a, b \in \mathbb{R}\}$$

Beispiel 1.1.4. (1) Die Folge aus Definition 1.1.1 ist gleich der Folge $\mathcal{F}_{0,1}$.

(2) Fuer $a = -2$ und $b = 2$ erhalten wir die Folge

$$\mathcal{F}_{-2,2} = (-2, 2, 0, 2, 2, 4, 6, 10, \dots)$$

(3) Fuer $a = -1$ und $b = \frac{1}{2}$ erhalten wir

$$\mathcal{F}_{-1, \frac{1}{2}} = \left(-1, \frac{1}{2}, -\frac{1}{2}, 0, -\frac{1}{2}, -\frac{1}{2}, -1, -\frac{3}{2}, -\frac{5}{2}, -4, \dots\right)$$

Definition 1.1.5. Es seien $\mathcal{F} = (a_n)_{n \geq 0}$, $\mathcal{G} = (b_n)_{n \geq 0} \in V$ und $\alpha \in \mathbb{R}$.

(1) Wir definieren die Summe von $\mathcal{F}$ und $\mathcal{G}$ als die Folge

$$\mathcal{F} + \mathcal{G} = (a_n + b_n)_{n \geq 0}.$$

(2) Wir definieren das Skalarprodukt von $\mathcal{F}$ und α als die Folge

$$\alpha \cdot \mathcal{F} = (\alpha \cdot a_n)_{n \geq 0}.$$

Satz 1.1.6.

(1) Es seien $\mathcal{F}, \mathcal{G} \in V$. Dann gilt $\mathcal{F} + \mathcal{G} \in V$.

(2) Es sei $\mathcal{F} \in V$ und $\alpha \in \mathbb{R}$. Dann gilt $\alpha \cdot \mathcal{F} \in V$.

Wir sagen: V hat die Struktur eines Vektorraums über $\mathbb{R}$.

Proof. (1) Wir schreiben $\mathcal{F} = (a_n)_{n \geq 0}$ und $\mathcal{G} = (b_n)_{n \geq 0}$, und es sei $c_n = a_n + b_n$. Wir muessen zeigen, dass

$$c_n = c_{n-1} + c_{n-2}$$

fuer all $n \geq 2$. Aber

$$\begin{aligned} c_n &= a_n + b_n \\ &= a_{n-1} + a_{n-2} + b_{n-1} + b_{n-2} \\ &= (a_{n-1} + b_{n-1}) + (a_{n-2} + b_{n-2}) \\ &= c_{n-1} + c_{n-2}. \end{aligned}$$

(2) kann durch ähnliche Argumentation bewiesen werden. □

Beachte 1.1.7. Der Beweis von Satz 1.1.6 zeigt, dass folgendes gilt:

$$\begin{aligned} \mathcal{F}_{a,b} + \mathcal{F}_{c,d} &= \mathcal{F}_{a+c, b+d} \\ \alpha \cdot \mathcal{F}_{a,b} &= \mathcal{F}_{\alpha a, \alpha b}. \end{aligned}$$

Übung 1.1.8. Zeigen Sie, dass die Menge $\{\mathcal{F}_{\alpha\alpha,\alpha b} : \alpha \in \mathbb{R}\} \subset V$ unter der Annahme $(a, b) \neq (0, 0)$ unendlich viele Elemente enthält, aber trotzdem nicht gleich V ist.

Satz 1.1.9. Sei $\mathcal{F} \in V$. Dann gibt es $a, b \in \mathbb{R}$ so dass

$$(1) \quad \mathcal{F} = a \cdot \mathcal{F}_{1,0} + b \cdot \mathcal{F}_{0,1}.$$

Wir sagen: die Gleichung (13) schreibt $\mathcal{F}$ als eine Linearkombination der Folgen $\mathcal{F}_{0,1}$ und $\mathcal{F}_{1,0}$.

Proof. Aufgrund der Definition von V gilt $\mathcal{F} = \mathcal{F}_{a,b}$ für $a, b \in \mathbb{R}$. Weiterhin gilt

$$\mathcal{F}_{a,b} = a\mathcal{F}_{1,0} + b\mathcal{F}_{0,1}.$$

□

Übung 1.1.10. Sei $\mathcal{F} \in V$. Dann gibt es Elemente $c, d \in \mathbb{R}$ so dass

$$\mathcal{F} = c \cdot \mathcal{F}_{1,1} + d \cdot \mathcal{F}_{1,-1}.$$

Wir wollen nun die Symmetrien des Raumes V betrachten. Unter einer Symmetrie verstehen wir eine Abbildung $T : V \rightarrow V$, die die Struktur von V respektiert.

Definition 1.1.11. Eine Abbildung $T : V \rightarrow V$ ist eine Symmetrie von V , wenn für alle $\mathcal{F}, \mathcal{G} \in V$ und $\alpha \in \mathbb{R}$ gilt, dass

$$T(\mathcal{F} + \mathcal{G}) = T(\mathcal{F}) + T(\mathcal{G}) \quad \text{und} \quad T(\alpha \cdot \mathcal{F}) = \alpha \cdot T(\mathcal{F}).$$

Wir sagen: T ist eine lineare Abbildung.

Beispiele 1.1.12.

(1) Die Identitäts-Abbildung

$$\begin{aligned} \text{id} : V &\rightarrow V, \\ \mathcal{F} &\mapsto \mathcal{F}. \end{aligned}$$

(2) Für $\alpha \in \mathbb{R}$, definiere

$$\begin{aligned} T_\alpha : V &\rightarrow V, \\ \mathcal{F} &\mapsto \alpha \cdot \mathcal{F}. \end{aligned}$$

Übung 1.1.13. Sei $\mathcal{G} \in V$, und definiere die Abbildung

$$\begin{aligned} M_\mathcal{G} : V &\rightarrow V, \\ \mathcal{F} &\mapsto \mathcal{F} + \mathcal{G}. \end{aligned}$$

Zeigen Sie, dass $M_\mathcal{G}$ nur dann eine Symmetrie von V ist, wenn $\mathcal{G} = \mathcal{F}_{0,0}$.

Lemma 1.1.14. Sei $\mathcal{F} = (a_0, a_1, \dots) \in V$. Dann ist die Folge $(a_1, a_2, \dots)$ ebenfalls in V .

Proof. Für $n \geq 0$ sei $b_n = a_{n+1}$; wir müssen zeigen, dass die Folge $(b_n)_{n \geq 0}$ ein Element von V ist. Für $n \geq 2$ gilt

$$b_n = a_{n+1} = a_n + a_{n-1},$$

da $(a_n)_{n \geq 0} \in V$. Aber $a_n + a_{n-1} = b_{n-1} + b_{n-2}$, d.h.

$$b_n = b_{n-1} + b_{n-2}.$$

□

Satz 1.1.15. Definiere die Verschiebungs-Abbildung¹

$$\begin{aligned} S : V &\rightarrow V, \\ (a_0, a_1, \dots) &\mapsto (a_1, a_2, \dots). \end{aligned}$$

Dann ist S eine Symmetrie von V .

Proof. Explizite Rechnung zeigt, dass S die Bedingungen von Definition 1.1.11 erfüllt. □

Beispiel 1.1.16. Es gilt

$$\begin{aligned} S(\mathcal{F}_{-2,2}) &= (2, 0, 2, 2, 4, 6, 10, \dots) \\ &= \mathcal{F}_{0,2}. \end{aligned}$$

¹Es ist klar, dass $S(\mathcal{G}) \in V$ ist für jede Fibonacci Folge $\mathcal{G} \in V$.

Definition 1.1.17. Set $T : V \rightarrow V$ eine Symmetrie. Eine Folge $\mathcal{F} \in V$, $\mathcal{F} \neq \mathcal{F}_{0,0}$ ist eine Eigenfolge wenn es ein Element $\alpha \in \mathbb{R}$ gibt so dass

$$T(\mathcal{F}) = \alpha \cdot \mathcal{F}.$$

In diesem Fall heisst α der Eigenwert der Folge $\mathcal{F}$.

Beispiele 1.1.18.

- (1) Alle $\mathcal{F} \in V$, $\mathcal{F} \neq \mathcal{F}_{0,0}$ sind Eigenfolgen der Identitätsabbildung mit Eigenwert 1.
- (2) Alle $\mathcal{F} \in V$, $\mathcal{F} \neq \mathcal{F}_{0,0}$ sind Eigenfolgen der Abbildung T_α mit Eigenwert α .

Theorem 1.1.19. Es seien $\varphi = \frac{1+\sqrt{5}}{2}$ und $\psi = \frac{1-\sqrt{5}}{2}$. Dann sind $\mathcal{F}_{1,\varphi}$ und $\mathcal{F}_{1,\psi}$ Eigenfolgen der Abbildung S , mit jeweiligen Eigenwerten φ und ψ . Mit anderen Worten, es gilt

$$(2) \quad \mathcal{F}_{1,\varphi} = (1, \varphi, \varphi^2, \varphi^3, \dots),$$

$$(3) \quad \mathcal{F}_{1,\psi} = (1, \psi, \psi^2, \psi^3, \dots).$$

Weiterhin gilt: wenn $\mathcal{A}$ eine Eigenfolge von S ist, dann gibt es ein Element $\beta \in \mathbb{R}$, $\beta \neq 0$ so dass entweder $\mathcal{A} = \beta \cdot \mathcal{F}_{1,\varphi}$ oder $\mathcal{A} = \beta \cdot \mathcal{F}_{1,\psi}$.

Proof. Nimm an, dass $\mathcal{A} = (a_0, a_1, \dots)$ eine Eigenfolge von S ist, mit Eigenwert α . Dann gilt

$$(a_1, a_2, a_3, \dots) = (\alpha \cdot a_0, \alpha \cdot a_1, \alpha \cdot a_2, \dots),$$

d.h. $a_n = \alpha \cdot a_{n-1}$ für alle $n \geq 1$. Mit anderen Worten, $\mathcal{A}$ ist eine geometrische Folge:

$$\mathcal{A} = (a_0, a_0\alpha, a_0\alpha^2, \dots).$$

Da $\mathcal{A} \neq \mathcal{F}_{0,0}$, gilt $a_0 \neq 0$.

Nun ist $\mathcal{A}$ ein Element von V , d.h. $a_n = a_{n-1} + a_{n-2}$ für alle $n \geq 2$. Insbesondere gilt

$$\begin{aligned} a_2 &= a_1 + a_0 & \Leftrightarrow & & a_0\alpha^2 &= a_0\alpha + a_0 \\ \Rightarrow & \alpha^2 - \alpha - 1 = 0. \end{aligned}$$

Die Lösungen dieser quadratischen Gleichung sind φ und ψ , d.h. entweder $\mathcal{A} = a_0 \cdot (1, \varphi, \varphi^2, \dots)$ oder $\mathcal{A} = a_0 \cdot (1, \psi, \psi^2, \dots)$. $\square$

Übung 1.1.20. Zeigen Sie, dass

$$(4) \quad \frac{1}{\varphi - \psi} \cdot \mathcal{F}_{1,\varphi} + \frac{1}{\psi - \varphi} \cdot \mathcal{F}_{1,\psi} = \mathcal{F}_{0,1}.$$

Korollar 1.1.21. Sei $\mathcal{F}_{0,1} = (a_0, a_1, \dots)$. Dann gilt

$$a_n = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n}{\sqrt{5}}.$$

Wir nennen dies eine geschlossene Form der Folge.

Proof. Der n te Eintrag der Folge $\mathcal{F}_{1,\varphi}$ (bzw. der Folge $\mathcal{F}_{1,\psi}$) ist laut (2) (bzw. (3)) gegeben durch φ^n (bzw. durch ψ^n). Es folgt daher von (4), dass

$$a_n = \frac{\varphi^n}{\varphi - \psi} + \frac{\psi^n}{\psi - \varphi}.$$

Indem wir die Werte von φ und ψ einsetzen, erhalten wir die Formel. $\square$

Wir sehen also, dass wir die Betrachtung des reellen Vektorraums V und seiner Symmetries einige sehr interessante (und konkrete) Sätze über Fibonacci-Folgen beweisen können. Wir werden diesem Phänomen in diesem Kurs häufig begegnen: durch sehr abstrakte Argumente erhalten wir Informationen über sehr konkrete mathematische Objekte. Andererseits sind die abstrakten Strukturen oft von sehr konkreten Objekten inspiriert: der Informationsfluss zwischen konkreten und abstrakten Strukturen geht also in beide Richtungen.

1.2. Mengenlehre.

Definition 1.2.1. Eine Menge ist eine Sammlung von Objekten. Die Objekte heissen die Elemente dieser Menge. Wenn x ein Element von M ist, schreiben wir $x \in M$. Wenn y kein Element von M ist, schreiben wir $y \notin M$.

Beispiele 1.2.2.

- (1) $\mathbf{Z} = \{0, 1, -1, 2, -2, 3, -3, 4, -4, 5, \dots\}$ ist die Menge der ganzen Zahlen.
- (2) $M = \{1, 13, -27\} = \{-27, 1, 13\}$ ist eine Menge mit drei Elementen.
- (3) $N = \{\text{Mathematikstudenten mit langen Haaren}\}$ ist eine Menge.
- (4) $P = \{1, 1, 1\} = \{1\}$ ist eine Menge mit einem Element.
- (5) Die leere Menge ist die Menge, die keine Elemente enthaelt: $\emptyset = \{\}$.

Beachte 1.2.3.

- (1) Es spielt keine Rolle, in welcher Reihenfolge wie die Elemente einer Menge aufzaehlen.²
- (2) Wiederholungen von Elementen in einer Menge werden ignoriert: jedes Element kommt nur einmal vor.
- (3) Auch Mengen selber koennen Elemente von Mengen sein.
- (4) Eine Menge muss nicht unbedingt Elemente der gleichen Art enthalten: auch

$$M = \{0, 1, \text{Ehringer Kuehe}\}$$

ist eine Menge.

Beispiel 1.2.4.

- (1) $M = \{\{1, 2\}, \{2, 3\}, \{-1, 0\}\}$ ist die Menge, die als Objekte die Mengen $\{1, 2\}$, $\{2, 3\}$, $\{-1, 0\}$ enthaelt.
- (2) $N = \{\emptyset\}$ ist die Menge, die als einziges Element die leere Menge enthaelt.

Manchmal koennen wir die Elemente einer Menge mit Hilfe einer Eigenschaft angeben:

Definition 1.2.5. $M = \{x : A(x)\}$ (manchmal auch geschrieben $\{x | A(x)\}$) ist die Menge aller x mit einer bestimmten Eigenschaft.

Lecture 2

Beispiel 1.2.6. $M = \{x : x \in \mathbb{Z} \text{ und } x^2 < 9\}$ ist die Menge $M = \{0, \pm 1, \pm 2, \pm 3\}$.

Definition 1.2.7. Es seien P und Q Mengen.

- (1) P ist eine Unter- oder Teilmenge von Q , geschrieben $P \subseteq Q$, wenn fuer alle $x \in P$ gilt: $x \in Q$.
- (2) P ist eine strikte Unter- oder Teilmenge von Q , geschrieben $P \subsetneq Q$, wenn $P \subseteq Q$ und $P \neq Q$.
- (3) Wir schreiben $P \not\subseteq Q$, wenn P keine Teilmenge von Q ist.

Beispiel 1.2.8. (1) Es sei $\mathcal{P}$ die Menge aller Primzahlen. Dann gilt $\mathcal{P} \subsetneq \mathbb{Z}$.
(2) Es sei $\mathbb{Q}_+$ die Menge aller positiven rationalen Zahlen. Dann gilt $\mathbb{Z} \not\subseteq \mathbb{Q}_+$.
(3) Die leere Menge ist eine Teilmenge von jeder Menge.
(4) Es sei $M = \{1, \{1, 2\}\}$. Die Teilmengen von M sind

$$\emptyset, \{1\}, \{\{1, 2\}\}, M.$$

Übung 1.2.9. Finden Sie alle Teilmengen folgender Mengen: $M_1 = \{1\}$, $M_2 = \{1, 2\}$, $M_3 = \{1, 2, 3\}$. Faellt Ihnen etwas auf?

Wenn wir mehrere Mengen haben, dann koennen wir aus ihnen neue Mengen konstruieren:

Definition 1.2.10. Es seien P, Q Mengen. Wir definieren

- (1) die Schnittmenge $P \cap Q = \{x | x \in P \text{ und } x \in Q\}$;
- (2) die Vereinigungsmenge $P \cup Q = \{x | x \in P \text{ oder } x \in Q\}$;
- (3) P ohne Q : $P - Q = \{x | x \in P \text{ und } x \notin Q\}$.

Wenn Q eine Teilmenge von P ist, dann nennen wir P ohne Q das Komplement von Q (in P), geschrieben Q^c .

Beispiel 1.2.11. (1) Es sei $\mathcal{P}$ die Menge aller Primzahlen und Q die Menge aller Quadratzahlen. Dann gilt $\mathcal{P} \cap Q = \emptyset$.

²Wenn die Reihenfolge doch eine Rolle spielen soll, dann betrachten wir *geordnete Mengen*.

- (2) Es sei M die Menge aller Professoren an der ETH und N die Teilmenge aller Professoren ohne Haare. Dann ist $N^c = M - N$ die Menge aller Professoren an der ETH mit mindestens einem Haar.
- (3) Fuer jede Menge M gilt $M - \emptyset = M$.

Wir koennen auch Schnitt- und Vereinigungsmengen von (moeglicherweise unendlichen) Familien (d.h Mengen) von Mengen bilden:

Definition 1.2.12. Es sei $\mathcal{A}$ eine Familie von Mengen, $\mathcal{A} \neq \emptyset$. Dann ist

$$\bigcup_{A \in \mathcal{A}} A = \{x : \text{es gibt } A \in \mathcal{A} \text{ so dass } x \in A\};$$

$$\bigcap_{A \in \mathcal{A}} A = \{x : x \in A \text{ fuer alle } A \in \mathcal{A}\}.$$

Beispiel 1.2.13. Fuer $n \in \mathbf{N}$ sei

$$A_n = \{x \in \mathbf{Z} : x > 0 \text{ und } x \text{ ist das Produkt von genau } n \text{ Primfaktoren}\}.$$

Dann gilt

$$\bigcup_n A_n = \mathbf{Z} - \{1\} \quad \text{und} \quad \bigcap_n A_n = \emptyset.$$

Definition 1.2.14. Es seien X, Y zwei Mengen. Das Cartesische Produkt von X und Y ist die Menge aller geordneten Paar (x, y) mit $x \in X$ und $y \in Y$:

$$X \times Y = \{(x, y) : x \in X, y \in Y\}.$$

Beispiel 1.2.15. (1) Es sei $X = \{-1, 0, 1\}$ und $Y = \{3, 5\}$. Dann ist

$$X \times Y = \{(-1, 3), (0, 3), (1, 3), (-1, 5), (0, 5), (1, 5)\}.$$

(2) Es sei $X = [0, 2]$ und $Y = (1, 2]$. Dann ist

$$X \times Y = \{(x, y) : x \in [0, 2], y \in (1, 2]\}.$$

Definition 1.2.16. Es sei X eine Menge. Das n -fache Cartesische Produkt von X ist die Menge

$$X^n = \{(x_1, \dots, x_n) : x_1, \dots, x_n \in X\}.$$

1.3. Funktionen.

Definition 1.3.1. Es seien X, Y Mengen. Eine Funktion (oder Abbildung) $f : X \rightarrow Y$ ist eine Relation, die jedem Element $x \in X$ genau ein Element $f(x) \in Y$ zuordnet; wir schreiben³ $x \mapsto f(x)$. Wir nennen X die Definitionsmenge und Y die Zielmenge der Funktion.

Beispiel 1.3.2. (1) $f : \mathbb{Z} \rightarrow \mathbb{Z}, n \mapsto n^2$ ist eine Funktion.

(2) $g : \mathbb{Z} \rightarrow \{\pm 1\}, x \mapsto \begin{cases} 1 & \text{wenn } x > 0 \\ -1 & \text{wenn } x < 0 \end{cases}$ ist keine Funktion.⁴

(3) $g : \mathbb{Z} \rightarrow \{\pm 1\}, x \mapsto \begin{cases} 1 & \text{wenn } x > 0 \\ -1 & \text{wenn } x \leq 0 \end{cases}$ ist eine Funktion.

(4) $\text{id}_X : X \rightarrow X, x \mapsto x$ ist die Identitätsfunktion von X .

(5) Es sei $Y \subseteq X$. Die charakteristische Funktion von Y ist die Funktion

$$\text{char}_Y : X \rightarrow \{0, 1\},$$

$$x \mapsto \begin{cases} 1 & \text{wenn } x \in Y \\ 0 & \text{wenn } x \notin Y \end{cases}$$

(6) Addition (bzw. Multiplikation) ist eine Funktion $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$.

Bemerkung 1.3.3. Zwei Funktionen $f : X \rightarrow Y$ und $g : X' \rightarrow Y'$ sind genau dann gleich, wenn $X = X'$, $Y = Y'$ und $f(x) = g(x)$ fuer alle $x \in X$.

Definition 1.3.4. Es sei $f : X \rightarrow Y$ eine Funktion und $A \subseteq X$. Die Beschränkung von f auf A ist die Funktion

$$f|_A : A \rightarrow Y, \quad f|_A(x) = f(x) \quad \text{fuer } x \in A.$$

Definition 1.3.5. Es sei $f : X \rightarrow Y$ eine Funktion.

(1) f ist injektiv wenn fuer alle $x_1, x_2 \in X$ mit $x_1 \neq x_2$ gilt $f(x_1) \neq f(x_2)$.

(2) f ist surjektiv wenn fuer alle $y \in Y$ es mindestens ein $x \in X$ gibt, so dass $f(x) = y$.

(3) f ist bijektiv wenn f injektiv und surjektiv ist; mit anderen Worten, fuer alle $y \in Y$ gibt es genau ein $x \in X$, so dass $f(x) = y$.

Beispiel 1.3.6. (1) Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2$ ist weder injektiv noch surjektiv.

(2) Die Funktion $g : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}, x \mapsto x^2$ ist injektiv aber nicht surjektiv.

(3) Die Funktion $h : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}, x \mapsto x^2$ ist bijektiv.

Definition 1.3.7. Es sei $f : X \rightarrow Y$ bijektiv. Die inverse Funktion $f^{-1} : Y \rightarrow X$ ist die Funktion, die einem Element $y \in Y$ das eindeutig bestimmte Element $x \in X$ zuordnet, fuer das gilt $f(x) = y$.

Beispiel 1.3.8. Die inverse Funktion der Funktion h aus Beispiel 1.3.6 (3) ist gegeben durch

$$h^{-1} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}, \quad x \mapsto \sqrt{x}.$$

Manchmal koennen Funktionen miteinander verknuepft werden:

Definition 1.3.9. Es seien $f : X \rightarrow Y, g : Y \rightarrow Z$ Funktionen. Die Komposition von f und g ist die Funktion

$$g \circ f : X \rightarrow Z, \quad x \mapsto g(f(x)).$$

Beispiel 1.3.10. (1) Es sei $f : X \rightarrow Y$ bijektiv mit inverser Funktion $f^{-1} : Y \rightarrow X$. Dann gilt

$$f^{-1} \circ f = \text{id}_X \quad \text{und} \quad f \circ f^{-1} = \text{id}_Y.$$

Bemerkung 1.3.11. Es seien $f : X \rightarrow Y, g : Y \rightarrow Z$ Funktionen. Dann ist $g \circ f$ definiert, aber $f \circ g$ ist nur dann definiert, wenn $Z = X$! Aber selbst dann ist es im Allgemeinen nicht wahr, dass $f \circ g = g \circ f$!

Beispiel 1.3.12. Es seien $f, g : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch $f(x) = x^2$ und $g(x) = x + 1$. Dann gilt

$$g \circ f(x) = x^2 + 1 \quad \text{aber} \quad f \circ g(x) = (x + 1)^2.$$

Satz 1.3.13. Es seien $f : X \rightarrow Y, g : Y \rightarrow Z$ Funktionen.

(1) Wenn f und g injektiv sind, dann ist auch $g \circ f$ injektiv.

³Auf Vorschlag eines Studenten wird $\mapsto$ gesprochen als Pfeil mit Fuss.

⁴Warum nicht?

(2) Wenn f und g surjektiv sind, dann ist auch $g \circ f$ surjektiv.

Proof. Übung.

□

Korollar 1.3.14. Wenn f und g bijektiv sind, dann ist auch $g \circ f$ bijektiv.

2.1. Körper. Sie haben bereits gesehen, dass $\mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ ähnlich Strukturen haben: sie besitzen Addition und Multiplikation. Wir wollen diese Gemeinsamkeiten durch Axiome beschreiben:

Definition 2.1.1. Ein Körper (eng. field) ist eine Menge $K \neq \emptyset$ mit zwei Abbildungen (oder Verknüpfungen)

$$\begin{aligned} + : K \times K &\rightarrow K & (\text{Addition}) \\ \cdot : K \times K &\rightarrow K & (\text{Multiplikation}) \end{aligned}$$

und zwei ausgezeichneten Elementen $0, 1 \in K$ mit $0 \neq 1$, die folgende Eigenschaften erfüllen:

- (1) Assoziativität der Addition: $x + (y + z) = (x + y) + z$ für alle $x, y, z \in K$,
- (2) Kommutativität der Addition: $x + y = y + x$ für alle $x, y \in K$,
- (3) Neutrales Element der Addition: $0 + x = x$ für alle $x \in K$,
- (4) Additives inverses Element: $\forall x \in K$ gibt es $x' \in K$, so dass $x + x' = 0$;
- (5) Assoziativität der Multiplikation: $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ für alle $x, y, z \in K$,
- (6) Kommutativität der Multiplikation: $x \cdot y = y \cdot x$ für alle $x, y \in K$,
- (7) Neutrales Element der Multiplikation: $1 \cdot x = x$ für alle $x \in K - \{0\}$,
- (8) Multiplikatives inverses Element: $\forall x \in K - \{0\}$ gibt es ein $x' \in K$ so dass $x \cdot x' = 1$;
- (9) Distributivität: $x \cdot (y + z) = x \cdot y + x \cdot z$ für alle $x, y, z \in K$.

Gibt es noch andere Körper? Die Antwort ist ja: es gibt viele verschiedene Arten von Körpern.

Übung 2.1.2. Die Menge

$$\mathbb{Q}(i) = \{a + bi : a, b \in \mathbb{Q}\}$$

mit der Addition und Multiplikation komplexer Zahlen ist ein Körper. Ebenso ist die Menge

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$$

ein Körper. Diese Arten von Körpern heißen Zahlkörper (engl. number fields); sie sind sehr wichtig in der algebraischen Zahlentheorie.

Eine weitere sehr interessante Klasse von Körpern sind die Körper mit endlich vielen Elementen.

Definition 2.1.3. Es sei p eine Primzahl. Der Körper $\mathbb{F}_p$ ist folgendermaßen definiert: die Elemente sind $\{0, \dots, p-1\}$. Für $x, y \in \mathbb{F}_p$ definieren wir $x + y \in \mathbb{F}_p$ (bzw. $xy \in \mathbb{F}_p$) als den Rest, der bei der Division durch p entsteht.

Bemerkung 2.1.4. Die einzige Schwierigkeit zu zeigen, dass $\mathbb{F}_p$ die Körperaxiome erfüllt, ist die Existenz eines multiplikativen inversen Elements.

Lemma 2.1.5. Es sei $x \in \mathbb{F}_p$, $x \neq 0$. Dann gibt es ein $y \in \mathbb{F}_p$ so dass $xy = 1$; wir schreiben $y = x^{-1}$.

Proof. Wir betrachten die Menge

$$M = \{1 \cdot x, 2 \cdot x, \dots, (p-1) \cdot x\}.$$

Behauptung 1. Keines der Elemente von M ist durch p teilbar.

Beweis von Behauptung 1. Offensichtlich, da p eine Primzahl ist.

Behauptung 2. Keine zwei Elemente von M haben den gleichen Rest bei Division durch p .

Beweis von Behauptung 2. Nimm an, dass $i \cdot x$ und $j \cdot x$ (mit $i \geq j$) den gleichen Rest bei der Division durch p haben. Dann ist $(i-j)x$ durch p teilbar, und da p eine Primzahl ist, bedeutet dies, dass $p|(i-j)$ oder $p|x$. Nun gilt aber

$$1 \leq i, j, x < p,$$

woraus wir folgern, dass $i = j$. Das beweist Behauptung 2.

Da die Menge M $p-1$ Elemente enthält, folgt aus Behauptung 2, dass es ein $1 \leq y < p$ gibt, so dass xy den Rest 1 bei der Division durch p hat. $\square$

Satz 2.1.6. $\mathbb{F}_p$ ist ein Körper.

Proof. Übung. $\square$

Beispiele 2.1.7.

- (1) Die Additions- und Multiplikationstabellen von $\mathbb{F}_2$ sind gegeben durch

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

- (2) Die Additions- und Multiplikationstabellen von $\mathbb{F}_3$ sind gegeben durch

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

·	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Bemerkung 2.1.8.

- (1) Der Körper $\mathbb{F}_p$ kann mit Hilfe von modularem Rechnen / Kongruenzen wesentlich eleganter definiert werden. Aber das gehört in den Bereich der Zahlentheorie.
- (2) Es sei $n \geq 2$. Sie können dann natürlich auch die Menge der Reste betrachten, die bei der Division durch n entstehen. Addition und Multiplikation können wie in dem Fall, wenn n eine Primzahl ist, definiert werden; wenn n keine Primzahl ist, dann hat aber nicht jedes Element, das nicht null ist, ein multiplikatives inverses Element.
- (3) Übung für Ehrgeizige: es sei M die Menge der Reste, die bei der Division durch 10 entstehen. Finden Sie heraus, welche Elemente von M ein multiplikatives inverses Element besitzen. Was stellen Sie fest?

Bemerkung 2.1.9. Eine Menge R mit Operationen $+$ und $\cdot$, die alle Axiome aus Definition 2.1.1 ausser (8) erfüllen, nennt man einen kommutativen Ring; so ist z.B. $\mathbb{Z}$ ein kommutativer Ring. Ebenso ist die Menge der Reste, die bei Division durch n entstehen, ein kommutativer Ring, der genau dann ein Körper ist, wenn n eine Primzahl ist.

2.2. Matrizen. Es sei K ein Körper.

Definition 2.2.1. Seien $m, n \geq 1$. Eine $(m \times n)$ Matrix A mit Werten in K ist eine Tabelle mit m Zeilen und n Spalten, deren Einträge Elemente von K sind. Wir schreiben

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

wobei a_{ij} den Eintrag in Reihe i und Spalte j bezeichnet. Die Matrix A ist quadratisch, wenn $m = n$ gilt.

Wir schreiben $M_{m \times n}(K)$ fuer die Menge aller $(m \times n)$ -Matrizen mit Werten in K .

Definition 2.2.2. Eine $(1 \times n)$ Matrix wird als Zeilenvektor, eine $(m \times 1)$ Matrix als Spaltenvektor bezeichnet.

Definition 2.2.3. Wir schreiben $0_{m \times n}$ fuer die $(m \times n)$ Matrix, deren Werte alle Null sind.

Wir schreiben 1_n fuer die Matrix $(a_{ij}) \in M_{n \times n}(K)$, fuer die $a_{ii} = 1 \ \forall 1 \leq i \leq n$ und $a_{ij} = 0 \ \forall 0 \leq i, j \leq n, i \neq j$ gilt.

Beispiel 2.2.4. Es ist $1_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ und $1_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Definition 2.2.5.

- (1) Seien $A = (a_{ij})$ und $B = (b_{ij})$ Elemente von $M_{m \times n}(K)$. Wir definieren die Summe $A + B$ als die $(m \times n)$ Matrix (c_{ij}) mit $c_{ij} = a_{ij} + b_{ij}$.
- (2) Sei $A = (a_{ij}) \in M_{m \times n}(K)$ und $\alpha \in K$. Dann ist das Skalarprodukt $\alpha \cdot A \in M_{m \times n}(K)$ als die Matrix mit Einträgen (αa_{ij}) definiert.

Lecture 4

Theorem 2.2.6. Es seien $A, B, C \in M_{m \times n}(K)$ und $\alpha, \beta \in K$. Dann gilt ⁵

- (1) (Kommutativitaet) $A + B = B + A$
- (2) (Assoziativitaet) $A + (B + C) = (A + B) + C$;
- (3) $A + 0_{m \times n} = 0_{m \times n} + A = A$;
- (4) $\alpha(\beta \cdot A) = (\alpha\beta) \cdot A$;
- (5) $(\alpha + \beta)A = \alpha A + \beta A$;
- (6) $\alpha(A + B) = \alpha A + \alpha B$.

Proof. (1) Wir schreiben $A = (a_{ij})$ und $B = (b_{ij})$. Dann ist

$$(A + B)_{ij} = a_{ij} + b_{ij} = b_{ij} + a_{ij} = (B + A)_{ij}.$$

Die anderen Aussagen koennen aehnlich bewiesen werden. □

Die Multiplikation von zwei Matrizen ist komplizierter:

Definition 2.2.7. Let $A \in M_{m \times n}(K)$ and $B \in M_{n \times p}(K)$. Dann ist das Produkt AB die $(m \times p)$ -Matrix⁶ $C = (c_{ik})$ mit

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{in}b_{nk} = \sum_{j=1}^n a_{ij}b_{jk}.$$

Beispiele 2.2.8. Es gilt

$$\begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y \\ 2x + 3y \end{pmatrix}$$

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

Theorem 2.2.9.

- (1) (Assoziativitaet) Seien $A \in M_{m \times n}(K)$, $B \in M_{n \times p}(K)$ und $C \in M_{p \times q}(K)$. Dann gilt
- $$A(BC) = (AB)C.$$

⁵In Kapitel 3 werden wir sehen, dass diese Eigenschaften genau die Vektorraumaxiome sind: $M_{m \times n}(K)$ sind ein K -Vektorraum.

⁶Wir werden in Kapitel 4 sehen, warum dies eine gute Definition ist: wenn man Matrizen als lineare Abbildungen interpretiert, dann entspricht die Matrix Multiplikation der Verknuepfung von Abbildungen.

(2) (Distributivitaet 1) Seien $A \in M_{m \times n}(K)$, $B \in M_{n \times p}(K)$ und $C \in M_{n \times p}(K)$. Dann gilt

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

(3) (Distributivitaet 2) Seien $A, B \in M_{m \times n}(K)$ und $C \in M_{n \times p}(K)$. Dann gilt

$$(A + B)C = AC + BC.$$

(4) Seien $A \in M_{m \times n}(K)$, $B \in M_{n \times p}(K)$ und $\alpha \in K$. Dann gilt

$$\alpha(AB) = (\alpha A)B = A(\alpha B).$$

Proof. Ueberpruefen Sie die Aussagen direkt von der Definition. □

Beachte 2.2.10. Es seien $A \in M_{m \times n}(K)$ und $B \in M_{n \times p}(K)$. Dann sind die beiden Produkte AB und BA dann und nur dann definiert, wenn $m = p$. In diesem Fall ist $AB \in M_{m \times m}(K)$ und $BA \in M_{n \times n}(K)$.

Beispiele 2.2.11.

(i) Es seien $A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$ und $B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$. Dann gilt

$$AB = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \quad BA = \begin{pmatrix} 1 & -2 & -2 \\ 0 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix}$$

(ii) Es seien nun A wie in (i) und $C = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$. Dann ist

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

aber CA ist nicht definiert.

Beispiel 2.2.12. Sei $A \in M_{n \times n}(K)$. Dann gilt

$$A\mathbf{1}_n = \mathbf{1}_n A = A.$$

Definition 2.2.13. Es seien A, B quadratische Matrizen. Dann kommutieren A und B , wenn $AB = BA$.

Warnung 2.2.14. Es sei $n \geq 2$. Dann gibt es viele Matrizen $A, B \in M_{n \times n}(K)$, die nicht kommutieren!

Übung 2.2.15. Sei $n \geq 2$. Finden Sie Matrizen $A, B \in M_{n \times n}(\mathbb{R})$ so dass $AB \neq BA$.

Definition 2.2.16. Es sei $A = (a_{ij}) \in M_{n \times n}(K)$.

(1) A ist diagonal falls $a_{ij} = 0$ fuer all $i \neq j$.

(2) A ist eine obere (untere) Dreiecksmatrix falls $a_{ij} = 0$ fuer alle $i > j$ (bzw. $i < j$).

Beispiel 2.2.17. Fuer all n ist die Einheitsmatrix $\mathbf{1}_n$ diagonal. Die Matrix $\begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ ist eine obere Dreiecksmatrix.

Lemma 2.2.18. Es seien $A, B \in M_{n \times n}(K)$ beide $\begin{cases} \text{diagonal} \\ \text{eine obere Dreiecksmatrix} \\ \text{eine untere Dreiecksmatrix} \end{cases}$. Dann ist AB von der gleichen Form.

Proof. Explizite Rechnung. □

Definition 2.2.19. Eine $(n \times n)$ -Matrix $A \in M_{n \times n}(K)$ ist invertierbar, wenn es eine $(n \times n)$ -Matrix $B \in M_{n \times n}(K)$ gibt, so dass⁷

$$(5) \quad AB = BA = \mathbf{1}_n.$$

⁷Wir werden spaeter im Kurs sehen, dass es ausreicht, wenn $AB = \mathbf{1}_n$; die Gleichung $BA = \mathbf{1}_n$ haelt dann automatisch (und umgekehrt).

Lemma 2.2.20. Wenn A invertierbar ist, dann gibt es genau ein B , dass Gleichung (41) erfuehlt. Wir nennen B die inverse Matrix von A , und wir schreiben A^{-1} .

Proof. Nehmen wir an, es gibt zwei $(n \times n)$ -Matrizen B und B' , die die Gleichung (41) erfuehlen. Dann gilt

$$B = B\mathbf{1}_n = B(AB') = (BA)B' = \mathbf{1}_n B' = B'.$$

□

Beispiele 2.2.21.

- (1) Die Matrix $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M_{2 \times 2}(\mathbb{Q})$ ist invertierbar, und $A^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.
- (2) Die Matrix $B = \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix} \in M_{2 \times 2}(K)$ fuer einen beliebigen Koerper K ist nicht invertierbar.
- (3) Warnung: Manchmal haengt es vom Grundkoerper ab, ob eine Matrix invertierbar ist oder nicht!
Es sei $C = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \in M_{2 \times 2}(K)$. Diese Matrix ist invertierbar fuer $K = \mathbb{C}$ oder $K = \mathbb{Q}$, aber sie ist nicht invertierbar fuer $K = \mathbb{F}_3$.

Bemerkung 2.2.22.

- (1) Wir werden uns spaeter im Kurs mit Fragen der Invertierbarkeit naecher beschaeftigen.
- (2) Man kann zeigen: wenn A eine invertierbare diagonale bzw. obere/untere Dreiecksmatrix ist, dann ist A^{-1} von der gleichen Form.

Theorem 2.2.23. Es seien A, B invertierbare $(n \times n)$ Matrizen. Dann gilt

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Proof. Übung.

□

Bemerkung 2.2.24. Die Umkehrung der Faktoren kennen wir aus dem Alltag: wenn wir morgens erst Socken und dann Schuhe anziehen, so ziehen wir sie abends in umgekehrter Reihenfolge wieder aus.⁸

⁸Diese Beobachtung stammt von Hermann Weyl.

2.3. Elementare Zeilenoperationen.

Definition 2.3.1. Es sei A eine $(m \times n)$ -Matrix. Die elementaren Zeilenumformungen (EZU) auf A sind folgendermassen definiert:

- $P(r, s)$ fuer $1 \leq r < s \leq m$: Vertauschen der Zeilen r und s ;
- $M(r, \lambda)$ fuer $1 \leq r \leq m$ und $\lambda \in K, \lambda \neq 0$: Multiplikation der Zeile r (d.h. aller Werte der Zeile r) mit λ ;
- $S(r, s, \lambda)$: fuer $1 \leq r, s \leq m, r \neq s$ und $\lambda \in K, \lambda \neq 0$: Addition von $\lambda \times$ (Zeile r) zur Zeile s .

Definition 2.3.2. Wir sagen, dass zwei Matrizen A und A' zeilen-aequivalent sind, wenn wir A' durch die Anwendung von endlich vielen (EZU)s auf A erhalten.

Beispiel 2.3.3. Es sei $A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$. Dann gilt

$$\begin{aligned} P(2, 3) &\rightarrow \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \\ M(1, -1) &\rightarrow \begin{pmatrix} -1 & -2 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \\ S(2, 1, 2) &\rightarrow \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}. \end{aligned}$$

Bemerkung 2.3.4. Diese Definition ist symmetrisch in A und A' : wenn wir A' durch die Anwendung von endlich vielen (EZU)s auf A erhalten, dann erhalten wir auch A durch die Anwendung von endlich vielen (EZU)s auf A' ; mit anderen Worten, die (EZU)s sind umkehrbar. Zeigen Sie das.

Definition 2.3.5. Eine $(m \times n)$ -Matrix ist in reduzierter Zeilenform RZF, wenn folgende Bedingung erfuehlt ist:

- (1) in jeder Zeile ist das erste von null verschiedene Element (falls es eins gibt) eine 1 (wir nennen es die fuehrende 1),
- (2) ausser einer fuehrenden 1 sind in dessen Spalte nur Nullen.

Lecture 5

Beispiele 2.3.6.

- (1) Die Matrix $A = \begin{pmatrix} 0 & 0 & 0 & 1 & 2 \\ 0 & 1 & -2 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ ist in reduzierter Zeilenform.
- (2) Die Matrix $B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ ist nicht in reduzierter Zeilenform.
- (3) Die Matrix $C = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix}$ ist nicht in reduzierter Zeilenform.
- (4) Die Matrix $D = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 5 \end{pmatrix}$ ist in reduzierter Zeilenform.

Theorem 2.3.7. Jede Matrix A ist zeilen-aequivalent zu einer Matrix in reduzierter Zeilenform.

Proof. Schreibe $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ mit $a_{ij} \in K$.

- Falls alle Werte in der ersten Zeile null sind, dann ist Bedingung (1) fuer diese Zeile erfuehlt.
- Falls es einen Wert in der ersten Zeile gibt, der nicht null ist, dann sei k der kleinste Index j , so dass $a_{1j} \neq 0$ (d.h. $a_{1k} \neq 0$, aber $a_{1j} = 0$ fuer alle $1 \leq j < k$). Wende die (EZU) $M(1, a_{1k}^{-1})$ an; dann ist Bedingung (1) erfuehlt.
- Wende (EZU) $S(1, i, -a_{ik})$ an fuer alle $2 \leq i \leq m$; damit sind alle anderen Werte in Spalte k gleich null.

Wir wenden jetzt die gleiche Strategie auf die zweite Zeile an.

- Falls alle Werte in der zweiten Zeile null sind, dann ist Bedingung (1) fuer diese Zeile erfuehlt.
- Falls es einen Wert in der zweiten Zeile gibt, der nicht null ist, dann sei ℓ der kleinste Index j , so dass $a_{2j} \neq 0$ (d.h. $a_{2\ell} \neq 0$, aber $a_{2j} = 0$ fuer alle $1 \leq j < \ell$). Wende die (EZU) $M(2, a_{2\ell}^{-1})$ an; dann ist Bedingung (1) erfuehlt. (Beachte: $\ell \neq k!$)
- Wende (EZU) $S(2, i, -a_{i\ell})$ an fuer alle $1 \leq i \leq m, i \neq 2$; damit sind alle anderen Werte in Spalte ℓ gleich null.

Wiederhole das Verfahren fuer die anderen Zeilen. $\square$

Definition 2.3.8. Eine $(m \times n)$ Matrix A ist in reduzierter Zeilenstufenform RZSF (eng: row-reduced echelon form), wenn folgende Bedingungen erfuehlt sind:

- (1) A ist in reduzierter Zeilenform;
- (2) alle Zeilen von A , deren Werte alle gleich null sind, liegen unter den Zeilen, die einen von Null verschiedenen Wert enthalten;
- (3) die fuehrende 1 einer Zeile liegt rechts von der fuehrenden 1 der Zeile darueber.

Mit anderen Worten, A hat folgende Form:

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

Theorem 2.3.9. Jede Matrix ist zeilen-aequivalent zu einer Matrix in reduzierter Zeilenstufenform.

Proof. In Satz 2.3.7 haben wir bewiesen, dass jede Matrix zeilen-aequivalent zu einer Matrix in reduzierter Zeilenform ist. Es sei nun A eine Matrix in reduzierter Zeilenform. Wir koennen sie in reduzierte Zeilenstufenform ueberfuehren, indem wir endlich oft die (EZU) $P(r, s)$ (d.h. Vertauschen von Zeilen) anwenden. $\square$

Beispiel 2.3.10. Es sei $A = \begin{pmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{pmatrix} \in M_{3 \times 4}(\mathbb{R})$.

$$\begin{aligned} M(1, -\frac{1}{9}) &\rightarrow \begin{pmatrix} 0 & 1 & -\frac{1}{3} & -\frac{4}{9} \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{pmatrix} \\ S(1, 2, -4) &\rightarrow \begin{pmatrix} 0 & 1 & -\frac{1}{3} & -\frac{4}{9} \\ 1 & 0 & \frac{4}{3} & \frac{7}{9} \\ 2 & 6 & -1 & 5 \end{pmatrix} \\ S(1, 3, -6) &\rightarrow \begin{pmatrix} 0 & 1 & -\frac{1}{3} & -\frac{4}{9} \\ 1 & 0 & \frac{4}{3} & \frac{7}{9} \\ 2 & 0 & 1 & \frac{23}{3} \end{pmatrix} \\ S(2, 3, -2) &\rightarrow \begin{pmatrix} 0 & 1 & -\frac{1}{3} & -\frac{4}{9} \\ 1 & 0 & \frac{4}{3} & \frac{7}{9} \\ 0 & 0 & -\frac{5}{3} & \frac{55}{9} \end{pmatrix} \\ M(3, -\frac{3}{5}) &\rightarrow \begin{pmatrix} 0 & 1 & -\frac{1}{3} & -\frac{4}{9} \\ 1 & 0 & \frac{4}{3} & \frac{7}{9} \\ 0 & 0 & 1 & -\frac{11}{3} \end{pmatrix} \\ S(3, 2, -\frac{4}{3}) &\rightarrow \begin{pmatrix} 0 & 1 & -\frac{1}{3} & -\frac{4}{9} \\ 1 & 0 & 0 & \frac{17}{3} \\ 0 & 0 & 1 & -\frac{11}{3} \end{pmatrix} \\ S(3, 1, \frac{1}{3}) &\rightarrow \begin{pmatrix} 0 & 1 & 0 & -\frac{5}{3} \\ 1 & 0 & 0 & \frac{17}{3} \\ 0 & 0 & 1 & -\frac{11}{3} \end{pmatrix} \\ P(1, 2) &\rightarrow \begin{pmatrix} 1 & 0 & 0 & \frac{17}{3} \\ 0 & 1 & 0 & -\frac{5}{3} \\ 0 & 0 & 1 & -\frac{11}{3} \end{pmatrix} \end{aligned}$$

Im naechsten Abschnitt werden wir sehen, dass diese Rechnungen mit Matrizen sehr nuetzlich sind, um lineare Gleichungssysteme zu loesen.

2.4. Lineare Gleichungssysteme.

Beachte 2.4.1. Betrachte das lineare Gleichungssystem

$$\begin{aligned} 2x - y &= 3 \\ -x + 3y &= -1 \end{aligned}$$

Wir koennen es mit Hilfe von Matrizen schreiben als

$$\begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

Allgemein gilt: gegeben seien $m, n \geq 1$ und $\forall 1 \leq i \leq m$ eine Gleichung

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

mit Unbekannten x_j . Dann koennen wir dieses Gleichungssystem in Matrixform ausdruecken:

$$Ax = b, \quad \text{wobei} \quad A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Definition 2.4.2. Es sei $(S) : Ax = b$ ein lineares Gleichungssystem, wobei A eine $(m \times n)$ -Matrix ist.

- Wir schreiben $L(S)$ fuer die Loesungen des Gleichungssystems, d.h.

$$L(S) = \left\{ x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} : x_i \in K, Ax = b \right\}.$$

- Die erweiterte Matrix $A|b$ ist die $(m \times (n+1))$ Matrix, bei der der Spaltenvektor b als $(n+1)$ ste Spalte der Matrix A hinzugefuegt wird.

Beachte 2.4.3. Das lineare Gleichungssystem $Ax = b$ ist durch die erweiterte Matrix $A|b$ vollstaendig bestimmt.

Bemerkung 2.4.4. Wenn $b = 0$, dann schreiben wir A anstatt der erweiterten Matrix $A|0$; man nennt ein lineares Gleichungssystem der Form $Ax = 0$ homogen.

Theorem 2.4.5. Es seien

$$(S) : Ax = b \quad \text{und} \quad (S') : A'x = b'$$

lineare Gleichungssysteme von jeweils m Gleichungen in n Unbekannten. Nimm an, dass die erweiterte Matrix $A'|b'$ zeilen-aequivalent zu $A|b$ ist. Dann gilt $L(S') = L(S)$.

Proof. Wir machen zunaechst folgende Beobachtungen:

- es ist ausreichend, den Satz zu zeigen, dass wenn (S') von (S) durch eine (Ezu) konstruiert ist;
- es folgt von Bemerkung 2.3.4, dass es ausreichend ist zu zeigen, dass $L(S) \subseteq L(S')$.⁹

Wir muessen also folgende Behauptung beweisen: wenn (S') von (S) durch eine (Ezu) konstruiert ist, dann gilt $L(S) \subseteq L(S')$. Wir analysieren dazu jede der drei Operationen einzeln: fuer die Operationen $P(r, s)$ und $M(r, \lambda)$ ist die Behauptung offensichtlich. Fuer die Operation $S(r, s, \lambda)$ argumentieren wir wie folgt: wenn (S') von (S) durch die Operation $S(r, s, \lambda)$ konstruiert wird, dann unterscheiden sich die erweiterten Matrizen $A|b$ und $A'|b'$ nur in der Zeile s :

$$a'_{sj} = a_{sj} + \lambda a_{rj} \quad \text{und} \quad b'_s = b_s + \lambda b_r.$$

Es sei x eine Loesung von (S) , d.h. $\forall 0 \leq i \leq m$ gilt

$$a_{i1}x_1 + \dots + a_{in}x_n = b_i.$$

Dann gilt ebenfalls

$$\begin{aligned} a_{s1}x_1 + \dots + a_{sn}x_n + \lambda(a_{r1}x_1 + \dots + a_{rn}x_n) &= b_s + \lambda b_r \\ \Leftrightarrow a'_{s1}x_1 + \dots + a'_{sn}x_n &= b'_s, \end{aligned}$$

d.h. x ist eine Loesung von (S') . □

⁹Durch Symmetrie folgt dann, dass $L(S') \subseteq L(S)$, und so $L(S) = L(S')$.

Wir koennen also versuchen, ein lineares Gleichungssystem $(S) : Ax = b$ zu loesen, indem wir die erweiterte Matrix durch (EZO)s in reduzierte Zeilenstufenform umformen.

Beispiel 2.4.6. Betrachten wir das lineare Gleichungssystem $(S) : Ax = 0$, wobei

$$A = \begin{pmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{pmatrix} \in M_{3 \times 4}(\mathbb{R}).$$

Wir haben bereits in Beispiel 2.3.10 gesehen, dass A zeilen-aequivalent zu folgender Matrix ist:

$$A' = \begin{pmatrix} 1 & 0 & 0 & \frac{17}{3} \\ 0 & 1 & 0 & -\frac{5}{3} \\ 0 & 0 & 1 & -\frac{11}{3} \end{pmatrix}.$$

Wir koennen nun also direkt die Loesungen ablesen:

$$L(S) = \left\{ x_1 = -\frac{17}{3}x_4, x_2 = \frac{5}{3}x_4, x_3 = \frac{11}{3}x_4 : x_4 \in K \right\}.$$

Betrachten wir nun das Gleichungssystem $(S') : Ax = b$, wobei $b = \begin{pmatrix} 9 \\ 5 \\ -5 \end{pmatrix}$. Die erweiterte Matrix ist

$$\left(\begin{array}{cccc|c} 0 & -9 & 3 & 4 & 9 \\ 1 & 4 & 0 & -1 & 5 \\ 2 & 6 & -1 & 5 & -5 \end{array} \right).$$

Mit den gleichen Schritten wie in Beispiel 2.3.10 wandeln wir sie in reduzierte Zeilenstufenform um:

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & \frac{17}{3} & -\frac{22}{5} \\ 0 & 1 & 0 & -\frac{5}{3} & \frac{12}{5} \\ 0 & 0 & 1 & -\frac{11}{3} & \frac{51}{5} \end{array} \right)$$

d.h.

$$x_1 + \frac{17}{3}x_4 = -\frac{22}{5}, \quad x_2 - \frac{5}{3}x_4 = \frac{12}{5}, \quad x_3 - \frac{11}{3}x_4 = \frac{51}{5},$$

und wir erhalten die Loesungsmenge

$$L(S') = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} : x_1 = -\frac{22}{5} - \frac{17}{3}x_4, x_2 = \frac{12}{5} + \frac{5}{3}x_4, x_3 = \frac{51}{5} + \frac{11}{3}x_4 : x_4 \in K \right\}.$$

Beispiel 2.4.7. Gegeben sei

$$(S) : \begin{pmatrix} 1 & -2 & 1 \\ 2 & 1 & 1 \\ 0 & 5 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix},$$

wobei b_1, b_2, b_3 Elemente von $\mathbb{R}$ sind. Die erweiterte Matrix ist

$$A = \left(\begin{array}{ccc|c} 1 & -2 & 1 & b_1 \\ 2 & 1 & 1 & b_2 \\ 0 & 5 & -1 & b_3 \end{array} \right).$$

Wir wandeln sie in reduzierte Zeilenstufenform um:

$$\begin{aligned}
 S(1, 2, -2) &\rightarrow \left(\begin{array}{ccc|c} 1 & -2 & 1 & b_1 \\ 0 & 5 & -1 & b_2 - 2b_1 \\ 0 & 5 & -1 & b_3 \end{array} \right) \\
 S(2, 3, -1) &\rightarrow \left(\begin{array}{ccc|c} 1 & -2 & 1 & b_1 \\ 0 & 5 & -1 & b_2 - 2b_1 \\ 0 & 0 & 0 & b_3 - b_2 + 2b_1 \end{array} \right) \\
 M(2, \frac{1}{5}) &\rightarrow \left(\begin{array}{ccc|c} 1 & -2 & 1 & b_1 \\ 0 & 1 & -\frac{1}{5} & \frac{1}{5}(b_2 - 2b_1) \\ 0 & 0 & 0 & b_3 - b_2 + 2b_1 \end{array} \right) \\
 S(2, 1, 2) &\rightarrow \left(\begin{array}{ccc|c} 1 & 0 & \frac{3}{5} & \frac{1}{5}(b_1 + 2b_2) \\ 0 & 1 & -\frac{1}{5} & \frac{1}{5}(b_2 - 2b_1) \\ 0 & 0 & 0 & b_3 - b_2 + 2b_1 \end{array} \right)
 \end{aligned}$$

In other words, (S) is equivalent to

$$x_1 = -\frac{3}{5}x_3 + \frac{1}{5}(b_1 + 2b_2), \quad x_2 = \frac{1}{5}x_3 + \frac{1}{5}(b_2 - 2b_1), \quad 0 = b_3 - b_2 + 2b_1.$$

Es koennen also folgende Faelle auftreten:

- wenn $b_3 - b_2 + 2b_1 \neq 0$, dann gibt es keine Loesung: $L(S) = \emptyset$;
- wenn $b_3 - b_2 + 2b_1 = 0$, dann ist das System aequivalent zu

$$x_1 = -\frac{3}{5}x_3 + \frac{1}{5}(b_1 + 2b_2), \quad x_2 = \frac{1}{5}x_3 + \frac{1}{5}(b_2 - 2b_1),$$

$$\text{d.h. } L(S) = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_1 = -\frac{3}{5}x_3 + \frac{1}{5}(b_1 + 2b_2), x_2 = \frac{1}{5}x_3 + \frac{1}{5}(b_2 - 2b_1) : x_3 \in \mathbb{R} \right\}.$$

3. VEKTORRAEUME

3.1. Definition und Beispiele.

Definition 3.1.1. Es sei K ein Koerper. Ein Vektorraum ueber K ist eine Menge V mit zwei Operationen

$$(6) \quad + : V \times V \rightarrow V, \quad (v_1, v_2) \mapsto v_1 + v_2,$$

$$(7) \quad \times : K \times V \rightarrow V \quad (\lambda, v) \mapsto \lambda v$$

so dass die folgenden Bedingungen erfuehlt sind:

- (VR1) $\forall v_1, v_2 \in V : v_1 + v_2 = v_2 + v_1$;
- (VR2) $\forall v_1, v_2, v_3 \in V : v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$;
- (VR3) $\exists 0_V \in V$ so dass $\forall v \in V$ gilt: $0_V + v = v$;
- (VR4) $\forall v \in V \exists w \in V$ so dass $v + w = 0_V$; with schreiben¹⁰ $w = -v$;
- (VR5) $\forall v_1, v_2 \in V$ und $\forall \lambda \in K$ gilt $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$;
- (VR6) $\forall v \in V$ and $\forall \lambda, \mu \in K$ gilt $(\lambda + \mu)v = \lambda v + \mu v$;
- (VR7) $\forall v \in V$ and $\forall \lambda, \mu \in K$ gilt $\lambda(\mu v) = (\lambda \mu)v$;
- (VR8) $\forall v \in V$ gilt $1v = v$.

Bemerkung 3.1.2. Die Operationen (6) und (7) heissen jeweils Vektor Addition und Skalar-Multiplikation.

Beispiele 3.1.3.

- (1) Die Menge aller Spaltenvektoren $\begin{pmatrix} x \\ y \end{pmatrix}$ mit $x, y \in K$ ist ein K -Vektorraum. Allgemeiner ist die Menge aller Spaltenvektoren mit n Eintraegen ein K -Vektorraum unter der Addition und Skalarmultiplikation von Matrizen (Def. 2.2.5); wir schreiben ihn als K^n .
- (2) Der triviale K -Vektorraum ist $\{0\}$.
- (3) Die Menge aller Fibonacci Folgen von Definition 1.1.3 ist ein $\mathbb{R}$ -Vektorraum.
- (4) Es sei K ein Koerper, and $K[x]^{\leq n}$ die Menge aller Polynome mit Koeffizienten in K vom Grad $\leq n$. Dann ist $K[x]^{\leq n}$ ein K -Vektorraum.
- (5) Es sei

$$V = \{f : (0, 1) \rightarrow \mathbb{R} : f \text{ ist stetig}\}.$$

Dann ist f ein $\mathbb{R}$ -Vektorraum unter der punktweisen Addition von Funktionen.

- (6) Es sei

$$U = \{f \in V : f(1/2) = 0\}.$$

Dann ist U ein $\mathbb{R}$ -Vektorraum: es ist ein Unterraum von V .

- (7) Es sei

$$U' = \{f \in V : f(1/2) = 1\}.$$

Dann ist U' **kein** $\mathbb{R}$ -Vektorraum (warum nicht?).

Satz 3.1.4. Es sei V ein K -Vektorraum. Dann gibt es genau ein Element mit der Eigenschaft (VR3); wir nennen es die additive Identitaet von V .

Proof. Nimm an, dass 0_V und $0'_V$ die Eigenschaft (VR3) haben. Dann gilt

$$0_V = 0_V + 0'_V = 0'_V.$$

□

Korollar 3.1.5. Let $v \in V$. Dann ist das Element in (VR4) eindeutig bestimmt; es heisst das (additive) Inverse¹¹ von v .

Proof. Nimm an, es gibt $u, u' \in V$ so dass

$$v + u = v + u' = 0_V.$$

Dann gilt

$$u = u + 0_V = u + (v + u') = (u + v) + u' = 0_V + u' = u'.$$

□

¹⁰Diese Schreibweise ist an dieser Stelle missverstaendlich, da (noch) nicht klar ist, dass es genau ein w gibt, so dass $v + w = 0_V$ gilt. Wir zeigen die Eindeutigkeit in Korollary 3.1.5.

¹¹Falls Sie etwas Gruppentheorie gelernt haben: $(V, +)$ ist eine abelsche Gruppe.

Satz 3.1.6. *Es sei V ein K Vektorraum, und es seien $\lambda \in K$ und $v \in V$. Dann gilt*

- (1) $\lambda 0_V = 0_V$;
- (2) $0v = 0_V$;
- (3) $(-\lambda)v = -(\lambda v) = \lambda(-v)$;
- (4) *if $\lambda v = 0_V$, then $\lambda = 0$ or $v = 0_V$.*

Proof. (1) - (3) Übung.

(4) Wir zeigen folgende Aussage¹²: wenn $\lambda v = 0_V$ und $\lambda \neq 0$, dann gilt $v = 0$. Da $\lambda \neq 0$, hat λ ein multiplikatives inverses Element $\lambda^{-1} \in K$. Dann gilt

- (8) $\lambda v = 0_V \quad \Leftrightarrow \quad \lambda^{-1}(\lambda v) = 0_V \quad \text{wegen (1)}$
- (9) $\quad \quad \quad \Leftrightarrow \quad (\lambda^{-1}\lambda)v = 0_V \quad \text{wegen (VR7)}$
- (10) $\quad \quad \quad \Leftrightarrow \quad 1.v = v = 0_V \quad \text{wegen (VR8)}$

□

¹²Vergewissern Sie sich, dass diese Aussage äquivalent zu (4) ist.

3.2. Unterräume.

Definition 3.2.1. Es sei V ein K -Vektorraum. Eine Teilmenge $U \subseteq V$ ist ein Unterraum wenn $U \neq \emptyset$ und wenn sie bezüglich Addition und Skalar-Multiplikation abgeschlossen ist, d.h. wenn sie folgende zwei Bedingungen erfuehlt:

- (UR1) $\forall u, v \in U$ gilt $u + v \in U$;
 (UR2) $\forall u \in U$ und $\forall \lambda \in K$ gilt $\lambda u \in U$.

Wir schreiben $U \leq V$.

Der folgende Satz ist nuetzlich um festzustellen, ob eine Teilmenge eines Vektorraumes ein Unterraum ist:

Satz 3.2.2. Eine Teilmenge U von einem Vektorraum V ist dann und nur dann ein Unterraum von V , wenn folgende Bedingungen erfuehlt sind:

- (1) $0_V \in U$;
 (2) $\forall u, v \in U$ und $\lambda \in K$ gilt $\lambda u + v \in U$.

Proof. Nimm an, dass U ein Unterraum ist. Da $U \neq \emptyset$, gibt es ein $u \in U$. Da U bezüglich der Salar-Multiplikation abgeschlossen ist, gilt

$$0u \in U \Leftrightarrow 0_V \in U,$$

da $0u = 0_V$ (Prop. 3.1.6 (2)). Es seien nun $u, v \in U$ und $\lambda \in K$. Dann ist $\lambda u \in U$ wegen (UR2) und $\lambda u + v \in U$ wegen (UR2).

Nimm nun an, dass die Bedingungen erfuehlt sind. Dann ist U nicht die leere Menge da $0_V \in U$. Es seien nun $u, v \in U$. Fuer $\lambda = 1$ erhalten wir $u + v \in U$. Weiterhin, fuer $\lambda \in K$ und $v = 0_V$ erhalten wir

$$\lambda u + 0_V = \lambda u \in U,$$

d.h. U ist ein Unterraum. □

Beispiele 3.2.3.

- (1) Es sei V ein Vektorraum. Dann sind $\{0\}$ und V die trivialen Unterraeume von V .
 (2) Es sei $V = \mathbb{R}^3$, und es sei

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in V : x + y + z = 0 \right\}.$$

Dann ist U ein Unterraum von V .

- (3) Wie im vorherigen Beispiel sei $V = \mathbb{R}^3$. Es sei $\lambda \in \mathbb{R}$, $\lambda \neq 0$. Dann ist

$$U' = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in V : x + y + z = \lambda \right\}$$

kein Unterraum von V , da er unter den Operationen der Addition und der Skalar-Multiplikation nicht abgeschlossen ist.

- (4) Das vorherige Beispiel hat folgende Verallgemeinerung: es sei $A \in M_{m \times n}(K)$ und $b \in M_{m \times 1}(K)$, und wir betrachten das lineare Gleichungssystem

$$(S) : Ax = b.$$

Die Loesungsmenge $L(S) \subseteq K^n$ ist dann und nur dann ein Unterraum von K^n , wenn $b = 0$; in diesem Fall nennen wir (S) ein *homogenes* lineares Gleichungssystem.

- (5) Es sei $V = M_{2 \times 2}(\mathbb{C})$ (ein komplexer Vektorraum unter der normalen Matrix-Addition) und U die Teilmenge der invertierbaren Matrizen. Dann ist U kein Unterraum, da die Null-Matrix nicht invertierbar ist.
 (6) Wieder sei $V = M_{2 \times 2}(\mathbb{C})$ und U' die Teilmenge der nicht invertierbaren Matrizen. Ist U' ein Unterraum? Geben Sie einen Beweis oder ein Gegenbeispiel. Wie ist es mit der Teilmenge U'' der oberen Dreiecksmatrizen?

Das folgende Beispiel ist sehr wichtig:

Lecture 7

Satz 3.2.4. Es sei V ein K -Vektorraum und $v_1, \dots, v_n \in V$. Es sei

$$U = \{\lambda_1 v_1 + \dots + \lambda_n v_n : \lambda_i \in K \forall 1 \leq i \leq n\}.$$

Dann ist U ein Unterraum; wir nennen ihn den von $v_1, \dots, v_n$ erzeugten Unterraum oder die lineare Huelle von $v_1, \dots, v_n$ und schreiben

$$U = \langle v_1, \dots, v_n \rangle.$$

Proof. Wir beweisen den Satz mit Hilfe von Satz 3.2.2: fuer $\lambda_1 = \dots = \lambda_n = 0$ erhalten wir $0_V \in U$. Es seien nun $u, w \in U$ und $\lambda \in K$. Aufgrund der Definition von U gibt es Skalare $\mu_1, \dots, \mu_n$ und $\nu_1, \dots, \nu_n$ so dass

$$u = \mu_1 v_1 + \dots + \mu_n v_n \quad \text{und} \quad w = \nu_1 v_1 + \dots + \nu_n v_n.$$

Daher

$$\begin{aligned} u + \lambda w &= \mu_1 v_1 + \dots + \mu_n v_n + \lambda(\nu_1 v_1 + \dots + \nu_n v_n) \\ &= (\mu_1 + \lambda \nu_1) v_1 + \dots + (\mu_n + \lambda \nu_n) v_n \end{aligned}$$

was ebenfalls ein Element von U ist. □

Bemerkung 3.2.5. Wir koennen diese Beispiel wie folgt verallgemeinern: es sei S eine (moeglicherweise unendliche) Teilmenge von V . Wir definieren die lineare Huelle $\langle S \rangle$ von S als¹³

$$\langle S \rangle = \{v \in V : \exists n \geq 0, v_1, \dots, v_n \in S \text{ und } \lambda_1, \dots, \lambda_n \in K \text{ so dass } v = \lambda_1 v_1 + \dots + \lambda_n v_n\}.$$

Insbesondere ist der Null-Vektorraum die lineare Huelle der leeren Menge.

Beachten Sie, dass S unendlich sein kann: wir erlauben trotzdem nur endliche Summen.¹⁴

Beispiel 3.2.6. Es sei $V = M_{2 \times 2}(\mathbb{R})$, und es sei U der von den Matrizen $e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ und $e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ erzeugt Unterraum. Dann ist U der Unterraum der oberen Dreiecksmatrizen.

Es seien nun $v_1 = e_{11} + e_{12}$, $v_2 = e_{12} + e_{22}$ und $v_3 = e_{22} + e_{11}$. Was ist $\langle v_1, v_2, v_3 \rangle$?

Beispiel 3.2.7. Es sei V der Vektorraum aller Folgen $(a_n)_{n \geq 1}$ mit Eintraegen in $\mathbb{Q}$, und fuer $i \geq 1$ sei $e_i \in V$ die Folge, die an der i ten Stelle eine 1 hat und ansonsten Nullen. Dann ist

$$\langle e_i : i \geq 1 \rangle \neq V,$$

da sich z.B. die konstante Folge $(1, 1, 1, 1, 1, \dots)$ nicht in der linearen Huelle enthalten ist. Andererseits sei V' der Unterraum von V aller *endlichen* Folgen. Dann gilt

$$V' = \langle e_i : i \geq 1 \rangle.$$

Beispiel 3.2.8. Es sei $V = \mathbb{R}^3$, und es sei $U = \left\langle \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix} \right\rangle$. Ist der Vektor $v = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}$ in U enthalten? Wir koennen diese Frage als lineares Gleichungssystem formulieren: hat das System

$$(S) : \begin{pmatrix} 1 & 2 \\ 2 & -2 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}$$

eine Loesung? *Uebung.*

Satz 3.2.9. Es sei V ein Vektorraum und $U \leq V$ ein Unterraum. Dann ist U ebenfalls ein Vektorraum. Ausserdem gilt: wenn $W \leq U$ und $U \leq V$, dann $W \leq V$.

Proof. Ueberpruefen Sie die Axiome. □

Aus zwei Unterraumen lassen sich auf mehrere Arten neue Unterraume konstruieren:

Theorem 3.2.10. Es seien $U, W \leq V$.

(1) Die Schnittmenge

$$U \cap W = \{v \in V : v \in U \wedge v \in W\}$$

ist ein Unterraum von V .

¹³gesprochen *Krokodilklaenger von S*

¹⁴Fuer unendliche Summen braucht man den Begriff der Konvergenz – diese fuehrt zu der Theorie von Banachraeumen.

(2) Die Summe von U und W ,

$$U + W = \{u + w : u \in U, w \in W\}$$

ist ein Unterraum von V .

Proof. Wir beweisen (1) mit Hilfe von Satz 3.2.2: da U und W Unterräume sind, enthalten beide das Element 0_V . Daher gilt $0_V \in U \cap W$. Es seien nun $u, w \in U \cap W$ und $\lambda \in K$. Da U und W Unterräume sind, ist das Element $u + \lambda w$ in beiden Unterräumen enthalten, und daher auch in $U \cap W$.

(2) kann mit ähnlichen Argumenten bewiesen werden (Übung). □

3.3. Basen von Vektorraeumen. Zurueck zum Satz 3.2.4:

Definition 3.3.1. Es sei V ein Vektorraum und $v_1, \dots, v_n \in V$. Ein Element der Form

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

mit $\lambda_i \in K$ ist eine Linearkombination von $v_1, \dots, v_n$.

Beispiel 3.3.2. Es sei $V = \mathbb{R}^3$ und $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$. Dann koennen wir v als eine Linearkombination

der Vektoren $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ und $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ schreiben:

$$v = x e_1 + y e_2 + z e_3.$$

Definition 3.3.3. Es sei V ein Vektorraum. Dann ist V endlich-dimensional wenn es endlich viele Vektoren $v_1, \dots, v_n \in V$ gibt so dass

$$V = \langle v_1, \dots, v_n \rangle :$$

mit anderen Worten, wenn jeder Vektor als Linearkombination der Vektoren $v_1, \dots, v_n$ geschrieben werden kann; diese Linearkombination muss nicht eindeutig sein. Wir nennen $v_1, \dots, v_n$ ein Erzeugendessystem.

Bemerkung. Es sei V ein Vektorraum und $U \leq V$. Dann ist U endlich-dimensional, wenn es $u_1, \dots, u_n \in U$ gibt, so dass

$$U = \langle u_1, \dots, u_n \rangle.$$

Es ist hierbei wichtig, dass die Elemente $u_1, \dots, u_n$ selber in U sind! Um z.B. zu zeigen, dass der Unterraum $U \leq M_{2 \times 2}(\mathbb{R})$,

$$U = \left\{ \begin{pmatrix} a & b \\ c & -a \end{pmatrix} : a, b, c \in \mathbb{R} \right\}$$

endlich-dimesnional ist, reicht es nicht zu sagen, dass sich jede Matrix in U als Linearkombination der Matrizen

$$e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad e_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

schreiben laesst, da e_{11} und e_{22} nicht in U sind.

Beispiele 3.3.4.

- (1) Beispiel 3.3.2 zeigt, dass e_1, e_2, e_3 ein Erzeugendensystem von $\mathbb{R}^3$ ist; daher ist $\mathbb{R}^3$ endlich-dimensional.
- (2) Auf gleiche Weise koennen wir zeigen, dass der Vektorraum $\mathbb{R}^n$ ist endlich-dimensional ist.
- (3) For $n \geq 1$ ist $M_{n \times n}(K)$ endlich-dimensional.
- (4) Es sei

$$U = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2 \times 2}(\mathbb{R}) : a = -d \right\}.$$

Dann ist U ein endlich dimensionaler Unterraum von $M_{2 \times 2}(\mathbb{R})$. (Uebung: beweisen Sie das, und finden Sie in beiden Faelle ein endliches Erzeugendessystem.)

- (5) Es sei V der Vektorraum aller unendlichen reellen Folgen

$$V = \{(a_0, a_1, a_n \dots) : a_i \in \mathbb{R} \forall i\},$$

wobei die Addition und Skalarmultiplikation wie in Kapitel 1.1 definiert sind. Wir werden in Beispiel 3.3.25 beweisen, dass V nicht endlich-dimensional ist. Allerdings haben wir schon gesehen, dass der Unterraum der Fibonacci-Folgen endlich-dimensional ist: die Folgen $\mathcal{F}_{0,1}$ und $\mathcal{F}_{1,0}$ sind ein Erzeugendessystem (Satz 1.1.9).

Beispiel 3.3.5. Sind die Vektoren

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

ein Erzeugendensystem von $\mathbb{R}^3$? Mit anderen Worten, hat das lineare Gleichungssystem

$$(11) \quad \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

eine Lösung fuer alle Vektoren $\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \in \mathbb{R}^3$? Die reduzierte Zeilenstufenform der erweiterten Matrix ist

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & b_1 \\ 0 & 1 & 1 & b_1 + b_2 \\ 0 & 0 & 0 & b_3 + b_1 - 2b_2 \end{array} \right).$$

Mit anderen Worten, (11) hat genau dann eine Lösung, wenn $b_3 = 2b_2 + b_1$; ansonsten hat es keine Lösung. Das bedeutet, dass z.B. der Vektor $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ nicht in der linearen Huelle von v_1, v_2, v_3 enthalten ist, d.h. v_1, v_2, v_3 ist kein Erzeugendensystem von $\mathbb{R}^3$.

Die Definition 3.3.3 wirft folgende Fragen auf:

- Wie (wenn ueberhaupt) koennen wir die Dimension eines Vektorraums definieren?
- Wenn V ein endlich-dimensionaler Vektorraum ist, ist dann auch jeder Unterraum von V endlich-dimensional?

Zurueck zum Beispiel 3.3.2: das Erzeugendensystem e_1, e_2, e_3 hat eine wichtige Eigenschaft, naemlich dass jedes Element in $\mathbb{R}^3$ *eindeutig* als eine Linearkombination der Vektoren dargestellt werden kann: es

sei $v = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$. Dann gilt

$$v = ae_1 + be_2 + ce_3,$$

und es gibt keine andere Moeglichkeit, v als Linearkombination der Vektoren e_1, e_2, e_3 zu schreiben.

Nun ist e_1, e_2, e_3, f mit $f = e_1 + e_2$ ebenfalls ein Erzeugendensystem: allerdings hat es den Nachteil, dass sich Vektoren in $\mathbb{R}^3$ nicht mehr eindeutig als Linearkombination der erzeugenden Vektoren schreiben lassen: z.B.

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0e_1 + 0e_2 + 0e_3 = e_1 + e_2 - f.$$

Definition 3.3.6. Es sei V ein Vektorraum, und es seien $v_1, \dots, v_n \in V$. Dann sind $v_1, \dots, v_n$ linear unabhaengig, wenn sich 0_V eindeutig als Linearkombination dieser Vektoren darstellen laesst: naemlich

$$0_V = 0e_1 + \dots + 0v_n.$$

Mit anderen Worten, $v_1, \dots, v_n$ sind linear unabhaengig wenn gilt

$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0 \quad \Rightarrow \quad \alpha_1 = \dots = \alpha_n = 0.$$

Wenn sie nicht linear unabhaengig sind, dann nennen wir sie linear abhaengig.

Bemerkung 3.3.7. Wir koennen diese Definition of unendliche Erzeugendensystem verallgemeinern: es sei $S \subseteq V$. Dann ist S linear unabhaengig wenn jede endliche Teilmenge von S linear unabhaengig ist. Zu ist z.B. die Menge $\{e_i : i \geq 1\}$ aus Beispiel 3.2.7 linear unabhaengig.

Wenn $v_1, \dots, v_n \in V$ linear unabhaengig sind, dann laesst sich *jeder* Vektor in der linearen Huelle (nicht nur der Nullvektor) von $v_1, \dots, v_n$ eindeutig als Linearkombination der Vektoren darstellen:

Satz 3.3.8. Es seien $v_1, \dots, v_n \in V$ linear unabhaengig. Dann kann jeder Vektor $v \in \langle v_1, \dots, v_n \rangle$ eindeutig als Linearkombination von $v_1, \dots, v_n$ dargestellt werden.

Proof. Nimm an, es gibt $v \in \langle v_1, \dots, v_m \rangle$ so dass

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n = \beta_1 v_1 + \dots + \beta_n v_n$$

für $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in K$. Dann gilt

$$\begin{aligned} & (\alpha_1 v_1 + \dots + \alpha_n v_n) - (\beta_1 v_1 + \dots + \beta_n v_n) = 0 \\ \Leftrightarrow & (\alpha_1 - \beta_1) v_1 + \dots + (\alpha_n - \beta_n) v_n = 0_V \\ \Rightarrow & \alpha_1 - \beta_1 = \dots = \alpha_n - \beta_n = 0, \end{aligned}$$

da $v_1, \dots, v_n$ linear unabhängig sind. Daher gilt

$$\alpha_1 = \beta_1, \quad \alpha_2 = \beta_2 \quad \dots \quad \alpha_n = \beta_n,$$

was zu beweisen war. □

Beispiele 3.3.9.

- (1) Es sei $V = \mathbb{R}^2$, und wir betrachten die Vektoren

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

Dann sind v_1, v_2, v_3 linear abhängig: $2v_1 - v_2 - v_3 = 0$. Hingegen sind die Vektoren v_1, v_2 linear unabhängig: die einzige Lösung des linearen Gleichungssystems

$$\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\text{ist } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

- (2) Es sei $V = \mathbb{R}^3$, und wir betrachten die Spaltenvektoren

$$v_1 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -9 \\ 4 \\ 6 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 4 \\ -1 \\ 5 \end{pmatrix}.$$

Sind diese Vektoren linear unabhängig? Mit anderen Worten, hat die Gleichung

$$(12) \quad x_1 v_1 + x_2 v_2 + x_3 v_3 + x_4 v_4 = 0$$

eine Lösung außer $x_1 = x_2 = x_3 = x_4 = 0$? Wir haben diese Frage bereits in Beispiel 2.3.5 beantwortet: $x_4 \in \mathbb{R}$ beliebig und

$$x_1 = \frac{17}{3} x_4, \quad x_2 = \frac{5}{3} x_4, \quad x_3 = \frac{11}{3} x_4$$

ist eine Lösung von (12). Daher sind die Vektoren linear abhängig.

- (3) Generell kann man zeigen, dass je drei Vektoren in $\mathbb{R}^2$ (allgemeiner: je $n+1$ Vektoren in $\mathbb{R}^n$) linear abhängig sind – wir werden später verstehen, warum dies so ist.
 (4) Per Konvention ist die leere Menge $\emptyset$ linear unabhängig, und es gilt $\langle \emptyset \rangle = \{0\}$.

Lemma 3.3.10. *Es seien $v_1, \dots, v_n \in V$.*

- Wenn zwei der Vektoren $v_1, \dots, v_n$ gleich sind, oder wenn ein Vektor ein Vielfaches von einem anderen Vektor ist, dann sind $v_1, \dots, v_n$ linear abhängig.
- Wenn $v_1, \dots, v_n$ linear unabhängig und $c_1, \dots, c_n \in K$ alle $\neq 0$ sind, dann sind auch $c_1 v_1, \dots, c_n v_n$ linear unabhängig.
- Wenn einer der Vektoren der Nullvektor 0_V ist, dann ist $v_1, \dots, v_n$ linear abhängig.

Proof. Übung. □

Lemma 3.3.11.

- (1) Es sei V ein Vektorraum und $v \in V, v \neq \{0\}$. Dann ist v linear unabhängig.
 (2) Der Nullvektor 0_V ist linear abhängig.

Proof. Klar. □

Satz 3.3.12. *Es seien $v_1, \dots, v_n \in V$ linear unabhängig, und es sei v_{n+1} ein Vektor in V , der nicht in $\langle v_1, \dots, v_n \rangle$ enthalten ist. Dann sind $v_1, \dots, v_n, v_{n+1}$ linear unabhängig.*

Proof. Nimm an, dass $v_1, \dots, v_n, v_{n+1}$ linear abhaengig sind, d.h. es gibt $\alpha_1, \dots, \alpha_{n+1} \in K$, nicht alle null, so dass

$$\alpha_1 v_1 + \dots + \alpha_n v_n + \alpha_{n+1} v_{n+1} = 0.$$

Dann ist $\alpha_{n+1} \neq 0$, da $v_1, \dots, v_n$ linear unabhaengig sind. Dann aber folgt, dass

$$v_{n+1} = -\alpha_{n+1}^{-1} (\alpha_1 v_1 + \dots + \alpha_n v_n),$$

d.h. $v_{n+1} \in \langle v_1, \dots, v_n \rangle$, was ein Widerspruch ist. Daher ist unsere Annahme falsch. $\square$

Definition 3.3.13. *Es sei V ein Vektorraum. Eine Basis von V ist ein linear unabhaengiges Erzeugendensystem von V .*

Beispiele 3.3.14.

- (1) e_1, e_2, e_3 ist eine Basis von $\mathbb{R}^3$; $S = \{e_1, e_2, e_3, f\}$ mit $f = e_1 - e_2$ ist hingegen keine Basis: so haben wir

$$e_1 = f + e_2,$$

d.h. die Darstellung des Vektors e_1 als Linearkombination der Vektoren in S ist nicht eindeutig.

- (2) $\mathcal{F}_{1,\varphi}$ und $\mathcal{F}_{1,\psi}$ sind eine Basis des Vektorraums von Fibonacci Folgen. Die Folgen $\mathcal{F}_{0,1}$ und $\mathcal{F}_{1,0}$ sind ebenfalls eine Basis.
 (3) Es seien e_{11} , e_{12} und e_{22} wie in Beispiel 3.2.6. Dann sind diese Vektoren eine Basis fuer den Unterraum U der oberen Dreiecksmatrizen in $M_{2 \times 2}(\mathbb{R})$.

Hat jeder Vektorraum eine Basis?

Theorem 3.3.15. *Es sei V ein endlich-dimensionaler Vektorraum. Dann enthaelt jedes endliche Erzeugendensystem eine Basis von V .*

Proof. Es sei S ein endliches Erzeugendensystem. Waehle eine Untermenge $T \subseteq S$, die linear unabhaengig ist und die groesstmoeegliche Anzahl von Elementen enthaelt. Es gibt zwei Moeglichkeiten:

- Wenn die lineare Huelle von T gleich V ist, dann ist T eine Basis.
- Wenn $\langle T \rangle \neq V = \langle S \rangle$, dann koennen wir ein $v \in S$ waehlen, dass nicht in $\langle T \rangle$ enthalten ist. Dann ist $\{v\} \cup S$ linear unabhaengig aufgrund von Satz 3.3.12, was ein Widerspruch zur Definition von T ist. Daher kann es ein solches v nicht geben, d.h. $\langle T \rangle = V$.

$\square$

Korollar 3.3.16. *Jeder endlich-dimensionale Vektorraum besitzt eine Basis.*¹⁵

Es liegt nahe, die Dimension eines Vektorraums als die Anzahl der Element einer Basis zu definieren. Aber ist das wohldefiniert? Koennnte ein endlich-dimensionaler Vektorraum nicht zwei Basen mit unterschiedlich vielen Elementen haben?

Lemma 3.3.17. (Austauschlemma) *Es sei V ein endlich-dimensionaler K -Vektorraum, und $S = \{v_1, \dots, v_n\}$ eine Basis. Es sei $w \in V$ nicht null, mit der Linearkombination*

$$(13) \quad w = \alpha_1 v_1 + \dots + \alpha_n v_n$$

mit $\alpha_i \in K$. Wenn $\alpha_j \neq 0$, dann ist auch

$$S' = \{v_1, \dots, v_{j-1}, w, v_{j+1}, \dots, v_n\}$$

eine Basis von V .

Proof. Wir muessen zeigen, dass S' ein Erzeugendensystem und linear unabhaengig ist.

- Erzeugendensystem: wir zeigen zunaechst, dass $v_j \in \langle v_1, \dots, v_{j-1}, w, v_{j+1}, \dots, v_n \rangle$. Von Gleichung (13) sehen wir, dass

$$(14) \quad \begin{aligned} \alpha_j v_j &= w - \sum_{i \neq j} \alpha_i v_i \\ \Rightarrow v_j &= \frac{1}{\alpha_j} w - \sum_{i \neq j} \frac{\alpha_i}{\alpha_j} v_i. \end{aligned}$$

¹⁵Es sei $V = \{0\}$. Dann ist $\emptyset$ eine Basis von V . Der Nullvektor ist ein Erzeugendensystem von V , aber keine Basis, da er nicht linear unabhaengig ist.

Es sei $u \in V$. Da S eine Basis ist, ist es insbesondere ein Erzeugendensystem, und so gibt es $\beta_1, \dots, \beta_n \in K$ so dass

$$u = \beta_1 v_1 + \dots + \beta_n v_n.$$

Indem wir Gleichung (14) dort einsetzen, erhalten wir einen Ausdruck fuer u als Linearkombination von $v_1, \dots, v_{j-1}, w, v_{j+1}, \dots, v_n$:

$$u = \frac{\beta_j}{\alpha_j} w + \sum_{i \neq j} \left(\beta_i - \beta_j \frac{\alpha_i}{\alpha_j} \right) v_i,$$

d.h. S' ist ein Erzeugendensystem.

- Lineare Unabhaengigkeit: Nimm an, dass es Skalare $\gamma_1, \dots, \gamma_n$ gibt, nicht alle gleich 0, so dass

$$\gamma_1 v_1 + \dots + \gamma_{j-1} v_{j-1} + \gamma_j w + \gamma_{j+1} v_{j+1} + \dots + \gamma_n v_n = 0.$$

Wir setzen Gleichung (13) fuer w dort ein und erhalten

$$\alpha_j \gamma_j v_j + \sum_{i \neq j} (\gamma_i + \gamma_j \alpha_i) v_i = 0.$$

Da $v_1, \dots, v_n$ linear unabhaengig sind, gilt daher

$$\alpha_j \gamma_j = 0 \quad \text{und} \quad \gamma_i + \gamma_j \alpha_i = 0 \quad \forall i \neq j.$$

Nun ist $\alpha_j \neq 0$ und daher $\gamma_j = 0$, was impliziert, dass ebenfalls $\gamma_i = 0$ fuer alle $i \neq j$. □

Beispiele 3.3.18.

- (1) Wir wissen aus Beispiel 3.3.14 (1), dass e_1, e_2, e_3 eine Basis von $\mathbb{R}^3$ ist. Indem wir Lemma 3.3.17 mehrfach anwenden, sehen wir, dass $e_1 + e_2, e_2 + e_3, e_3 + e_1$ ebenfalls eine Basis ist.
- (2) Wir verstehen jetzt, warum sowohl $\mathcal{F}_{1,\varphi}, \mathcal{F}_{1,\psi}$ als auch $\mathcal{F}_{1,0}, \mathcal{F}_{0,1}$ Basen des Raumes aller Fibonacci Folgen sind: von Satz 1.1.9 wissen wir, dass

$$\mathcal{F}_{1,\varphi} = \mathcal{F}_{1,0} + \varphi \mathcal{F}_{0,1} \quad \text{und} \quad \mathcal{F}_{1,\psi} = \mathcal{F}_{1,0} + \psi \mathcal{F}_{0,1}.$$

Daher folgt aus dem Austauschlemma, dass $\mathcal{F}_{1,\varphi}, \mathcal{F}_{1,\psi}$ ebenfalls eine Basis ist.

- (3) Es sei V ein endlich-dimensionaler $\mathbb{R}$ -Vektorraum, der nicht der Null-Vektorraum ist. Dann hat V unendlich viele verschiedene Basen.

Theorem 3.3.19. (Austauschsatz) *Es sei V ein endlich-dimensionaler Vektorraum, und es sei $v_1, \dots, v_n$ eine Basis von V . Es seien $w_1, \dots, w_k$ linear unabhaengiger Vektoren in V . Dann gilt $k \leq n$, und es gibt $(n - k)$ Basisvektoren, die zusammen mit $w_1, \dots, w_k$ eine Basis von V bilden.*

Proof. Wir beweisen den Satz ueber Induktion nach k .

$k = 1$: Das ist Lemma 3.3.17.

Wir nehmen nun an, dass der Satz fuer k gilt. Es seien $w_1, \dots, w_{k+1}$ linear unabhaengiger Vektoren in V . Dann sind auch $w_1, \dots, w_k$ linear unabhaengig, und es gilt $k < n$ und es gibt $(n - k)$ Basisvektoren, die zusammen mit $w_1, \dots, w_k$ eine Basis von V bilden. (Beachten Sie, dass der Fall $k = n$ nicht eintreten kann: dann waeren $w_1, \dots, w_k$ selber eine Basis, und $w_1, \dots, w_{k+1}$ koennen nicht linear unabhaengig sein.)

Dann koennen wir w_{k+1} als eine Linearkombination dieser Vektoren schreiben:

$$w_{k+1} = \alpha_1 w_1 + \dots + \alpha_k w_k + \alpha_{k+1} v_{k+1} + \dots + \alpha_n v_n$$

mit $\alpha_i \in K$ nicht alle null. Da $w_1, \dots, w_{k+1}$ linear unabhaengig sind, ist w_{k+1} nicht in dem von $w_1, \dots, w_k$ erzeugten Unterraum enthalten. Es gibt also einen Index j , $k+1 \leq j \leq n$, so dass $\alpha_j \neq 0$. Dann koennen wir nach Lemma 3.3.17 den Vektor v_j gegen den Vektor w_{k+1} austauschen und erhalten eine Basis

$$w_1, \dots, w_k, v_{k+1}, \dots, v_{j-1}, w_{k+1}, v_{j+1}, \dots, v_n.$$

□

Korollar 3.3.20. *Es sei V ein endlich-dimensionaler Vektorraum, und es seien $v_1, \dots, v_n$ und $w_1, \dots, w_m$ Basen von V . Dann gilt $m = n$. Mit anderen Worten, alle Basen von V haben die gleiche Anzahl von Elementen.*

Wir koennen nun endlich die Dimension eines Vektorraums definieren:

Definition 3.3.21. Es sei V ein endlich-dimensionaler Vektorraum. Die Dimension von V ist die Anzahl von Elementen einer Basis von V ; wir schreiben $\dim_K V$. Wenn V nicht endlich-dimensional ist, dann schreiben wir $\dim_K V = \infty$.

Beispiele 3.3.22.

- (1) $\mathbb{R}^3$ ist ein 3-dimensionaler $\mathbb{R}$ -Vektorraum. Allgemeiner: $\mathbb{R}^n$ ist ein n -dimensionaler $\mathbb{R}$ -Vektorraum; die *Standardbasis* ist gegeben durch die Spaltenvektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

- (2) Der Raum aller Fibonacci Folgen ist 2-dimensional.
 (3) Es sei K ein Koerper und $V = K[x]^{\leq n}$. Dann ist $\dim_K V = n + 1$.
 (4) $\mathbb{C}$ ist ein 1-dimensionaler $\mathbb{C}$ -Vektorraum (was ist eine Basis?), aber ein 2-dimensionaler $\mathbb{R}$ -Vektorraum (mit Basis $1, i$).¹⁶
 (5) Es gilt $\dim_K \{0\} = 0$, da $\emptyset$ eine Basis von $\{0\}$ ist und $|\emptyset| = 0$. Der Null-Vektorraum ist der einzige Vektorraum der Dimension 0.
 (6) Es sei

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x + y + z = 0 \right\}.$$

Dann ist U ein zwei-dimensionaler Unterraum von $\mathbb{R}^3$, mit Basis $\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Theorem 3.3.19 hat ein paar schoene und interessante Konsequenzen:

Satz 3.3.23. Es sei V ein endlich-dimensionaler K -Vektorraum mit Dimension n . Es seien $v_1, \dots, v_n \in V$. Dann sind folgende Aussagen aequivalent:

- (i) $v_1, \dots, v_n$ sind linear unabhaengig;
- (ii) $v_1, \dots, v_n$ sind ein Erzeugendensystem von V ;
- (iii) $v_1, \dots, v_n$ sind eine Basis von V .

Proof. (iii) $\Rightarrow$ (i),(ii) ist klar von der Definition.

(i) $\Rightarrow$ (iii): Aufgrund von Theorem 3.3.19 koennen wir $\{v_1, \dots, v_n\}$ zu einer Basis von V erweitern. Aber jede Basis von V hat n Elemente, daher ist $\{v_1, \dots, v_n\}$ selber eine Basis.

(ii) $\Rightarrow$ (iii): Aus Theorem 3.3.15 folgt, dass $\{v_1, \dots, v_n\}$ eine Basis enthaelt. Aber jede Basis von V hat n Elemente, daher ist $\{v_1, \dots, v_n\}$ selber eine Basis. $\square$

Satz 3.3.24. Es sei V ein n -dimensionaler Vektorraum, und es seien $v_1, \dots, v_k \in V$.

- (i) Wenn $k < n$, dann sind $v_1, \dots, v_k$ kein Erzeugendensystem fuer V .
- (ii) Wenn $k > n$, dann sind $v_1, \dots, v_k$ linear abhaengig.

Proof. (i) Nimm an, dass $v_1, \dots, v_k$ ein Erzeugendensystem fuer V ist. Dann enthaelt es eine Basis (Theorem 3.3.15). Da aber jede Basis n Elemente enthaelt, erhalten wir einen Widerspruch.

(ii) Nimm an, dass $v_1, \dots, v_k$ linear unabhaengig sind. Dann koennen wir $\{v_1, \dots, v_k\}$ zu einer Basis von V erweitern (Theorem 3.3.19). Aber jede Basis enthaelt n Elemente, was wiederum einen Widerspruch ergibt. $\square$

Beispiele 3.3.25.

- (1) Zurueck zum Beispiel 3.3.9 (3): es folgt direkt von Satz 3.3.24, dass alle $n + 1$ Vektoren in $\mathbb{R}^n$ linear abhaengig sind.
 (2) Es sei V der Vektorraum aller reellen Folgen. Wir koennen jetzt beweisen, dass V nicht endlich-dimensional ist: fuer $n \geq 1$ sei $\mathcal{F}_n$ die Folge, deren n ter Wert eine 1 ist und alle anderen Werte null sind. Dann ist die unendliche Menge $\{\mathcal{F}_n : n \geq 1\}$ linear unabhaengig. Es folgt aus Proposition 3.3.24 (ii), dass V nicht endlich-dimensional sein kann.

¹⁶Man kann zeigen, dass $\mathbb{R}$ ein unendlich-dimensional $\mathbb{Q}$ -Vektorraum ist. Diese haengt eng mit der sogenannten *Kontinuumshypothese* zusammen.

3.4. Basen von Unterräumen.

Beachte 3.4.1. Es sei V ein Vektorraum und U ein Unterraum von V , und es seien $u_1, \dots, u_n \in U$. Wenn $u_1, \dots, u_n$ linear unabhängig sind als Elemente von U , dann sind sie ebenfalls linear unabhängig als Elemente von V .

Satz 3.4.2. Es sei V ein endlich-dimensionaler K -Vektorraum, und es sei U ein Unterraum von V . Dann ist U endlich-dimensional, und

$$\dim_K U \leq \dim_K V,$$

mit Gleichheit dann und nur dann, wenn $U = V$.

Proof. Es sei $n = \dim_K V$. Wenn $U = \{0\}$, dann ist $\dim_K U = 0$ (Beispiel 3.3.22 (4)), und die Aussage folgt.

Nimm also an, dass $U \neq \{0\}$. Sei $u_1 \in U$, $u_1 \neq 0_V$. Dann gibt es zwei Möglichkeiten: entweder ist $U = \langle u_1 \rangle$: in diesem Fall ist u_1 eine Basis fuer U , d.h. $\dim_K U = 1$, und die Aussage folgt. Oder $\langle u_1 \rangle \subsetneq U$: dann wähle $u_2 \in U - \langle u_1 \rangle$. Es folgt von Satz 3.3.12, dass u_1, u_2 linear unabhängig sind.

Dann gibt es wieder zwei Möglichkeiten: entweder ist $U = \langle u_1, u_2 \rangle$; in diesem Fall sind u_1, u_2 eine Basis von U . Oder $\langle u_1, u_2 \rangle \subsetneq U$; in diesem Fall wählen wir ein $u_3 \in U - \langle u_1, u_2 \rangle$ usw.

Aufgrund von Satz 3.3.24 muss dieser Prozess nach höchstens n Schritten enden. Wir erhalten also eine Liste $u_1, \dots, u_k$ fuer $k \leq n$ von linear unabhängigen Vektoren so dass $U = \langle u_1, \dots, u_k \rangle$. Es folgt von der Definition, dass $u_1, \dots, u_k$ eine Basis von U sind. Daher gilt

$$\dim_K U = k \leq n = \dim_K V.$$

Wenn $U = V$, dann gilt natuerlich $k = n$. Nimm umgekehrt an, dass $k = n$. Die Vektoren $u_1, \dots, u_k$ sind eine Basis fuer U und daher linear unabhängig; sie sind ebenfalls linear unabhängig als Elemente von V (Note 3.4.1). Dann folgt von Satz 3.3.23, dass sie eine Basis von V sind, und daher $U = V$. $\square$

Wir haben in Theorem 3.2.10 gesehen, dass sich aus zwei Unterräumen U, W von V die Unterräume $U \cap W$ und $U + W$ bilden lassen. Was koennen wir ueber deren Dimensionen sagen? Die naive Vermutung, dass $\dim(U + W) = \dim U + \dim W$ gilt, ist (im allgemeinen) falsch: wenn beispielsweise $U = W \subsetneq V$ und $U \neq \{0\}$, dann gilt

$$\dim(U + W) = \dim U < 2 \dim U.$$

Sehen wir uns ein weiteres Beispiel an, um die Phaenomene, die auftreten koennen, besser zu verstehen:

Beispiel 3.4.3. Betrachte den Vektorraum $\mathbb{R}^4$ mit der Standardbasis e_1, e_2, e_3, e_4 .

- Es seien $U = \langle e_1, e_2 \rangle$ und $W = \langle e_3 \rangle$. Dann ist

$$U + W = \langle e_1, e_2, e_3 \rangle,$$

die Vektoren e_1, e_2, e_3 sind linear unabhängig, und daher gilt

$$\dim(U + W) = 3 = 2 + 1 = \dim U + \dim W.$$

- Es seien nun

$$U' = \langle e_2, e_3 \rangle \quad \text{und} \quad W' = \langle e_2 + e_3, e_4 \rangle.$$

Insbesondere gilt $\dim U' = \dim W' = 2$. Per Definition gilt

$$U' + W' = \langle e_2, e_3, e_2 + e_3, e_4 \rangle.$$

Nun sind die Vektoren $e_1, e_2, e_3, e_2 + e_3, e_4$ offensichtlich nicht linear unabhängig: $e_2 + e_3 \in \langle e_2, e_3, e_4 \rangle$, d.h.

$$U' + W' = \langle e_2, e_3, e_4 \rangle,$$

und wir wissen, dass die Vektoren e_2, e_3, e_4 linear unabhängig sind. Daher ist $U' + W'$ 3-dimensional. Was ist passiert?

Der Grund, warum $\dim(U' + W') < \dim U' + \dim W'$, ist der, dass die Schnittmenge von U' und W' nicht der Null-Vektorraum ist: der Vektor $e_2 + e_3$ ist sowohl in U' als auch in W' enthalten, und er ist eine Basis fuer $U' \cap W'$ (Uebung), d.h. $\dim U' \cap W' = 1$. Haben Sie eine Vermutung, wie die allgemeine Formel fuer $\dim(U' + W')$ lautet?

Theorem 3.4.4. Es sei V ein endlich-dimensionaler Vektorraum, und es seien U, W Unterräume von V . Dann gilt

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

Proof. Es seien $\dim_K U = \ell$, $\dim_K W = m$ und $\dim_K(U \cap W) = k$. Wähle eine Basis $v_1, \dots, v_k$ von $U \cap W$. Aufgrund von Theorem 3.3.19 können wir diese Basis zu Basen

$$v_1, \dots, v_k, u_1, \dots, u_{\ell-k}, \quad \text{bzw.} \quad v_1, \dots, v_k, w_1, \dots, w_{m-k}$$

von U , bzw. von W erweitern. Wir behaupten nun, dass

$$v_1, \dots, v_k, u_1, \dots, u_{\ell-k}, w_1, \dots, w_{m-k}$$

eine Basis ist von $U + W$.

Da $U + W = \{u + w : u \in U, w \in W\}$ ist klar, dass

$$U + W = \langle v_1, \dots, v_k, u_1, \dots, u_{\ell-k}, w_1, \dots, w_{m-k} \rangle.$$

Wir zeigen die lineare Unabhängigkeit. Nimm an, dass

$$\alpha_1 v_1 + \dots + \alpha_k v_k + \beta_1 u_1 + \dots + \beta_{\ell-k} u_{\ell-k} + \gamma_1 w_1 + \dots + \gamma_{m-k} w_{m-k} = 0_V.$$

Dann folgt, dass

$$\alpha_1 v_1 + \dots + \alpha_k v_k + \beta_1 u_1 + \dots + \beta_{\ell-k} u_{\ell-k} = -(\gamma_1 w_1 + \dots + \gamma_{m-k} w_{m-k}).$$

Beachte nun, dass die linke Seite dieser Gleichung in U ist, die rechte aber in W . Schreibe $x = \alpha_1 v_1 + \dots + \alpha_k v_k + \beta_1 u_1 + \dots + \beta_{\ell-k} u_{\ell-k}$. Dann gilt $x \in U \cap W$. Da $v_1, \dots, v_k$ eine Basis ist für $U \cap W$, gibt es $\delta_1, \dots, \delta_k \in K$, so dass

$$x = \delta_1 v_1 + \dots + \delta_k v_k.$$

Aber ebenfalls gilt $x = -(\gamma_1 w_1 + \dots + \gamma_{m-k} w_{m-k})$, so dass

$$\delta_1 v_1 + \dots + \delta_k v_k + \gamma_1 w_1 + \dots + \gamma_{m-k} w_{m-k} = 0_V.$$

Da $v_1, \dots, v_k, w_1, \dots, w_{m-k}$ eine Basis von W ist, folgt daraus, dass

$$\delta_1 = \dots = \delta_k = \gamma_1 = \dots = \gamma_{m-k} = 0.$$

Daher gilt

$$\alpha_1 v_1 + \dots + \alpha_k v_k + \beta_1 u_1 + \dots + \beta_{\ell-k} u_{\ell-k} = 0_V.$$

Nun ist aber $v_1, \dots, v_k, u_1, \dots, u_{\ell-k}$ eine Basis von U , daher folgt

$$\alpha_1 = \dots = \alpha_k = \beta_1 = \dots = \beta_{\ell-k} = 0.$$

□

Warnung 3.4.5. Es ist **nicht wahr**, dass für drei Unterräume U, W, X von V folgende Formel gilt:¹⁷

$$\dim(U + W + X) = \dim U + \dim W + \dim X - \dim(U \cap W) - \dim(U \cap X) - \dim(W \cap X) + \dim(U \cap W \cap X).$$

Finden Sie ein Gegenbeispiel, indem Sie geeignete Unterräume von $\mathbb{R}^2$ betrachten. Hier ist ein Gegenbeispiel: es sei $V = \mathbb{R}^2$, und wir betrachten die drei ein-dimensionalen Unterräume

$$(15) \quad U = \langle e_1 \rangle, \quad W = \langle e_2 \rangle, \quad X = \langle e_1 + e_2 \rangle.$$

Dann ist $U + W + X = \mathbb{R}^2$, und

$$U \cap W = W \cap X = X \cap U = U \cap W \cap X = \{0\},$$

d.h. in (15) erhalten wir $2 = 3$.¹⁸

Philipp Steiner und Marco Vaccaro haben eine schöne Formel für den Schnitt von beliebig vielen Unterräumen gefunden; sie ist allerdings nicht symmetrisch in den A_i 's.

Korollar 3.4.6. Es sei V ein endlich-dimensionaler Vektorraum, und es seien U, W Unterräume von V . Dann sind folgende Aussagen äquivalent:

- (i) $\dim(U + W) = \dim U + \dim W$;
- (ii) $\dim(U \cap W) = 0$;
- (iii) $U \cap W = \{0_V\}$.
- (iv) für jedes $v \in U + W$ gibt es genau ein $u \in U$ und $w \in W$ so dass $v = u + w$;
- (v) die Gleichung $u + w = 0_V$ mit $u \in U$ und $w \in W$ hat $u = w = 0_V$ als einzige Lösung.

¹⁷Selbst viele professionelle Mathematiker wissen das nicht!

¹⁸Das Problem ist, dass im allgemeinen

$$(U + W) \cap X \neq U \cap X + W \cap X.$$

Proof. (i) $\Leftrightarrow$ (ii): folgt direkt von Theorem 3.4.4.

(ii) $\Leftrightarrow$ (iii): klar, da es genau einen Vektorraum der Dimension null gibt.

(iii) $\Rightarrow$ (iv) Es sei $v \in V$, und nimm an, dass es $u, u' \in U$ und $w, w' \in W$ gibt, so dass

$$v = u + w = u' + w'.$$

Dann gilt $u - u' = w' - w$. Nun ist $u - u' \in U$ und $w' - w \in W$, d.h.

$$u - u', w' - w \in U \cap W.$$

Da $U \cap W = \{0_V\}$, folgt dass $u = u'$ und $w = w'$.

(iv) $\Rightarrow$ (v): die Gleichung $u + w = 0_V$ hat die Lösung $u = w = 0_V$, und da U, W Unterräume sind, enthalten beide das Element 0_V . Indem wir nun (iv) auf $v = 0_V$ anwenden, sehen wir, dass dieses die einzige Lösung ist.

(v) $\Rightarrow$ (iii): nimm an, dass (iii) nicht wahr ist, d.h. es gibt einen Vektor $v \in U \cap W$, der nicht der Nullvektor ist. Dann gilt

$$0_V = v + (-v),$$

und $v \in U$ und $-v \in W$, was einen Widerspruch zu (v) ist. $\square$

Definition 3.4.7. Es sei V ein K -Vektorraum und $U \leq V$. Ein Unterraum $W \leq V$ ist ein Komplement von U wenn $V = U + W$ und $U \cap W = \{0\}$.

Warnung. Es ist *nicht* wahr, dass das Komplement von $U \leq V$ gegeben ist durch $V - U$!

Satz 3.4.8. Es sei $U \leq V$. Dann gibt es einen Unterraum $W \leq V$, der ein Komplement zu U ist.

Proof. Es sei $u_1, \dots, u_\ell$ eine Basis von U . Aufgrund von Theorem 3.3.19 wissen wir, dass wir diese Basis von U zu einer Basis von V erweitern können: es gibt Vektoren $w_1, \dots, w_m \in V$, so dass $u_1, \dots, u_\ell, w_1, \dots, w_m$ eine Basis von V ist. Es sei $W = \langle w_1, \dots, w_m \rangle$. Dann ist W ein Komplement von U : da $u_1, \dots, u_\ell, w_1, \dots, w_m$ eine Basis von V ist, gilt $V = U + W$. Nimm nun an, dass $v \in U \cap W$. Dann gibt es Skalare $\alpha_1, \dots, \alpha_\ell$ und $\lambda_1, \dots, \lambda_m$ so dass

$$\begin{aligned} v &= \sum_{i=1}^{\ell} \alpha_i u_i = \sum_{j=1}^m \lambda_j w_j \\ \Rightarrow \quad \alpha_1 v_1 + \dots + \alpha_\ell v_\ell - (\lambda_1 w_1 + \dots + \lambda_m w_m) &= 0_V \\ \Rightarrow \quad \alpha_1 = \dots = \alpha_\ell = \lambda_1 = \lambda_m &= 0, \end{aligned}$$

da die Vektoren $u_1, \dots, u_\ell, w_1, \dots, w_m$ linear unabhängig sind. $\square$

Beispiel 3.4.9. Es sei $V = \mathbb{R}^2$ und $U = \langle e_1 \rangle$. Es sei $w \in V$ ein Vektor, so dass u, w linear unabhängig sind (d.h. w ist kein Vielfaches von e_1). Dann ist $W = \langle w \rangle$ ein Komplement von U .

Bemerkung 3.4.10. Das Komplement eines Unterraumes ist im allgemeinen nicht eindeutig bestimmt.

So ist in Beispiel 3.4.9 z.B. $\langle e_2 \rangle$ ein Komplement von U . Das Gleiche gilt für $\left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle$ oder für $\left\langle \begin{pmatrix} 3 \\ -1 \end{pmatrix} \right\rangle$.

3.5. Zusammenfassung der Rechenmethoden. Wir haben in den vorherigen Abschnitten gesehen, dass wir lineare Gleichungssysteme fuer die Analyse von Vektorraeumen benutzen koennen. Hier ist eine Zusammenfassung der bisherigen Ergebnisse: es sei $V = K^m$ und $v_1, \dots, v_n \in V$.

- (1) Gegeben sei $w \in V$. Ist $w \in \langle v_1, \dots, v_n \rangle$?

Antwort: $w \in \langle v_1, \dots, v_n \rangle$ ist aequivalent zu der Aussage: es gibt $\lambda_1, \dots, \lambda_n \in K$, so dass

$$(16) \quad w = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Schreiben wir $w = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix}$ und $v_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$ fuer $1 \leq j \leq n$, dann ist (16) aequivalent zu folgendem linearen Gleichungssystem:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix},$$

das wir mit Hilfe der reduzierten Zeilenspaltenform loesen koennen.

- (2) Koennen wir bestimmen, welche Vektoren in $\langle v_1, \dots, v_n \rangle$ enthalten sind?

Antwort: es sei $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ ein beliebiger Vektor in V . Dann betrachten wir das lineare Gleichungssystem

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

und bestimmen, fuer welche Vektoren b eine Loesung existiert.

- (3) Koennen wir bestimmen, ob $v_1, \dots, v_n$ linear unabhaengig sind?

Antwort: $v_1, \dots, v_n$ sind genau dann linear unabhaengig, wenn das lineare Gleichungssystem

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

nur die triviale Loesung $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ besitzt.

3.6. Zeilen und Spaltenräume.

Bemerkung 3.6.1. Wir können ein Element von K^n entweder als Zeilen- oder Spaltenvektor mit n Werten schreiben.

Definition 3.6.2. Es seien $m, n \geq 1$ und $A \in M_{m \times n}(K)$. Es seien $u_1, \dots, u_m \in K^n$ die Zeilen von A und $v_1, \dots, v_n \in K^m$ die Spalten von A . Wir definieren

$$\begin{aligned}\text{Zeilen}(A) &= \langle u_1, \dots, u_m \rangle \leq K^n, \\ \text{Spalten}(A) &= \langle v_1, \dots, v_n \rangle \leq K^m.\end{aligned}$$

Was können wir über die Dimensionen dieser Unterräume sagen?

Definition 3.6.3. Wir definieren

$$\begin{aligned}\text{Zeilenrang}(A) &= \dim_K \text{Zeilen}(A), \\ \text{Spaltenrang}(A) &= \dim_K \text{Spalten}(A).\end{aligned}$$

Beispiele 3.6.4.

(1) Für $A = \begin{pmatrix} 1 & -2 & 1 \\ -1 & 2 & -1 \end{pmatrix}$ gilt

$$\text{Zeilenrang}(A) = 1 = \text{Spaltenrang}(A).$$

(2) Für die Matrix $B = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$ gilt

$$\text{Zeilenrang}(B) = 2 = \text{Spaltenrang}(B).$$

(3) Für die Matrix $C = \begin{pmatrix} 1 & 2 & 2 & -2 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 3 & 0 \end{pmatrix}$ gilt

$$\text{Zeilenrang}(C) = 3 = \text{Spaltenrang}(C).$$

(4) Für die Matrix $D = \begin{pmatrix} 1 & 2 & 2 \\ 0 & -1 & 2 \\ 0 & 2 & -4 \end{pmatrix}$ gilt

$$\text{Zeilenrang}(D) = 2 = \text{Spaltenrang}(D).$$

(5) Für die Matrix $E = \begin{pmatrix} 1 & 2 \\ 0 & 0 \\ -3 & -6 \end{pmatrix}$ gilt

$$\text{Zeilenrang}(E) = 1 = \text{Spaltenrang}(E).$$

Was fällt Ihnen auf?

Hier ist das erste wirklich überraschende Resultat in diesem Kurs:¹⁹

Theorem 3.6.5. Es seien $m, n \geq 1$ und $A \in M_{m \times n}(K)$. Dann gilt

$$\text{Zeilenrang}(A) = \text{Spaltenrang}(A).$$

Für Matrizen in reduzierter Zeilenstufenform ist dieses Resultat offensichtlich.

Lemma 3.6.6. Theorem 3.6.5 gilt, wenn A in reduzierter Zeilenstufenform ist.

Proof. Aufgrund der Definition hat A folgende Form: [picture]

Es seien $j_1 < \dots < j_r$ die Nummern derjenigen Spalten, die eine führende Eins enthalten.

Es seien $u_1, \dots, u_r$ die Zeilenvektoren, die einen von Null verschiedenen Eintrag enthalten. Dann enthält jede dieser Zeilen eine führende 1, und die Zeilenvektoren sind linear unabhängig, weil $\forall 1 \leq k \leq r$ der erste von Null verschiedene Eintrag von u_k die führende 1 an der Stelle j_k ist, und $j_1 < \dots < j_r$. Mit anderen Worten,

$$\text{Zeilenrang}(A) = r.$$

¹⁹Ich gebe zu, dass ich es intuitiv immer noch nicht wirklich verstehe.

Was gilt fuer den Spaltenrang? Die Spaltenvektoren $v_{j_1}, \dots, v_{j_r}$ sind ein Erzeugendensystem von $\text{Spalten}(A)$; tatsaechlich sind sie die ersten r Standardvektoren $e_1, \dots, e_r$ der Standardbasis $e_1, \dots, e_m$ von K^m . Sie sind linear unabhaengig und daher eine Basis von $\text{Spalten}(A)$, d.h.

$$\text{Spaltenrang}(A) = r.$$

□

Der Beweis fuer eine beliebige Matrix A ist ziemlich kompliziert. Um die Aussage elegant beweisen zu koennen, brauchen wir die Theorie von linearen Abbildungen.

Bemerkung 3.6.7. *Lasse Raeuftlin, Sofia Espina und Ruben Leuzinger haben ein Beispiel gefunden, dass sich das Resultat nicht auf Wuerfel uebertragen laesst: betrachte den $(2 \times 2 \times 2)$ Wuerfel $A = (a_{ijk})$, fuer den*

$$A_{1,j,k} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \quad \text{und} \quad A_{2,j,k} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Dann sind der Zeilen- und Spaltenrang von $A_{1,j,k}$ gleich 1, aber der Zeilen- und Spaltenrang von $A_{i,1,k}$ gleich 2.

4.1. Definition und Beispiele.

Definition 4.1.1. Es seien V, W Vektorräume über K . Eine Funktion $T : V \rightarrow W$ ist eine lineare Abbildung, wenn sie folgende Bedingungen erfüllt:

- (i) es gilt $T(v_1 + v_2) = T(v_1) + T(v_2)$ für alle $v_1, v_2 \in V$;
- (ii) es gilt $T(\alpha v) = \alpha T(v)$ für alle $v \in V, \alpha \in K$.

Mit anderen Worten, T respektiert die Vektoraddition und Skalarmultiplikation.

Eine lineare Abbildung $T : V \rightarrow V$ ist ein Endomorphismus von V .

Bemerkung 4.1.2. Wir schreiben manchmal Tv anstatt $T(v)$.

Beispiele 4.1.3.

- (1) Es sei V ein Vektorraum. Die Identitätsabbildung $\text{id}_V : V \rightarrow V, v \mapsto v$ ist linear.
- (2) Es seien V, W Vektorräume. Die Null-Abbildung $V \rightarrow W, v \mapsto 0_W$ ist linear.
- (3) Es sei $K[x]^{\leq n}$ der Vektorraum aller Polynome vom Grad $\leq n$. Die Ableitungsabbildung

$$D : K[x]^{\leq n} \rightarrow K[x]^{\leq n}, \quad a_0 + a_1x + \dots + a_nx^n \mapsto a_1 + 2a_2x + \dots + na_nx^{n-1}$$

ist linear.

- (4) Es sei V der Vektorraum aller Fibonacci Folgen. Dann ist die Verschiebungsabbildung (c.f. Proposition 1.1.15) linear.
- (5) Es sei $n \geq 1$ und $V = M_{n \times n}(K)$. Definiere die Spur-Abbildung

$$\text{Tr} : V \rightarrow K, \quad A = (a_{ij})_{1 \leq i, j \leq n} \mapsto a_{11} + \dots + a_{nn}.$$

Dann ist Tr linear.

Das folgende Beispiel zeigt, dass Matrizen eine wichtige Quelle von linearen Abbildungen sind.

Definition 4.1.4. Es seien $m, n \geq 1$ und $A \in M_{m \times n}(K)$. Definiere die Abbildung

$$T_A : K^n \rightarrow K^m, \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Lemma 4.1.5. Die Abbildung T_A ist linear.

Proof. Übung. □

Bemerkung 4.1.6. Eine Abbildung $T : V \rightarrow W$ ist genau dann linear, wenn

$$T(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 T(v_1) + \alpha_2 T(v_2)$$

für alle $v_1, v_2 \in V$ und $\alpha_1, \alpha_2 \in K$.

Satz 4.1.7. Es seien V, W Vektorräume über K und $T : V \rightarrow W$ eine lineare Abbildung.

(i) Es seien $v_1, \dots, v_n \in V$ und $\alpha_1, \dots, \alpha_n \in K$. Dann gilt

$$T(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_1 T(v_1) + \dots + \alpha_n T(v_n).$$

(ii) Es gilt $T(0_V) = 0_W$.

Proof. (i) ist klar.

(ii) Es gilt

$$T(0_V) = T(0 \cdot 0_V) = 0 \cdot T(0_V) = 0_W,$$

wobei die letzte Gleichung von Satz 3.1.6 (2) folgt. □

Korollar 4.1.8. Es seien V, W Vektorräume über K , V endlich-dimensional, und es sei $v_1, \dots, v_n$ eine Basis von V . Es sei $T : V \rightarrow W$ eine lineare Abbildung. Dann ist T durch $T(v_1), \dots, T(v_n)$ eindeutig bestimmt.²⁰

²⁰Mit anderen Worten, wenn wir $T(v_1), \dots, T(v_n)$ kennen, dann kennen wir auch $T(v)$ für alle $v \in V$.

Proof. Es sei $v \in V$. Da $v_1, \dots, v_n$ eine Basis ist, schreiben wir v als Linearkombination $v = \sum_{i=1}^n \alpha_i v_i$. Dann gilt aufgrund von Satz 4.1.7 (i)

$$T(v) = \sum_{i=1}^n \alpha_i T(v_i),$$

was zu beweisen war. □

Tatsächlich koennen wir auf diese Weise lineare Abbildungen konstruieren:

Theorem 4.1.9. *Es seien V, W Vektorraeume ueber K , V endlich-dimensional, mit Basis $v_1, \dots, v_n$. Es seien $w_1, \dots, w_n \in W$. Dann gibt es genau eine lineare Abbildung $T : V \rightarrow W$ so dass*

$$T(v_i) = w_i \quad \forall 1 \leq i \leq n.$$

Proof. Wir definieren die Abbildung T wie folgt: es sei $v \in V$. Da $v_1, \dots, v_n$ eine Basis ist, koennen wir v eindeutig als Linearkombination $v = \sum_{i=1}^n \alpha_i v_i$ schreiben. Wir definieren

$$T(v) = \sum_{i=1}^n \alpha_i w_i.$$

Dann ist T wohl-definiert, da die Linearkombination eindeutig ist, und bei Definition gilt $T(v_i) = w_i$.

Ueberpruefen wir, dass T linear ist:

- Wenn $v, v' \in V$, schreiben wir

$$v = \sum_{i=1}^n \alpha_i v_i \quad \text{und} \quad v' = \sum_{i=1}^n \beta_i v_i.$$

Dann gilt

$$\begin{aligned} T(v + v') &= T\left(\sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^n \beta_i v_i\right) \\ &= T\left(\sum_{i=1}^n (\alpha_i + \beta_i) v_i\right) \\ &= \sum_{i=1}^n (\alpha_i + \beta_i) w_i \\ &= \sum_{i=1}^n \alpha_i w_i + \sum_{i=1}^n \beta_i w_i \\ &= T(v) + T(v'). \end{aligned}$$

- Wenn $v = \sum_{i=1}^n \alpha_i v_i \in V$ und $\beta \in K$, dann gilt

$$\begin{aligned} T(\beta v) &= T\left(\beta \sum_{i=1}^n \alpha_i v_i\right) \\ &= \sum_{i=1}^n \beta \alpha_i w_i \\ &= \beta \sum_{i=1}^n \alpha_i w_i \\ &= \beta T\left(\sum_{i=1}^n \alpha_i v_i\right) \\ &= \beta T(v). \end{aligned}$$

Die Eindeutigkeit von T folgt von Korollar 4.1.8. □

Warnung. Das Theorem ist im Allgemeinen falsch, wenn $v_1, \dots, v_n$ ein Erzeugendensystem aber nicht lienar unabhaengig ist. Betrachte z.B. das Erzeugendensystem $e_1, e_2, e_1 + e_2 \in \mathbb{R}^2$. Gibt es eine lineare Abbildung $T : \mathbb{R}^2 \rightarrow \mathbb{R}$, die jeden dieser Vektoren auf 1 abbildet? Nein, denn

$$0 = T(0_{\mathbb{R}^2}) = T(e_1 + e_1 - (e_1 + e_2)) = T(e_1) + T(e_2) - T(e_1 + e_2) = 1,$$

was einen Widerspruch ergibt.

Noch eine weitere Eigenschaft von linearen Abbildungen ist wichtig: lineare Abbildungen lassen sich verknuepfen.

Lemma 4.1.10. *Es seien $T : V \rightarrow U$ und $S : U \rightarrow W$ lineare Abbildungen. Dann ist die Abbildung*

$$S \circ T : V \rightarrow W$$

ebenfalls linear.

Proof. Es seien $v_1, v_2 \in V$ und $\alpha_1, \alpha_2 \in K$. Dann gilt

$$\begin{aligned} S \circ T(\alpha_1 v_1 + \alpha_2 v_2) &= S(T(\alpha_1 v_1 + \alpha_2 v_2)) \\ &= S(\alpha_1 T(v_1) + \alpha_2 T(v_2)) \\ &= \alpha_1 ST(v_1) + \alpha_2 ST(v_2) \\ &= \alpha_1 S \circ T(v_1) + \alpha_2 S \circ T(v_2). \end{aligned}$$

□

4.2. Kernel and Image.

Definition 4.2.1. Es seien V, W Vektorräume über K , und es sei $T : V \rightarrow W$ eine lineare Abbildung.

(1) Der Kern von T ist

$$\ker(T) = \{v \in V : T(v) = 0_W\} \subseteq V.$$

(2) Das Bild von T ist

$$\operatorname{im}(T) = \{T(v) : v \in V\} \subseteq W.$$

Beispiele 4.2.2.

(1) Es sei $A = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$, und es sei $T_A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ die dazugehörige lineare Abbildung

(Definition 4.1.4). Was ist $\ker(T_A)$? Per Definition ist

$$\ker(T_A) = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : Ax = 0_{\mathbb{R}^3} \right\},$$

d.h. $\ker(T_A)$ ist die Lösung eines linearen Gleichungssystems. Wir haben die Lösung dieses Gleichungssystems bereits in Beispiel 3.3.5 bestimmt: die Lösung ist

$$\ker(T_A) = \left\langle \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right\rangle.$$

Was ist $\operatorname{im}(T_A)$? Laut Definition ist

$$\operatorname{im}(T_A) = \left\{ b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} : \exists x \in \mathbb{R}^3 \text{ so dass } Ax = b \right\}.$$

Auch dieses haben wir in Beispiel 3.3.5 bereits bestimmt: $b \in \operatorname{im}(T_A)$ genau dann, wenn $b_3 = 2b_2 + b_1$, mit anderen Worten

$$\operatorname{im}(T_A) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \right\rangle.$$

(2) Es sei nun $B = \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 2 & 1 \end{pmatrix}$, und es sei $T_B : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ die dazugehörige lineare Abbildung. Eine einfache Rechnung zeigt, dass

$$\ker(T_B) = \{0_{\mathbb{R}^2}\}.$$

Lemma 4.2.3.

(1) $\ker(T)$ ist ein Unterraum von V .

(2) $\operatorname{im}(T)$ ist ein Unterraum von W .

Proof. Wir benutzen Satz 3.2.2.

(1) Es folgt von Satz 4.1.7, dass $T(0_V) = 0_W$, d.h. $0_V \in \ker(T)$. Es seien nun $u, v \in \ker(T)$ und $\lambda \in K$. Da T linear ist, gilt

$$T(u + \lambda v) = T(u) + \lambda T(v) = 0_W + \lambda 0_W = 0_W,$$

d.h. $u + \lambda v \in \ker(T)$.

(2) Da $T(0_V) = 0_W$, sehen wir, dass $0_W \in \operatorname{im}(T)$. Es seien nun $w_1, w_2 \in \operatorname{im}(T)$; dann gibt es $v_1, v_2 \in V$, so dass $T(v_i) = w_i$ für $i = 1, 2$. Es sei nun $\lambda \in K$. Dann gilt

$$w_1 + \lambda w_2 = T(v_1) + \lambda T(v_2) = T(v_1) + T(\lambda v_2) = T(v_1 + \lambda v_2),$$

d.h. $w_1 + \lambda w_2 \in \operatorname{im}(T)$. □

Beispiel 4.2.4. Es sei $A \in M_{m \times n}(K)$, und wir betrachten das homogene lineare Gleichungssystem

$$S : Ax = 0_{K^m}.$$

Dann ist $L(S) = \ker(T_A)$ ein Unterraum von K^n .

Satz 4.2.5. Es sei $T : V \rightarrow W$ eine lineare Abbildung. Wenn V endlich-dimensional ist, dann sind auch $\ker(T)$ und $\operatorname{im}(T)$ endlich-dimensional.²¹

Proof. Die Aussage fuer $\ker(T)$ folgt von Satz 3.4.2.

Um die Aussage fuer $\operatorname{im}(T)$ zu beweisen, benutzen wir Korollar 4.1.8: es sei $v_1, \dots, v_n$ eine Basis von V . Dann folgt aus dem Korollar, dass $T(v_1), \dots, T(v_n)$ ein Erzeugendensystem von $\operatorname{im}(T)$ ist. Daher ist $\operatorname{im}(T)$ endlich-dimensional. $\square$

Der Kern von T zeigt an, ob T eine injektive Abbildung ist:

Satz 4.2.6. Es sei $T : V \rightarrow W$ eine lineare Abbildung zwischen Vektorraeumen. Dann ist T genau dann injektiv, wenn $\ker(T) = \{0_V\}$.

Proof. $\Rightarrow$: klar von der Definition.

$\Leftarrow$: nimm an, dass $T(v) = T(u)$ fuer $u, v \in V$. Dann gilt

$$T(u) - T(v) = 0_W \quad \Leftrightarrow \quad T(u - v) = 0_W,$$

d.h. $u - v \in \ker(T)$. Wenn $\ker(T) = \{0_V\}$, dann folgt daraus, dass $u = v$, mit anderen Worten T ist injektiv. $\square$

Beispiel 4.2.7. Die Abbildung T_B aus Beispiel 4.2.2 (2) ist injektiv; die Abbildung T_A aus dem gleichen Beispiel (1) ist nicht injektiv.

Definition 4.2.8. Es sei $T : V \rightarrow W$ eine lineare Abbildung mit V endlich-dimensional. Der Rang $\operatorname{rk}(T)$ von T ist $\dim_K \operatorname{im}(T)$.

Der folgende Satz bringt $\operatorname{rk}(T)$ und die Dimension von $\ker(T)$ miteinander in Verbindung:

Theorem 4.2.9. Es seien V, W endlich-dimensionale Vektorraeume ueber K , und es sei $T : V \rightarrow W$ eine lineare Abbildung. Dann gilt

$$\dim_K V = \dim_K \ker(T) + \operatorname{rk}(T).$$

Proof. Es sei $u_1, \dots, u_n$ eine Basis fuer $\ker(T)$. Aufgrund von Theorem 3.3.19 koennen wir es zu einer Basis $u_1, \dots, u_n, v_1, \dots, v_r$ fuer V erweitern (d.h. $\dim_K V = n + r$). Behauptung: $w_1, \dots, w_r$ ist eine Basis von $\operatorname{im}(T)$.

Aufgrund von Korollar 4.1.8 ist T durch $T(u_1), \dots, T(u_n), T(v_1), \dots, T(v_r)$ eindeutig bestimmt. Da $T(u_i) = 0_W \ \forall 1 \leq i \leq n$, folgern wir, dass $T(v_1), \dots, T(v_r)$ ein Erzeugendensystem von $\operatorname{im}(T)$ ist.

Fuer $1 \leq i \leq r$ sei $w_i = T(v_i)$. Wie ueberpruefen lineare Unabhaengigkeit: nimm an, dass

$$\alpha_1 w_1 + \dots + \alpha_r w_r = 0_V \quad \Leftrightarrow \quad T(\alpha_1 v_1 + \dots + \alpha_r v_r) = 0_V.$$

Dann ist $v = \alpha_1 v_1 + \dots + \alpha_r v_r \in \ker(T)$, d.h. es gibt $\beta_1, \dots, \beta_n \in K$ so dass

$$\beta_1 u_1 + \dots + \beta_n u_n = v.$$

Wir erhalten die Gleichung

$$\alpha_1 v_1 + \dots + \alpha_r v_r - (\beta_1 u_1 + \dots + \beta_n u_n) = 0_V.$$

Da die Vektoren linear unabhaengig sind, folgt, dass

$$\alpha_1 = \dots = \alpha_r = \beta_1 = \dots = \beta_n = 0.$$

Insbesondere sind $w_1, \dots, w_r$ linear unabhaengig. $\square$

Korollar 4.2.10. Es sei $T : V \rightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen Vektorraeumen. Dann gilt:

- (1) Wenn $\dim W < \dim V$, dann ist T nicht injektiv.
- (2) Wenn $\dim W > \dim V$, dann ist T nicht surjektiv.
- (3) Wenn $\dim W = \dim V$, dann sind folgende Aussagen aequivalent: T bijektiv $\Leftrightarrow T$ ist injektiv $\Leftrightarrow T$ ist surjektiv.

²¹Wie nehmen hier *nicht* an, dass W endlich-dimensional ist.

Proof. (1) Da $\text{im}(T) \leq W$, gilt $\text{rk}(T) \leq \dim W$. Daraus folgt, dass

$$\dim \ker(T) = \dim V - \text{rk}(T) \geq \dim V - \dim W > 0$$

und T ist nicht injektiv (Satz 4.2.6).

(2) Aus Theorem 4.2.9 folgt, dass

$$\text{rk im}(T) = \dim V - \dim \ker(T) \leq \dim V < \dim W,$$

so dass T nicht surjektiv sein kann.

(3)

$$\begin{aligned} T \text{ ist injektiv} &\Leftrightarrow \ker(T) = \{0_V\} \\ &\Leftrightarrow \text{rk}(T) = \dim V \\ &\Leftrightarrow \text{rk}(T) = \dim W \\ &\Leftrightarrow \text{im}(T) = W \\ &\Leftrightarrow T \text{ ist surjektiv} \end{aligned}$$

□

Beispiele 4.2.11.

(1) Es sei $A = \begin{pmatrix} 1 & -2 & 1 & 1 \\ 2 & 0 & 0 & 3 \\ -1 & -1 & 2 & 0 \end{pmatrix}$. Dann ist die lineare Abbildung T_A nicht injektiv. Is sie surjektiv?

(2) Es sei $B = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$. Dann ist T_B bijektiv: wir haben in Beispiel 2.3.3 gesehen, dass

B zeilen-äquivalent ist zu der Matrix $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Daher ist $0_{\mathbb{R}^3}$ die einzige Lösung des homogenen linearen Gleichungssystems $B \cdot x = 0_{\mathbb{R}^3}$, d.h. $\ker(T_B) = \{0_{\mathbb{R}^3}\}$.

Folgende Beobachtung ist eine einfache Folgerung von Theorem 4.2.9:

Korollar 4.2.12. *Es sei V ein endlich-dimensionaler Vektorraum und $T : V \rightarrow W$ eine injektive lineare Abbildung. Dann gilt fuer jeden Unterraum U von V*

$$\dim_K U = \dim_K T(U).$$

Definition 4.2.13. *Eine lineare Abbildung $T : V \rightarrow W$ ist ein Isomorphismus, wenn es eine lineare Abbildung $S : W \rightarrow V$ gibt, so dass*

$$S \circ T = \text{id}_V \quad \text{und} \quad T \circ S = \text{id}_W;$$

in diesem Fall schreiben wir $S = T^{-1}$.

Wir sagen, dass V und W isomorph sind, wenn es einen Isomorphismus $T : V \rightarrow W$ gibt; in diesem Fall schreiben wir $V \cong W$.

Bemerkung 4.2.14. *Es sei X die Menge aller Vektorräume ueber K . Dann ist " $\cong$ " eine Äquivalenzrelation auf X .*

Beispiel 4.2.15. Es sei $V = \mathbb{R}[x]^{\leq 2}$. Es sei e_1, e_2, e_3 die Standardbasis von $\mathbb{R}^3$. Dann koennen wir dank Theorem 4.1.9 folgendermassen eine lineare Abbildung $V \rightarrow \mathbb{R}^3$ konstruieren: wir wissen, dass $1, x, x^2$ eine Basis von V ist. Dann gibt es genau eine lineare Abbildung $T : V \rightarrow \mathbb{R}^3$ fuer die gilt

$$T(1) = e_1, \quad T(x) = e_2, \quad T(x^2) = e_3.$$

Um zu zeigen, dass T ein Isomorphismus ist, konstruieren wir eine inverse lineare Abbildung: Definiere $S : \mathbb{R}^3 \rightarrow V$ durch

$$S : \begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto a + bx + cx^2.$$

Dann ist klar, dass $S \circ T = \text{id}_V$ und $T \circ S = \text{id}_{\mathbb{R}^3}$.

Beispiel 4.2.16. Es sei V der reelle Vektorraum aller Fibonacci-Folgen. Dann definiert die Abbildung $\mathcal{F}_{a,b} \mapsto \begin{pmatrix} a \\ b \end{pmatrix}$ einen Isomorphismus $V \cong \mathbb{R}^2$.

Bemerkung 4.2.17. Ein Isomorphismus $T : V \rightarrow V$ wird auch Automorphismus genannt.

Bemerkung 4.2.18. Wir koennen den Beweis von Theorem 4.2.9 folgendermassen interpretieren: es sei X ein Komplement von $\ker(T)$ in V . Dann induziert die Abbildung $T : V \rightarrow W$ einen Isomorphismus $T : X \cong \text{im}(T)$.

Frage. Ist jede lineare Bijektion $T : V \rightarrow W$ ein Isomorphismus?

Lemma 4.2.19. Es sei $T : V \rightarrow W$ eine bijektive lineare Abbildung. Dann ist die inverse Abbildung $T^{-1} : W \rightarrow V$ ebenfalls linear, d.h. jede bijektive lineare Abbildung ist automatisch ein Isomorphismus.

Proof. Es seien $w_1, w_2 \in W$ und $\alpha_1, \alpha_2 \in K$. Dank Bemerkung 4.1.6 muessen wir zeigen, dass

$$T^{-1}(\alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 T^{-1}(w_1) + \alpha_2^{-1} T^{-1}(w_2).$$

Fuer $i = 1, 2$ sei $v_i = T^{-1}(w_i)$. Da T linear ist, gilt

$$T(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 T(v_1) + \alpha_2 T(v_2) = \alpha_1 w_1 + \alpha_2 w_2.$$

Wenden wir T^{-1} auf diese Gleichung an, dann folgt, dass

$$T^{-1}(\alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 v_1 + \alpha_2 v_2 = \alpha_1 T^{-1}(w_1) + \alpha_2 T^{-1}(w_2).$$

□

Beispiel 4.2.20. Um zu zeigen, dass die Abbildung T in Beispiel 4.2.15 ein Isomorphismus ist, reicht es daher zu zeigen, dass sie bijektiv ist. Nun zeigt eine einfache Rechnung, dass $\ker(T) = \{0_V\}$, daher ist T bijektiv aufgrund von Korollar 4.2.10 (3).

Lecture 13

Beispiel 4.2.21. Die Abbildung T_B aus Beispiel 4.2.11 ist also ein Isomorphismus. Was ist die inverse lineare Abbildung?

Folgendes Theorem ist sehr wichtig: es klassifiziert alle n -dimensionalen K -Vektorraeume.

Theorem 4.2.22. Es seien V, W n -dimensionale K -Vektorraeume. Dann gilt $V \cong W$.

Proof. Waehle jeweilige Basen $v_1, \dots, v_n$ und $w_1, \dots, w_n$ von V und W . Aus Theorem 4.1.9 folgt, dass es eine lineare Abbildung $T : V \rightarrow W$ gibt, so dass $T(v_i) = w_i \ \forall 1 \leq i \leq n$. Dann ist T surjectiv: wenn $\alpha_1 w_1 + \dots + \alpha_n w_n \in W$, dann gilt

$$T(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_1 w_1 + \dots + \alpha_n w_n.$$

Es folgt von Korollar 4.2.10, dass T bijektiv ist und daher dank Lemma 15.1.4 ein Isomorphismus. □

Bemerkung 4.2.23. Insbesondere ist jeder n -dimensionale K -Vektorraum isomorph zu K^n .

4.3. Lineare Abbildungen als Matrizen.

Definition 4.3.1. Es sei $T : V \rightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen K -Vektorräumen, und es sei $\mathcal{B} = (v_1, \dots, v_n)$, bzw. $\mathcal{C} = (w_1, \dots, w_m)$ eine Basis von V , bzw. von W . Die Abbildungsmatrix von T bezüglich der Basen $\mathcal{B}$ und $\mathcal{C}$ ist die Matrix $[T]_{\mathcal{C}}^{\mathcal{B}} = (a_{ij}) \in M_{m \times n}(K)$, deren Einträge definiert sind durch

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

Mit anderen Worten, die Einträge in der Spalte j sind die Skalare, die wir erhalten, wenn wir Tv_j als Linearkombination von $w_1, \dots, w_m$ schreiben.

Beispiele 4.3.2. Es sei V ein endlich-dimensionaler K -Vektorraum mit Basis $\mathcal{B} = (v_1, \dots, v_n)$.

- (1) Die Abbildungsmatrix der Null-Abbildung bezüglich jeder Basis ist die Nullmatrix.
- (2) Die Abbildungsmatrix der Identitätsabbildung bezüglich der Basis $\mathcal{B}$ ist die Einheitsmatrix: $[\text{id}]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{1}_n$.

Beispiel 4.3.3. Es seien nun $V = \mathbb{R}[x]^{\leq 3}$ und $W = \mathbb{R}[x]^{\leq 2}$, mit jeweiligen Standardbasen $\mathcal{E} = (1, x, x^2, x^3)$ und $\mathcal{E}' = (1, x, x^2)$. Wir betrachten die Ableitungsabbildung $D : V \rightarrow W$. Dann ist

$$[D]_{\mathcal{E}'}^{\mathcal{E}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Beispiel 4.3.4. Es sei V der Vektorraum aller Fibonacci-Folgen und $S : V \rightarrow V$ die Verschiebungs-Abbildung. Es sei $\mathcal{B}$ die Basis $(\mathcal{F}_{1,0}, \mathcal{F}_{0,1})$. Dann ist

$$[S]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Es sei nun $\mathcal{C}$ die Basis $(\mathcal{F}_{1,\varphi}, \mathcal{F}_{1,\psi})$. In dieser Basis hat S eine besonders schöne Form: sie ist diagonal.

$$[S]_{\mathcal{C}}^{\mathcal{C}} = \begin{pmatrix} \varphi & 0 \\ 0 & \psi \end{pmatrix}.$$

Wie hängen die beiden Matrizen zusammen? Wir werden diese Frage in Abschnitt 4.5 untersuchen.

Bemerkung 4.3.5. Mit Hilfe der Abbildungsmatrix $[T]_{\mathcal{C}}^{\mathcal{B}}$ können wir sehr leicht berechnen, wohin ein Vektor $v \in V$ abgebildet wird. Schreibe v als Linearkombination der Basis $\mathcal{B}$: $v = \beta_1 v_1 + \dots + \beta_n v_n$, und es sei

$$\begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}.$$

Dann ist

$$\begin{aligned} T(v) &= T(\beta_1 v_1 + \dots + \beta_n v_n), \\ &= \sum_{i=1}^n \beta_i T(v_i) \\ &= \sum_{j=1}^m \sum_{i=1}^n a_{ij} \beta_j \cdot w_i \\ &= \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} \beta_j \right) \cdot w_i \\ &= \gamma_1 w_1 + \dots + \gamma_m w_m \end{aligned}$$

Beispiel 4.3.6. Zurück zu Beispiel 4.3.3. Was ist das Bild von einem beliebigen Element $v \in \mathbb{R}[x]^{\leq 3}$ unter der Abbildung D ? Schreibe v als Linearkombination der Basis $\mathcal{E}$: $v = a + bx + cx^2 + dx^3$. Dann gilt

$$D(v) = aD(1) + bD(x) + cD(x^2) + dD(x^3),$$

mit anderen Worten, wir berechnen

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} b \\ 2c \\ 3d \end{pmatrix},$$

und diese sind genau die Koordinaten von $D(v)$ in der Basis $\mathcal{E}'$: $D(v) = b + 2cx + 3dx^2$.

Hier ist noch ein Beispiel: Betrachte K -Vektorräume V und W mit jeweiligen Basen $\mathcal{B} = (v_1, v_2)$ und $\mathcal{C} = (w_1, w_2, w_3)$. Es sei $T : V \rightarrow W$ die lineare Abbildung gegeben durch

$$\begin{aligned} Tv_1 &= 2w_2 - w_3 \\ Tv_2 &= w_1 - w_2 - w_3. \end{aligned}$$

Dann ist die Abbildungsmatrix gegeben durch

$$[T]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 0 & 1 \\ 2 & -1 \\ -1 & -1 \end{pmatrix}.$$

Was ist $T(v)$ fuer $v = av_1 + bv_2 \in V$? Laut Bemerkung 4.3.5 berechnen wir

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

das heisst,

$$T(av_1 + bv_2) = bw_1 + (2a - b)w_2 + (-a - b)w_3.$$

Dank der Abbildungsmatrix koennen wir die Verknuepfung von linearen Abbildungen mit Hilfe von Matrizen darstellen:

Satz 4.3.7. *Es seien V , W und U drei endlich-dimensionale Vektorraeume ueber K mit jeweiligen Basen $\mathcal{A} = (v_1, \dots, v_n)$, $\mathcal{B} = (w_1, \dots, w_m)$ und $\mathcal{C} = (u_1, \dots, u_p)$. Es seien $T : V \rightarrow W$ und $S : W \rightarrow U$ lineare Abbildungen. Dann gilt*

$$[S \circ T]_{\mathcal{C}}^{\mathcal{A}} = [S]_{\mathcal{C}}^{\mathcal{B}} \cdot [T]_{\mathcal{B}}^{\mathcal{A}},$$

wobei $\cdot$ die Matrix-Multiplikation bezeichnet.

Proof. Schreibe

$$[T]_{\mathcal{B}}^{\mathcal{A}} = (a_{ij}) \in M_{m \times n}(K), \quad (b_{ij}) = [S]_{\mathcal{C}}^{\mathcal{B}} \in M_{p \times m}(K), \quad (c_{ij}) = [S \circ T]_{\mathcal{C}}^{\mathcal{A}} \in M_{p \times n}(K).$$

Dann gilt aufgrund der Definition der Darstellungsmatrizen durch

$$(17) \quad T(v_j) = a_{1j}w_1 + \dots + a_{mj}w_m,$$

$$(18) \quad S(w_i) = b_{1i}u_1 + \dots + b_{pi}u_p,$$

$$(19) \quad (S \circ T)(v_j) = c_{1j}u_1 + \dots + c_{pj}u_p$$

gegeben. Aber $(S \circ T)(v_j)$ ist ebenfalls gegeben durch

$$\begin{aligned} (S \circ T)(v_j) &= S(T(v_j)) \\ &= S(a_{1j}w_1 + \dots + a_{mj}w_m) \\ &= a_{1j}S(w_1) + \dots + a_{mj}S(w_m) \\ &= \sum_{i=1}^m a_{ij}S(w_i) \\ &= \sum_{i=1}^m a_{ij} \sum_{k=1}^p b_{ki}u_k \\ &= \sum_{i=1}^m \sum_{k=1}^p a_{ij}b_{ki}u_k, \end{aligned}$$

das heisst $c_{kj} = \sum_{i=1}^m b_{ki}a_{ij}$, was genau der Wert an der Stelle (k, j) der Matrix $[S]_{\mathcal{C}}^{\mathcal{B}} \cdot [T]_{\mathcal{B}}^{\mathcal{A}}$ ist. Quod erat demonstrandum. $\square$

4.4. Matrizen als Lineare Abbildungen.

Bemerkung 4.4.1. Es seien V und W endlich-dimensionale Vektorräume mit jeweiligen Basen $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{C} = (w_1, \dots, w_m)$. Wir haben bereits gesehen, dass jede lineare Abbildung $T : V \rightarrow W$ eine Abbildungsmatrix $[T]_{\mathcal{C}}^{\mathcal{B}}$ hat.

Umgekehrt koennen wir aus einer Matrix $A \in M_{n \times m}(K)$ eine lineare Abbildung $L_A : V \rightarrow W$ konstruieren: wir definieren L_A als die lineare Abbildung $V \rightarrow W$, fuer die gilt

$$L_A(v_j) = \sum_{i=1}^m a_{ij} w_i \quad \forall 1 \leq j \leq n.$$

Wichtig: Die lineare Abbildung L_A haengt ebenfalls von der Wahl der Basen $\mathcal{B}$ und $\mathcal{C}$ ab!

Bemerkung 4.4.2. In dem Spezialfall, dass $V = K^n$, $W = K^m$ und $\mathcal{B}$ and $\mathcal{C}$ die jeweiligen Standardbasen sind, ist L_A die lineare Abbildung T_A aus Beispiel 4.1.4.

Lemma 4.4.3. Es seien V und W endlich-dimensionale Vektorräume mit jeweiligen Basen $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{C} = (w_1, \dots, w_m)$.

(1) Es sei $T : V \rightarrow W$ eine lineare Abbildung. Dann gilt $L_{[T]_{\mathcal{C}}^{\mathcal{B}}} = T$.

(2) Es sei $A \in M_{n \times m}(K)$. Dann gilt $[L_A]_{\mathcal{C}}^{\mathcal{B}} = A$: die Abbildungsmatrix bezueglich der Basen $\mathcal{B}$ und $\mathcal{C}$ ist A selber.

Proof. Klar von den Definitionen.²² □

Bemerkung 4.4.4. Wir koennen Lemma 4.4.3 folgendermassen zusammenfassen: es gibt eine 1 : 1-Korrespondenz

$$(20) \quad \{m \times n\text{-Matrizen mit Werten in } K\} \quad \leftrightarrow \quad \{\text{lineare Abbildungen } V \rightarrow W\}$$

$$(21) \quad A \quad \mapsto \quad L_A$$

$$(22) \quad [T]_{\mathcal{C}}^{\mathcal{B}} \quad \leftarrow \quad T$$

Wichtig: diese 1 : 1 Korrespondenz haengt von der Wahl der Basen von V und W ab – sie ist nicht kanonisch.

Satz 4.4.5. Es sei V ein n -dimensionaler K -Vektorraum mit Basis $\mathcal{B} = (v_1, \dots, v_n)$. Es sei $T : V \rightarrow V$ eine lineare Abbildung. Dann ist T genau dann ein Isomorphismus, wenn $[T]_{\mathcal{B}}^{\mathcal{B}}$ eine invertierbare Matrix ist. In diesem Fall ist die Abbildungsmatrix von T^{-1} bezueglich der Basis $\mathcal{B}$ gegeben durch $([T]_{\mathcal{B}}^{\mathcal{B}})^{-1}$.

Proof. $\Rightarrow$ Nimm an, dass T ein Isomorphismus ist, d.h. es gibt eine lineare Abbildung $T^{-1} : V \rightarrow V$ so dass

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

Dann gilt aufgrund von Satz 4.3.7, dass

$$[\text{id}_V]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{1}_n = [T]_{\mathcal{B}}^{\mathcal{B}} \cdot [T^{-1}]_{\mathcal{B}}^{\mathcal{B}} = [T]_{\mathcal{B}}^{\mathcal{B}} \cdot [T^{-1}]_{\mathcal{B}}^{\mathcal{B}},$$

das heisst $[T^{-1}]_{\mathcal{B}}^{\mathcal{B}}$ ist die inverse Matrix von $[T]_{\mathcal{B}}^{\mathcal{B}}$.

$\Leftarrow$ Nimm an, dass $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ invertierbar ist, d.h. es gibt eine Matrix A^{-1} , so dass

$$A \cdot A^{-1} = A^{-1} \cdot A = \mathbf{1}_n.$$

Dann ist $L_{A^{-1}}$ die inverse Abbildung zu T : aufgrund von Satz 4.3.7 gilt

$$[L_{A^{-1}} \circ T]_{\mathcal{B}}^{\mathcal{B}} = [L_{A^{-1}}]_{\mathcal{B}}^{\mathcal{B}} \cdot [T]_{\mathcal{B}}^{\mathcal{B}} = A^{-1} \cdot A = \mathbf{1}_n = [\text{id}_V]_{\mathcal{B}}^{\mathcal{B}},$$

wobei die zweite Gleichung von Lemma 4.4.3 (2) folgt. Daher gilt $L_{A^{-1}} \circ T = \text{id}_V$. Auf aehnliche Weise koennen wir zeigen, dass $T \circ L_{A^{-1}} = \text{id}_V$. Daher ist T ein Isomorphismus. □

²²Um (1) zu zeigen, weisen Sie nach, dass fuer alle $1 \leq j \leq n$ gilt $L_{[T]_{\mathcal{C}}^{\mathcal{B}}}(v_j) = T(v_j)$. Fuer (2) berechnen sie den (i, j) Eintrag der Matrix $[L_A]_{\mathcal{C}}^{\mathcal{B}}$ und zeigen, dass es gleich dem (i, j) Eintrag von A ist.

4.5. Basiswechsel. Wir wollen nun folgende Frage untersuchen: es sei $T : V \rightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen K -Vektorräumen, und es seien $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{B}' = (v'_1, \dots, v'_n)$, bzw. $\mathcal{C} = (w_1, \dots, w_m)$ und $\mathcal{C}' = (w'_1, \dots, w'_m)$ zwei verschiedene Basen von V , bzw. von W . Was ist die Beziehung zwischen den Matrizen $[T]_{\mathcal{C}}^{\mathcal{B}}$ und $[T]_{\mathcal{C}'}^{\mathcal{B}'}$?

Definition 4.5.1. Es sei V ein endlich-dimensionaler K -Vektorraum mit Basis $\mathcal{B} = (v_1, \dots, v_n)$. Es sei $\mathcal{B}' = (v'_1, \dots, v'_n)$ eine andere Basis von V , und es sei $A = [\text{id}]_{\mathcal{B}'}^{\mathcal{B}}$, d.h. $A = (a_{ij})$ ist die Matrix, deren Einträge definiert sind durch die Gleichungen

$$v_j = \sum_{i=1}^n a_{ij} v'_i \quad \text{für } 1 \leq j \leq n.$$

Die Matrix A heisst Basiswechselmatrix von $\mathcal{B}$ nach $\mathcal{B}'$.

Bemerkung 4.5.2. Wir koennen mit Hilfe der Basiswechselmatrix $A = [\text{id}]_{\mathcal{B}'}^{\mathcal{B}}$ eine Linearkombination bezueglich der Basis $\mathcal{B}$ als eine Linearkombination bezueglich der Basis $\mathcal{B}'$ ausdruecken: es sei $v = \alpha_1 v_1 + \dots + \alpha_n v_n \in V$. Dann gilt

$$v = \beta_1 v'_1 + \dots + \beta_n v'_n,$$

wobei

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Satz 4.5.3. Die Basiswechselmatrix $[\text{id}]_{\mathcal{B}'}^{\mathcal{B}}$ ist invertierbar, mit Inversem

$$([\text{id}]_{\mathcal{B}'}^{\mathcal{B}})^{-1} = [\text{id}]_{\mathcal{B}}^{\mathcal{B}'}$$

Proof. Von Theorem 4.3.7 folgt, dass

$$\begin{aligned} [\text{id}]_{\mathcal{B}}^{\mathcal{B}} &= \mathbf{1}_n = [\text{id}]_{\mathcal{B}'}^{\mathcal{B}} \circ [\text{id}]_{\mathcal{B}}^{\mathcal{B}'} \\ [\text{id}]_{\mathcal{B}'}^{\mathcal{B}'} &= \mathbf{1}_n = [\text{id}]_{\mathcal{B}}^{\mathcal{B}'} \circ [\text{id}]_{\mathcal{B}'}^{\mathcal{B}}. \end{aligned}$$

□

Beispiel 4.5.4. Es sei $\mathcal{B} = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right)$, und es sei $\mathcal{E} = (e_1, e_2)$ die Standardbasis von $\mathbb{R}^2$. Dann ist

$$[\text{id}]_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

Wir koennen natuerlich auch die Basiswechselmatrix von $\mathcal{B}'$ nach $\mathcal{B}$ betrachten: es gilt

$$[\text{id}]_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}.$$

Wir ueberpruefen Satz 4.5.3:

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

Theorem 4.5.5. Es sei $T : V \rightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen K -Vektorräumen, und es seien $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{B}' = (v'_1, \dots, v'_n)$, bzw. $\mathcal{C} = (w_1, \dots, w_m)$ und $\mathcal{C}' = (w'_1, \dots, w'_m)$ zwei verschiedene Basen von V , bzw. von W . Dann gilt

$$[T]_{\mathcal{C}'}^{\mathcal{B}'} = [\text{id}_W]_{\mathcal{C}'}^{\mathcal{C}} \cdot [T]_{\mathcal{C}}^{\mathcal{B}} \cdot [\text{id}_V]_{\mathcal{B}}^{\mathcal{B}'}$$

Proof. Das Theorem folgt durch mehrfache Anwendung von Satz 4.3.7.

Alternativ koennen Sie ovn den Definitionen argumentieren: let $A = [T]_{\mathcal{C}}^{\mathcal{B}}$, $P = [\text{id}_V]_{\mathcal{B}}^{\mathcal{B}'}$ and $Q = [\text{id}_W]_{\mathcal{C}'}^{\mathcal{C}}$. Diese Matrizen sind durch folgende Eigenschaften definiert (vergl. Definition 4.3.1 und Beispiel 4.3.2 (3)):

$$\begin{aligned} v'_\ell &= p_{1\ell} v_1 + \dots + p_{n\ell} v_n, \\ T(v_k) &= a_{1k} w_1 + \dots + a_{mk} w_m, \\ w_j &= q_{1j} w'_1 + \dots + q_{mj} w'_m \end{aligned}$$

fuer alle $1 \leq k, \ell \leq n$ und $1 \leq j \leq m$. Dann gilt

$$\begin{aligned}
 T(v'_\ell) &= \sum_{k=1}^n p_{k\ell} T(v_k) \\
 &= \sum_{k=1}^n \sum_{j=1}^m p_{k\ell} a_{jk} w_j \\
 &= \sum_{k=1}^n \sum_{j=1}^m \sum_{i=1}^m p_{k\ell} a_{jk} q_{ij} w'_i
 \end{aligned}
 \tag{23}$$

Aber $B = [T]_{\mathcal{C}}^{\mathcal{B}'}$ ist die Matrix mit der Eigenschaft, dass

$$T(v'_\ell) = b_{1\ell} w'_1 + \cdots + b_{m\ell} w'_m.$$

Indem wir die Formeln (23) und (24) vergleichen, erhalten wir

$$b_{i\ell} = \sum_{k=1}^n \sum_{j=1}^m p_{k\ell} a_{jk} q_{ij}.$$

□

Bemerkung 4.5.6. Wir koennen Theorem 4.5.5 folgendermassen zusammenfassen: $T : V \rightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen K -Vektorraeumen, und es seien $\mathcal{B}$ und $\mathcal{B}'$, bzw. $\mathcal{C}$ und $\mathcal{C}'$ zwei verschiedene Basen von V , bzw. von W . Dann gibt es invertierbare Matrizen $P \in M_{n \times n}(K)$ und $Q \in M_{m \times m}(K)$, so dass

$$[T]_{\mathcal{C}'}^{\mathcal{B}'} = Q \cdot [T]_{\mathcal{C}}^{\mathcal{B}} \cdot P.$$

Hier ist P die Basiswechselmatrix von $\mathcal{B}'$ nach $\mathcal{B}$ und Q die Basiswechselmatrix von $\mathcal{C}$ nach $\mathcal{C}'$.

Als Spezialfall erhalten wir einen Formel fuer die Matrix eines Endomorphismus nach Basiswechsel:

Korollar 4.5.7. Es sei $T : V \rightarrow V$ ein Endomorphismus, und es seien $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{B}' = (v'_1, \dots, v'_n)$ Basen von V . Es sei $P = [\text{id}]_{\mathcal{B}'}^{\mathcal{B}}$. Dann gilt

$$[T]_{\mathcal{B}'}^{\mathcal{B}'} = [\text{id}]_{\mathcal{B}'}^{\mathcal{B}} \cdot [T]_{\mathcal{B}}^{\mathcal{B}} \cdot [\text{id}]_{\mathcal{B}}^{\mathcal{B}'}.$$

Beispiel 4.5.8. Es sei $A = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$, und wir betrachten die Abbildung $T_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Mit anderen Worten, wenn $\mathcal{E} = (e_1, e_2)$, dann gilt

$$A = [T_A]_{\mathcal{E}}^{\mathcal{E}}.$$

Es sei nun $\mathcal{B} = (v_1, v_2) = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right)$ (vergl. Beispiel 4.5.4). Was ist die Matrix von T_A bezueglich der Basis $\mathcal{B}$? Es gilt

$$[\text{id}]_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad [\text{id}]_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}.$$

Von Korollar 4.5.7 erhalten wir

$$[T_A]_{\mathcal{B}}^{\mathcal{B}} = [\text{id}]_{\mathcal{B}}^{\mathcal{E}} \cdot [T_A]_{\mathcal{E}}^{\mathcal{E}} \cdot [\text{id}]_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -3 & -10 \\ 2 & 5 \end{pmatrix}.$$

Diese Matrix bedeutet, dass

$$T_A : v_1 \mapsto -3v_1 + 2v_2 \quad \text{und} \quad T_A : v_2 \mapsto -10v_1 + 4v_2.$$

Beispiel 4.5.9. Wir koennen jetzt einige der Rechnungen mit Fibonacci Folgen besser verstehen: es seien $\mathcal{B} = (\mathcal{F}_{1,0}, \mathcal{F}_{0,1})$ und $\mathcal{C} = (\mathcal{F}_{1,\varphi}, \mathcal{F}_{1,\psi})$ Basen des Raumes V von Fibonacci-Folgen, und es sei $S : V \rightarrow V$ die Verschiebungsabbildung. Wir haben bereits gesehen, dass

$$[S]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Die Basiswechselmatrix von $\mathcal{B}$ nach $\mathcal{C}$ ist gegeben durch

$$[\text{id}]_{\mathcal{B}}^{\mathcal{C}} = \begin{pmatrix} 1 & 1 \\ \varphi & \varphi \end{pmatrix}$$

mit inverser Matrix

$$[\text{id}]_{\mathcal{C}}^{\mathcal{B}} = \frac{1}{\sqrt{5}} \begin{pmatrix} -\psi & 1 \\ \varphi & -1 \end{pmatrix}.$$

Daher gilt

$$[S]_{\mathcal{C}}^{\mathcal{C}} = [\text{id}]_{\mathcal{C}}^{\mathcal{B}} \cdot [S]_{\mathcal{B}}^{\mathcal{B}} \cdot [\text{id}]_{\mathcal{B}}^{\mathcal{C}} = \frac{1}{\sqrt{5}} \begin{pmatrix} -\psi & 1 \\ \varphi & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \varphi & \varphi \end{pmatrix} = \begin{pmatrix} \varphi & 0 \\ 0 & \psi \end{pmatrix},$$

wie erwartet. (Rechnen Sie nach!)

Definition 4.5.10.

- (1) Zwei Matrizen $A, B \in M_{n \times n}(K)$ sind *ähnlich*, wenn es eine invertierbare Matrix $P \in M_{n \times n}(K)$ gibt, so dass

$$B = P^{-1} \cdot A \cdot P.$$

- (2) Zwei Matrizen $A, B \in M_{m \times n}(K)$ sind *äquivalent*, wenn es invertierbare Matrizen $P \in M_{m \times m}(K)$ und $Q \in M_{n \times n}(K)$ gibt, so dass

$$B = P \cdot A \cdot Q.$$

Beispiele 4.5.11.

- (1) Die Matrizen $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ und $\begin{pmatrix} \varphi & 0 \\ 0 & \psi \end{pmatrix}$ aus Beispiel 4.5.9 sind ähnlich.
- (2) Die Matrizen $A = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 1 & 0 \end{pmatrix}$ und $B = \begin{pmatrix} -1 & -1 & 1 \\ 3 & 4 & 0 \end{pmatrix}$ sind äquivalent, weil

$$B = \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix} \cdot A \cdot \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Satz 4.5.12.

- (1) ‘Ähnlichkeit’ definiert eine Äquivalenzrelation auf $M_{n \times n}(K)$.
- (2) ‘Äquivalenz’ definiert eine Äquivalenzrelation auf $M_{m \times n}(K)$.

Proof. Wir beweisen (1); der Beweis von (2) ist eine Übung. Schreibe $A \sim B$ wenn A ähnlich zu B ist.

Lecture 15

- Es gilt $A \sim A$, da $A = \mathbf{1}_n^{-1} A \mathbf{1}_n$.
- Wenn $A \sim B$, dann gilt

$$B = P^{-1} \cdot A \cdot P \quad \Rightarrow \quad A = (P^{-1})^{-1} B P^{-1},$$

und daher $B \sim A$.

- Wenn $A \sim B$ und $B \sim C$, dann gibt es invertierbare Matrizen P, Q so dass

$$B = P^{-1} A P \quad \text{und} \quad C = Q^{-1} B Q,$$

woraus folgt, dass

$$C = Q^{-1} P^{-1} A P Q = (PQ)^{-1} A PQ,$$

d.h. $A \sim C$.

□

4.6. Eine Bemerkung zu Koordinaten. Viele von Ihnen haben wahrscheinlich in der Schule gelernt, dass ein Vektor gleichbedeutend ist mit einem Spaltenvektor. Das ist allerdings nur bedingt richtig. Richtig ist: die Menge aller Spaltenvektoren der Länge n mit Koeffizienten in einem Körper K ist ein Vektorraum, nämlich K^n . Allerdings gibt es noch viele andere Vektorräume.

Ein *Vektor* ist zunächst nur ein Element u eines Vektorraums V . **Nachdem** wir eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ gewählt haben, können wir u einen Spaltenvektor zuordnen: schreibe u als Linearkombination bezüglich der Basis,

$$u = \alpha_1 v_1 + \dots + \alpha_n v_n \quad \alpha_i \in K \forall 1 \leq i \leq n.$$

Dann sind die Einträge des Vektors $\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \in K^n$ die *Koordinaten* von u bezüglich der gegebenen Basis.

Wenn wir die Basis von V ändern, dann ändern sich natürlich auch die Koordinaten von u : die neuen Koordinaten sind gegeben durch die Formel in 4.5.2.

Formal lässt sich das folgendermassen formulieren:

Satz 4.6.1. *Es sei $\mathcal{B}$ eine Basis von V . Definiere*

$$\Phi_{\mathcal{B}} : V \rightarrow K^n$$

als die Abbildung, die einen Vektor auf seinen Koordinaten bezüglich der Basis $\mathcal{B}$ abbildet. Dann ist $\Phi_{\mathcal{B}}$ ein nicht-kanonischer Isomorphismus.

Proof. Da $\mathcal{B}$ eine Basis ist, lässt sich jeder Vektor eindeutig als Linearkombination der Basiselemente schreiben. Daher ist $\Phi_{\mathcal{B}}$ injektiv. Um zu zeigen, dass sie surjektiv ist, konstruieren wir eine inverse Abbildung: es sei $\mathcal{B} = (v_1, \dots, v_n)$. Definiere

$$\Psi_{\mathcal{B}} : K^n \rightarrow V, \\ \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \mapsto \alpha_1 v_1 + \dots + \alpha_n v_n.$$

Daher ist $\Phi_{\mathcal{B}}$ surjektiv und daher ein Isomorphismus aufgrund von Lemma 4.2.19. □

Wenn wir nun Vektorräume V und W haben mit jeweiligen Basen $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{C} = (w_1, \dots, w_m)$, und es ist $A \in M_{m \times n}(K)$, dann können wir die lineare Abbildung $L_A : K^n \rightarrow K^m$ aus Bemerkung 4.4.1 folgendermassen mit der Abbildung $T_A : K^n \rightarrow K^m$ in Verbindung bringen: es gilt

$$L_A = \Psi_{\mathcal{C}} \circ T_A \circ \Phi_{\mathcal{B}},$$

das heisst, L_A ist definiert als die Komposition

$$V \xrightarrow{\Phi_{\mathcal{B}}} K^n \xrightarrow{T_A} K^m \xrightarrow{\Psi_{\mathcal{C}}} W.$$

4.7. Zeilenrang gleich Spaltenrang. Wir koennen jetzt den Satz aus Abschnitt 3.6 beweisen, naemlich dass fuer jede Matrix $A \in M_{m \times n}(K)$ gilt

$$(25) \quad \text{Zeilenrang}(A) = \text{Spaltenrang}(A).$$

Der Beweis erfordert etwas Vorbereitung. Wir beginninen mit folgenden Bemerkungen:

Bemerkung 4.7.1. Von Definition 4.1.4 wissen wir, dass wir von A eine lineare Abbildung $T_A : K^n \rightarrow K^m$ erhalten. Es seien $\mathcal{E}$ und $\mathcal{F}$ die jeweiligen Standardbasen von K^n und K^m . Dann folgt von Lemma 4.4.3²³, dass

$$[T_A]_{\mathcal{F}}^{\mathcal{E}} = A.$$

Bemerkung 4.7.2. Es gilt

$$\text{im}(T_A) = \text{Spalten}(A)$$

und daher $\text{rk}(T_A) = \text{Spaltenrang}(A)$.

Theorem 4.7.3. Es sei $A \in M_{m \times n}(K)$, und es sei $r = \text{Spaltenrang}(A)$. Dann gibt es invertierbare Matrizen $P \in M_{n \times n}(K)$ und $Q \in M_{m \times m}(K)$, so dass QAP die Form $\begin{pmatrix} \mathbf{1}_r & 0_{r \times s} \\ 0_{t \times r} & 0_{t \times s} \end{pmatrix}$ hat, wobei $s = n - r$ und $t = m - r$.

Proof. Es sei $(u_1, \dots, u_s)$ eine Basis fuer $\ker(T_A)$. Wie in dem Beweis von Theorem 4.2.9 erweitern wir die Basis zu einer Basis

$$\mathcal{B} = (v_1, \dots, v_r, u_1, \dots, u_s)$$

von K^n . Dein Einfachheit halber sei $v_{r+i} = u_i$ fuer $1 \leq i \leq s$, d.h.

$$\mathcal{B} = (v_1, \dots, v_n).$$

Fuer $1 \leq i \leq r$ sei $w_i = L_A(v_i)$; dann ist $(w_1, \dots, w_r)$ eine Basis von $\text{im}(T_A)$. Wir erweitern diese Basis zu einer Basis

$$\mathcal{C} = (w_1, \dots, w_r, w_{r+1}, \dots, w_m)$$

von K^m .

Was ist die Matrix von T_A bezueglich der Basen $\mathcal{B}$ und $\mathcal{C}$? Schreiben wir $[T_A]_{\mathcal{C}}^{\mathcal{B}} = (c_{ij})$, dann gilt per Definition

$$T_A(v_j) = c_{1j}w_1 + \dots + c_{mj}w_m.$$

Nun wissen wir aber, dass

$$T_A(v_j) = \begin{cases} w_j & \text{fuer } 1 \leq j \leq r \\ 0_{K^m} & \text{fuer } r < j \leq n \end{cases}$$

Mit anderen Worten,

$$[T_A]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} \mathbf{1}_r & 0_{r \times s} \\ 0_{t \times r} & 0_{t \times s} \end{pmatrix}.$$

Nun wissen wir von Theorem 4.5.5, dass

$$[T_A]_{\mathcal{C}}^{\mathcal{B}} = [\text{id}_{K^m}]_{\mathcal{C}}^{\mathcal{F}} \cdot [T_A]_{\mathcal{F}}^{\mathcal{E}} \cdot [\text{id}_{K^n}]_{\mathcal{E}}^{\mathcal{B}}.$$

Da $[T_A]_{\mathcal{F}}^{\mathcal{E}} = A$, erhalten wir das Resultat fuer $Q = [\text{id}_{K^m}]_{\mathcal{C}}^{\mathcal{F}}$ und $P = [\text{id}_{K^n}]_{\mathcal{E}}^{\mathcal{B}}$. □

Dieses Theorem hat eine interessante Konsequenz: erinnern Sie sich (Def. 4.5.10 (2)), dass zwei Matrizen $A, B \in M_{m \times n}(K)$ aequivalent sind, wenn es invertierbare Matrizen $P \in M_{m \times m}(K)$ und $Q \in M_{n \times n}(K)$ gibt, so dass

$$B = P \cdot A \cdot Q;$$

wir schreiben $A \sim B$.

Korollar 4.7.4. Es seien $a, b \in M_{m \times n}(K)$. Dann sind A und B genau dann aequivalent, wenn $\text{Spaltenrang}(A) = \text{Spaltenrang}(B)$. Insbesondere zerfaellt $M_{m \times n}(K)$ in $\min\{m, n\} + 1$ Äquivalenzklassen.

²³weil T_A nichts anderes ist als die Abbildung $L_A : K^n \rightarrow K^m$ bezueglich der Basen $\mathcal{E}$ und $\mathcal{F}$.

Proof. $\Leftarrow$: Es sei r der gemeinsame Spaltenrang. By Theorem 4.7.3 gilt

$$a \sim \begin{pmatrix} \mathbf{1}_r & 0_{r \times s} \\ 0_{t \times r} & 0_{t \times s} \end{pmatrix} \quad \text{und} \quad B \sim \begin{pmatrix} \mathbf{1}_r & 0_{r \times s} \\ 0_{t \times r} & 0_{t \times s} \end{pmatrix}.$$

Da ' $\sim$ ' eine Aequivalenzrelation ist (Proposition 4.5.12), folgt daraus, dass $A \sim B$.

$\Rightarrow$: Nimm an, dass $A \sim B$, d.h. es gibt invertierbare Matrizen $P \in M_{m \times m}(K)$, $Q \in M_{n \times n}(K)$ so dass $B = PAQ$. Dann gilt

$$T_B = T_P \circ T_A \circ T_Q,$$

wobei $T_\star$ die von $\star$ induzierte lineare Abbildung bezueglich der Einheitsbasen ist. Wir wollen zeigen, dass $\dim T_A = \dim T_B$. Es gilt

$$\begin{aligned} \operatorname{im}(T_B) &= T_B(K^n) \\ &= T_P \circ T_A(T_Q(K^n)) \\ &= T_P \circ T_A(K^n) && \text{da } T_Q \text{ ein Isomorphismus ist} \\ &= T_P(\operatorname{im}(T_A)) \end{aligned}$$

Daraus folgt, dass

$$\dim \operatorname{im}(T_B) = \dim T_P(\operatorname{im}(T_A)),$$

Doch da T_P ein Isomorphismus ist, folgt von Korollar 4.2.12, dass

$$\dim T_P(\operatorname{im}(T_A)) = \dim \operatorname{im}(T_A).$$

□

Übung 4.7.5. *Koennen Sie die Aequivalenzklassen von aehnlichen Matrizen beschreiben?*

Um (25) zu zeigen, muessen wir nun beweisen, dass sich der Zeilen- bzw. Spaltenrang einer Matrix nicht aendern, wenn wir die Matrix von links, bzw. von rechts mit einer invertierbaren Matrix multiplizieren.

Lecture 16

Satz 4.7.6. *Es sei $A \in M_{m \times n}(K)$, und es sei $Q \in M_{m \times m}(K)$ invertierbar. Dann gilt*

- (1) $\operatorname{Zeilenrang}(QA) = \operatorname{Zeilenrang}(A)$;
- (2) $\operatorname{Spaltenrang}(QA) = \operatorname{Spaltenrang}(A)$.

Proof. (1) Wir schreiben $Q = (q_{ij})$. Es seien $x_1, \dots, x_m$ die Zeilenvektoren von A . Dann ist die i te Zeile von QA gegeben durch

$$q_{i1}x_1 + \dots + q_{im}x_m,$$

mit anderen Worten, die Zeilen von QA sind Linearkombinationen von den Zeilen von A und daher

$$\operatorname{Zeilenrang}(QA) \leq \operatorname{Zeilenrang}(A).$$

Da Q invertierbar ist, koennen wir dieses Argument auch Multiplikation bei $Q^{-1}A$ anwenden und erhalten

$$\operatorname{Zeilenrang}(Q^{-1}QA) \leq \operatorname{Zeilenrang}(QA),$$

das heisst $\operatorname{Zeilenrang}(A) \leq \operatorname{Zeilenrang}(QA)$. Daraus folgt, dass

$$\operatorname{Zeilenrang}(QA) = \operatorname{Zeilenrang}(A).$$

(2) Es seien nun $y_1, \dots, y_n$ die Spalten von A , und es sei $U = \langle y_1, \dots, y_n \rangle \leq K^m$. Dann sind die Spalten von QA gegeben durch $Qy_1, \dots, Qy_n$, das heisst durch $L_Q(y_1), \dots, L_Q(y_n)$, wobei $L_Q : K^m \rightarrow K^m$ die durch Q gegebene lineare Abbildung bezueglich der Standardbasis ist. Nun ist aber Q invertierbar, was bedeutet, dass L_Q ein Isomorphismus ist und insbesondere injektiv. Daher folgt von Korollar 4.2.12, dass

$$\dim_K U = \dim_K L_Q(U),$$

und daher

$$\operatorname{Spaltenrang}(QA) = \operatorname{Spaltenrang}(A).$$

□

Um das analoge Resultat fuer AP zu zeigen, benutzen wir einen Trick:

Definition 4.7.7. *Es sei $B = (b_{ij}) \in M_{m \times n}(K)$. Definiere die transponierte Matrix B^t als die $(n \times m)$ -Matrix, deren (i, j) Eintrag durch b_{ji} gegeben ist.*

Bemerkung 4.7.8. *Mit anderen Worten, die Zeilen von B sind die Spalten von B^t und umgekehrt.*

Beispiel 4.7.9.

- (1) Es sei $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$. Dann ist $A^t = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$.
- (2) Es sei $B = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 2 \\ 0 & 0 & -1 \end{pmatrix}$. Dann ist $B^t = \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 3 & 2 & -1 \end{pmatrix}$.

Lemma 4.7.10.

- (1) Für eine Matrix $B \in M_{m \times n}(K)$ gilt $(B^t)^t = B$.
- (2) Es sei $B \in M_{m \times n}(K)$ und $C \in M_{n \times p}(K)$. Dann gilt
- $$(BC)^t = C^t B^t.$$
- (3) Eine Matrix $B \in M_{n \times n}(K)$ ist genau dann invertierbar, wenn B^t invertierbar ist, und in diesem Fall ist $(B^t)^{-1} = (B^{-1})^t$.
- (4) Es gilt

$$\text{Zeilenrang}(B) = \text{Spaltenrang}(B^t) \quad \text{und} \quad \text{Zeilenrang}(B^t) = \text{Spaltenrang}(B)$$

Proof. Übung. □

Wir können nun folgenden Satz als eine einfache Konsequenz von Satz 4.7.6 beweisen.

Satz 4.7.11. Es sei $A \in M_{m \times n}(K)$, und es sei $P \in M_{n \times n}(K)$ invertierbar. Dann gilt

- (1) $\text{Zeilenrang}(AP) = \text{Zeilenrang}(A)$;
 (2) $\text{Spaltenrang}(AP) = \text{Spaltenrang}(A)$.

Proof. Dank Lemma 4.7.10 (4) ist es ausreichend, den Satz in dem Fall zu zeigen, wenn wir AP durch $(AP)^t$, bzw. A durch A^t ersetzen. Aber

$$(AP)^t = P^t A^t,$$

und P^t ist invertierbar dank Lemma 4.7.10 (2) und (3). Der Satz folgt nun von Satz 4.7.6. □

Korollar 4.7.12. Es sei $A \in M_{m \times n}(K)$, und es seien $Q \in M_{m \times m}(K)$ und $P \in M_{n \times n}(K)$ invertierbar. Dann gilt

$$\text{Zeilenrang}(QAP) = \text{Zeilenrang}(A) \quad \text{und} \quad \text{Spaltenrang}(QAP) = \text{Spaltenrang}(A).$$

Folgender Satz ist nun eine einfache Konsequenz:

Theorem 4.7.13. Es sei $A \in M_{m \times n}$. Dann gilt

$$\text{Zeilenrang}(A) = \text{Spaltenrang}(A).$$

Proof. Wir haben in Theorem 4.7.3 gesehen, dass es invertierbare Matrizen $P \in M_{n \times n}(K)$ und $Q \in M_{m \times m}(K)$ gibt, so dass QAP die Form $\begin{pmatrix} \mathbf{1}_r & 0_{r \times s} \\ 0_{t \times r} & 0_{t \times s} \end{pmatrix}$ hat, wobei $s = n - r$ und $t = m - r$. Der Zeilen- und Spaltenrang dieser Matrix ist offensichtlich r . Dank Korollar 4.7.12 wissen wir, dass die Matrix QAP den gleichen Zeilen- bzw. Spaltenrang hat wie A . □

Definition 4.7.14. Es sei $A \in M_{m \times n}(K)$. Der gemeinsame Wert von Zeilen- und Spaltenrang von A ist der Rang $\text{rk}(A)$ von A .

Bemerkung 4.7.15. Obwohl für $A \in M_{m \times n}(K)$ gilt

$$\text{Zeilenrang}(A) = \text{Spaltenrang}(A),$$

sind die Räume $\text{Spalten}(A)$ und $\text{Zeilen}(A)$ verschieden! Zunächst ist $\text{Spalten}(A) \leq K^m$ und $\text{Zeilen}(A) \leq K^n$. Aber auch wenn $m = n$ erhalten wir im allgemeinen verschiedene Unterräume von K^n .

4.8. Zurueck zu linearen Gleichungssystemen. Es sei $A \in M_{m \times n}$, und wir betrachten das homogene lineare Gleichungssystem

$$(S_A) : A \cdot x = 0_{K^m}.$$

Wir schreiben $L(S_A) \leq K^n$ fuer die Loesungen des Gleichungssystems.

Lemma 4.8.1. *Es gilt*

$$\dim L(S_A) = n - \text{Rang}(A).$$

Proof. Es sei T_A die durch A gegebene lineare Abbildung $K^n \rightarrow K^m$ bezueglich der Standardbasen. Beachte (Bemerkung 4.7.2), dass $L(S_A) = \ker(T_A)$ und $\text{Rang}(A) = \dim \text{im}(T_A)$. Aber aufgrund von Theorem 4.2.9 gilt

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

□

Wir wissen nun von Theorem 2.3.9, dass A zeilen-aequivalent ist zu einer Matrix $A' = (a'_{ij})$ in reduzierter Zeilenstufenform, und dank Theorem 2.4.5 gilt

$$L(S_A) = L(S_{A'}).$$

Wir wir bereits im Beweis von Lemma 3.6.6 gesehen haben, ist $r = \text{Rang}(A')$ die Anzahl der fuehrenden Einsen von A' ; nimm an, dass sie in den Spalten

$$1 \leq j_1 < j_2 < \dots < j_r \leq n.$$

auftreten. Die Loesungsmenge $L(S_{A'})$ hat dann $\ell = n - r$ freie Variablen, sagen wir $x_{i_1}, \dots, x_{i_\ell}$ fuer

$$1 \leq i_1 < \dots < i_\ell \leq n.$$

Bemerkung 4.8.2. *Fuer alle Werte der freien Variablen $x_{i_1}, \dots, x_{i_\ell} \in K$ sind die Werte der Variablen $x_{j_1}, \dots, x_{j_r}$ eindeutig bestimmt, und zwar durch*

$$(26) \quad x_{j_k} = - \sum_{1 \leq q \leq \ell: i_q > j_k} a'_{k,q} x_{i_q}.$$

Wir erhalten daher eine Abbildung

$$\Phi : K^\ell \rightarrow L(S_A),$$

die ein ℓ -Tupel $(\lambda_1, \dots, \lambda_\ell)$ auf diejenige eindeutige Loesung

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in L(S_A) = L(S_{A'})$$

abbildet, fuer die gilt

$$x_{i_1} = \lambda_1, \dots, x_{i_\ell} = \lambda_\ell.$$

Beispiel 4.8.3. Es sei $A' \in M_{3 \times 6}(\mathbb{R})$ gegeben durch

$$A' = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Dann ist A' in reduzierter Zeilenstufenform, und es gilt $r =$ und $\ell = 6 - 3 = 3$. Die fuehrenden Einsen sind in den Spalten $j_1 = 2, j_2 = 4, j_3 = 5$, das heisst, fuer $x = \begin{pmatrix} x_1 \\ \vdots \\ x_6 \end{pmatrix} \in L(S_{A'})$ haben wir freie Variablen x_1, x_3, x_6 , und die Werte von x_2, x_4, x_5 sind bestimmt durch

$$x_2 = x_3 - 2x_6$$

$$x_4 = x_6$$

$$x_5 = -x_6$$

Mit anderen Worten, die Abbildung $\Phi : \mathbb{R}^3 \rightarrow L(S_{A'})$ ist gegeben durch

$$(a, b, c) \mapsto x = \begin{pmatrix} a \\ b - 2c \\ b \\ c \\ -c \\ c \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + c \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}.$$

Satz 4.8.4. *The Abbildung Φ ist linear und ein Isomorphismus.*

Proof. Gleichung (26) zeigt, dass $x_{j_1}, \dots, x_{j_r}$ linear von den freien Variablen abh ngen; daher ist Φ eine lineare Abbildung.

Um zu zeigen, dass Φ ein Isomorphismus ist, reicht es zu zeigen, dass $\ker(\Phi) = \{0_{K^\ell}\}$. Aber das folgt unmittelbar von der Definition von Φ : wenn $\Phi(\lambda_1, \dots, \lambda_\ell) = 0_{K^m}$, dann folgt $\lambda_i = 0$ for all i , da jedes λ_i ein Eintrag in $\Phi(\lambda_1, \dots, \lambda_\ell)$ ist. $\square$

Beispiel 4.8.5. Bezogen auf Beispiel 4.8.3 heisst das, dass jedes Element $x \in L(S_{A'})$ von der Form

$$a \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + c \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

f r irgendwelche $a, b, c \in \mathbb{R}$ ist.

Bemerkung 4.8.6. *Die inverse Abbildung Φ^{-1} ist folgendermassen gegeben: es sei $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in L(S_A)$.*

Dann gilt

$$\Phi^{-1}(x) = \begin{pmatrix} x_{i_1} \\ \vdots \\ x_{i_\ell} \end{pmatrix}.$$

Korollar 4.8.7. *Es seien $A \in M_{m \times n}(K)$, $b \in K^m$, und schreibe $(S_{A,b})$ f r das lineare Gleichungssystem $Ax = b$.*

- (1) *Es gilt $L(S_{A,b}) \neq \emptyset$ genau dann, wenn $b \in \text{Spalten}$.*
- (2) *Wenn $L(S_{A,b}) \neq \emptyset$ und $y \in L(S_{A,b})$, dann gilt*

$$L(S_{A,b}) = y + L(S_A) = \{x + y : x \in L(S_A)\}.$$

Proof. (1) Es seien $y_1, \dots, y_n$ die Spalten von A . Dann gilt f r einen beliebigen Vektor $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$:

$$Ax = x_1 y_1 + \dots + x_n y_n,$$

das heisst, $\text{im}(L_A) = \langle y_1, \dots, y_n \rangle$.

- (2) Es sei $y \in L(S_{A,b})$ und $z \in L(S_A)$. Dann gilt

$$A(y + z) = Ay + Az = b + 0_{K^m} = b,$$

so dass $y + z \in L(S_{A,b})$.

Wenn umgekehrt $y' \in L(S_{A,b})$, dann gilt

$$A(y - y') = Ay - Ay' = b - b = 0_{K^m},$$

mit anderen Worten $y' \in y + L(S_A)$. $\square$

Bemerkung 4.8.8. *Wir sehen also: $L(S_{A,b})$ ist gegeben durch die Verschiebung des Unterraums $L(S_A)$ entlang y . Eine Untermenge von K^n der Form $y + L(S_A)$ ist die durch y erzeugte Nebenklasse (eng. coset) von $L(S_A)$. Sie werden solchen Nebenklassen in der Algebra noch oft begegnen, zum Beispiel als die Elemente von Quotientenr umen.*

Satz 4.8.9. *Es sei $A \in M_{m \times n}(K)$, $b \in K^m$, und es sei $A|b$ die erweiterte Matrix. Folgende Bedingungen sind äquivalent:*

- (1) *Es gilt $\text{rk}(A|b) = \text{rk}(A)$.*
- (2) *Das lineare Gleichungssystem*

$$S_{A,b} : A.x = b$$

hat eine Lösung.

Proof. Es seien $y_1, \dots, y_n$ die Spaltenvektoren von A . Dann gilt

$$\begin{aligned} & \text{rk}(A|b) = \text{rk}(A) \\ \Leftrightarrow & \text{Spaltenrang von } A|b = \text{Spaltenrang von } A \\ \Leftrightarrow & \langle y_1, \dots, y_n, b \rangle = \langle y_1, \dots, y_n \rangle \\ \Leftrightarrow & b \in \langle y_1, \dots, y_n \rangle \\ \Leftrightarrow & b \in \text{Spalten}(A) \\ \Leftrightarrow & L(S_{A,b}) \neq \emptyset, \end{aligned}$$

wobei die letzte Äquivalenz von Korollary 4.8.7 (1) folgt. □

Korollar 4.8.10. *Es sei $A \in M_{m \times n}(K)$, $b \in K^m$, und es sei $A|b$ die erweiterte Matrix. Folgende Bedingungen sind äquivalent:*

- (1) *Es gilt $\text{rk}(A|b) = \text{rk}(A) = n$.*
- (2) *Das lineare Gleichungssystem*

$$S_{A,b} : A.x = b$$

hat genau eine Lösung.

Proof. Übung. □

5.1. Gruppen.

Definition 5.1.1. Eine Gruppe ist eine Menge G zusammen mit einer Operation $\star : G \times G \rightarrow G$, die folgende Axiome erfuehlt:

- (1) Assoziativitaet: $\forall g, h, k \in G$ gilt $(g \star h) \star k = g \star (h \star k)$;
- (2) Existenz eines neutralen Elements: es gibt ein Element $e \in G$, so dass $\forall g \in G$ gilt $g \star e = e \star g = g$;²⁴
- (3) Existenz eines Inversen: $\forall g \in G$ gibt es ein Element g^{-1} so dass $g \star g^{-1} = g^{-1} \star g = e$.²⁵

Eine Gruppe ist kommutativ oder abelsch, wenn $g \star h = h \star g$ fuer alle $g, h \in G$.

Beispiel 5.1.2.

- Die triviale Gruppe ist die Gruppe mit einem Element.
- Es sei V ein Vektorraum. Dann ist V eine Gruppe unter der Vektoraddition. Diese Gruppe ist abelsch (= kommutativ): es gilt $v + u = u + v$ fuer all $u, v \in V$.
- Insbesondere ist jeder Koerper unter der Addition eine kommutative (= abelsche) Gruppe.
- Auch $(\mathbb{Z}, +)$ und $(\mathbb{F}_p, +)$ sind abelsche Gruppen. Insbesondere ist $(\mathbb{F}_p, +)$ das erste Beispiel einer endlichen Gruppe.
- Es sei K ein Koerper, und es sei $K^\times = K \setminus \{0\}$. Dann ist $K^\times$ eine abelsche Gruppe bezueglich der Multiplikation.
- Die Menge $\{\pm 1\}$ ist eine Gruppe unter Multiplikation; es ist die einzige Gruppe der Ordnung 2 und die kleinste nicht-triviale Gruppe.

Diese Beispiele sind aber alle abelsch und etwas langweilig. Interessante Beispiele kommen aus der Geometrie und aus der Theorie von Matrizen:

Definition 5.1.3. Es sei X ein geometrischer Koerper. Eine Symmetrie von X ist eine Abbildung von X auf sich selber, die die Distanzen von Punkten in X zueinander nicht aendert.

Beispiel 5.1.4. Beispiele fuer Symmetrien sind Drehungen und Spiegelungen.

Beispiele 5.1.5.

- (1) Die Symmetrien des Quadrats sind eine nicht-abelsche endliche Gruppe der Ordnung 8, die sogenannte *dihedrale Gruppe* D_8 .
- (2) Es gibt 48 Symmetrien eines Wuerfels; sie bilden eine nicht-abelsche Gruppe, die *symmetrische Gruppe* S_4 .
- (3) Die Symmetriegruppe des Ikosaeder hat 60 Elemente; es ist die *alternierende Gruppe* A_5 , eine Untergruppe der symmetrischen Gruppe S_5 .²⁶
- (4) Es sei $n \geq 1$, und es sei $GL_n(K)$ die Menge aller invertierbaren Matrizen in $M_{n \times n}(K)$. Dann ist $GL_n(K)$ eine Gruppe unter Matrix-Multiplikation; sie heisst die *allgemeine lineare Gruppe vom Grad n* . Wenn $n > 1$, dann ist $GL_n(K)$ nicht abelsch. (Was ist $GL_1(K)$?) Die Gruppe $GL_n(K)$ had viele interessante Untergruppen: zum Beispiel die Untergruppe aller oberen Dreiecksmatrizen, oder die (abelsche) Untergruppe diagonalen Matrizen. Weitere Beispiele fuer Matrixgruppen sind symplektische Gruppen, unitaere Gruppen und orthogonale Gruppen. Letztere werden Sie in LA2 kennenlernen.

Übung 5.1.6. Wieviele Elemente hat die Gruppe $GL_n(\mathbb{F}_p)$?

Bemerkung 5.1.7. Gruppen finden sich in der Algebra ueberall, und Gruppentheorie ist ein riesiges Forschungsgebiet. Ein paar Beispiele:

- (1) Vor 200 Jahren hat der franzoesische Mathematiker Evariste Galois im Alter von 20 Jahren mit Gruppentheorie die Frage beantwortet, ob sich ein Polynom durch Radikale (verallgemeinerte Wurzeln) loesen laesst. Seine Idee war, jedem Polynom eine endliche Gruppe zuzuordnen und die Frage der Loesbarkeit mit Radikalen mit Hilfe der Eigenschaften dieser Gruppe zu beantworten. Galois' Arbeit hat ein neues Forschungsgebiet geschaffen: die algebraische Zahlentheorie.

²⁴Man kann zeigen, dass das neutrale Element e eindeutig bestimmt ist.

²⁵Man kann zeigen, dass g^{-1} eindeutig bestimmt ist; es heisst das Inverse von g .

²⁶Ueber alternierende Gruppen werden wir spaeter im Zusammenhang mit Determinanten noch mehr erfahren.

- (2) Von 1955 bis 2004 hat eine Gruppe von Forschern alle sog. einfachen endlichen Gruppen klassifiziert (eine Art Periodensystem der Gruppentheorie); der Beweis erstreckt sich ueber mehrere 10000 Seiten. Diese Gruppen lassen sich in Familien einteilen, mit Ausnahme der 26 sog. sporadischen Gruppen, die in keine der Familien gehoeren. Die groessten dieser sporadischen Gruppens sind die Monstergruppe (ca. $8 \cdot 10^{53}$ Elemente²⁷) und das Babymonster (ca. $4 \cdot 10^{33}$ Elemente).
- (3) Matrixgruppen lassen sich mit Hilfe der linearen Algebra studieren: dieses sehr attraktive Gebiet heisst Darstellungstheorie.
- (4) 1998 hat Richards Borchers die Fieldsmedallie dafuer bekommen, dass er die sogenannte Moonshine Vermutung von Conway bewiesen hat, die die Darstellungen der Monstergruppe mit den Werten einer in der Zahlentheorie wichtigen Funktion in Verbindung gebracht hat.
- (5) Eines der wichtigsten heutigen Forschungsgebiete ist die sogenannte Langlands-Korrespondenz, das Eigenschaften der allgemeinen Linearen Gruppen mit Objekten in der Galoistheorie miteinander in Beziehung bringt. Einer der einfachsten Faelle dieser Korrespondenz ist die sogenannte Taniyama–Shimura Vermutung, die Wiles’ Beweis des Fermatschen Satzes zugrunde liegt.

²⁷Zum Vergleich: die Erde besteht aus ca $6 \cdot 10^{49}$ Atomen.

5.2. Ringe.

Definition 5.2.1. *Ein Ring ist eine Menge R mit zwei Operationen $+$ (Addition) und $\times$ (Multiplikation), die folgende Axiome erfuellen:*

(1) $(R, +)$ ist eine abelsche Gruppe;

(2) Multiplikation ist assoziativ: $\forall a, b, c \in R$ gilt

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

(3) Es gibt ein Element $1_R \in R$ fuer dass gilt²⁸

$$1_R \times a = a \times 1_R = a.$$

(4) Multiplikation ist distributiv bezueglich der Addition, d.h.

$$a \times (b + c) = a \times b + a \times c \quad \text{und} \quad (b + c) \times a = b \times a + c \times a.$$

Bemerkung 5.2.2. *Hier multiplizieren wir nicht Elemente von R mit Skalaren, sonder wir multiplizieren zwei Elemente von R . Ein Vektorraum ist also im Allgemeinen kein Ring!*

Eine Ringe haben Sie schon kennengelernt:

Beispiele 5.2.3.

(1) $\mathbb{Z}$ ist ein (kommutativer) Ring.

(2) Ebenso ist jeder Koerper ein kommutativer Ring.

(3) Es sei $n \geq 1$. Dann ist $M_{n \times n}(K)$ ein Ring unter Matrix-Addition und Matrix-Multiplikation.

Wenn $n > 1$, dann ist der Ring nicht kommutativ (d.h. die Multiplikation ist nicht kommutativ).

Bemerkung 5.2.4. *Auch Ringe finden sich ueberall in der Algebra. Der ‘Herr der Ringe’ ist der franzoesische Mathematiker Jean-Marc Fontaine²⁹, da er aeussert wichtige sogenannte ‘Periodenringe’ (mit den schoenen Namen $\mathbf{B}_{\text{dR}}$, $\mathbf{B}_{\text{HT}}$, $\mathbf{B}_{\text{cris}}$, ...) in der p -adischen Hodge Theorie eingefuehrt hatte.*

²⁸Das Element ist die multiplikative Identitaet.

²⁹ganz egal, was Tolkien dazu sagt

6.1. Definition und erste Eigenschaften.

Definition 6.1.1. Es seien V, W K -Vektorraeume. Wir schreiben $\text{Hom}_K(V, W)$ fuer die Menge aller linearen Abbildungen von V nach W .

Satz 6.1.2. Es seien V, W K -Vektorraeume. Dann hat $\text{Hom}_K(V, W)$ ebenfalls die Struktur eines K -Vektorraums, mit den folgenden Operationen:

(1) Es seien $T_1, T_2 \in \text{Hom}_K(V, W)$. Dann ist $T_1 + T_2$ definiert durch

$$(T_1 + T_2)(v) = T_1(v) + T_2(v) \quad \forall v \in V.$$

(2) Es seien $T \in \text{Hom}_K(V, W)$ und $\alpha \in K$. Dann ist

$$(\alpha T)(v) = \alpha T(v) \quad \forall v \in V.$$

Proof. Wir zeigen zunaechst, dass $T_1 + T_2$ und αT ebenfalls Elemente von $\text{Hom}_K(V, W)$ sind. Dazu benutzen wir Bemerkung 4.1.6. Es seien $u, v \in V$ und $\lambda \in K$.

(1) Es gilt

$$\begin{aligned} (T_1 + T_2)(v + \lambda u) &= T_1(v + \lambda u) + T_2(v + \lambda u) \\ &= T_1(v) + \lambda T_1(u) + T_2(v) + \lambda T_2(u) \\ &= T_1(v) + T_2(v) + \lambda (T_1(u) + T_2(u)) \\ &= (T_1 + T_2)(v) + \lambda (T_1 + T_2)(u), \end{aligned}$$

das heisst, $T_1 + T_2$ ist linear.

(2) Es gilt

$$\begin{aligned} (\alpha T)(v + \lambda u) &= \alpha T(v + \lambda u) \\ &= \alpha T(v) + \alpha \lambda T(u) \\ &= (\alpha T)(v) + \lambda (\alpha T)(u), \end{aligned}$$

das heisst αT ist linear.

Wir muessen nun zeigen, dass $\text{Hom}_K(V, W)$ mit diesen Operationen alle Vektorraumaxiome erfuehlt. Wir zeigen die Existenz des neutralen Elements; der Rest ist eine Uebung.

Es sei $0 : V \rightarrow W$ die Null-Abbildung, d.h. $0(v) = 0_W$ fuer alle $v \in V$. Klarerweise ist 0 linear. Wir zeigen nun, dass $T + 0 = 0 + T = T$ fuer all $T \in \text{Hom}_K(V, W)$. Tatsaechlich gilt

$$(T + 0)(v) = T(v) + 0(v) = T(v) + 0_W = T(v),$$

das heisst $T + 0 = T$. Aehnlich koennen wir zeigen, dass $0 + T = T$. □

Theorem 6.1.3. Es seien V, W endlich-dimensionale K -Vektorraeume, und es sei $\mathcal{B} = (v_1, \dots, v_n)$, bzw. $\mathcal{C} = (w_1, \dots, w_m)$ eine Basis von V , bzw. W . Dann ist die Abbildung

$$\Psi_{\mathcal{C}}^{\mathcal{B}} : \text{Hom}_K(V, W) \rightarrow M_{m \times n}(K), \quad T \mapsto [T]_{\mathcal{C}}^{\mathcal{B}}$$

linear und ein Isomorphismus.

Proof. Selbstverstandlich gilt

$$\Psi_{\mathcal{C}}^{\mathcal{B}}(T_1 + \alpha T_2) = \Psi_{\mathcal{C}}^{\mathcal{B}}(T_1) + \alpha \Psi_{\mathcal{C}}^{\mathcal{B}}(T_2),$$

das heisst die Abbildung ist linear.

Wir muessen daher dank Lemma 15.1.4 nur zeigen, dass $\Psi_{\mathcal{C}}^{\mathcal{B}}$ bijektiv ist. Aber das ist der Inhalt von Bemerkung 4.4.4. □

Korollar 6.1.4. Wenn V, W endlich-dimensionale Vektorraeume sind, dann gilt das gleiche fuer $\text{Hom}(V, W)$, und es gilt

$$\dim_K \text{Hom}(V, W) = \dim_K V \cdot \dim_K W.$$

Proof. Es seien $n = \dim V$ und $m = \dim W$. Dank Theorem 6.1.3 wissen wir, dass $\text{Hom}_K(V, W) \cong M_{m \times n}(K)$. Aber $M_{m \times n}(K)$ ist endlich-dimensional, mit $\dim M_{m \times n}(K) = mn$ (was ist eine Basis?). □

In dem Fall, wenn $W = V$, laesst sich Theorem 6.1.3 noch verfeinern:

Satz 6.1.5. *Es sei V ein endlich-dimensionaler Vektorraum. Dann ist $\text{Hom}(V, V)$ ein Ring unter der Addition und Komposition von Funktionen. Weiterhin ist die Abbildung $\Psi_{\mathcal{B}}^{\mathcal{B}} : \text{Hom}_K(V, V) \rightarrow M_{n \times n}(K)$, fuer eine Basis $\mathcal{B}$ von V , ein Ring-Isomorphismus.*

Proof. (Skizze) Die Ring-Struktur von $\text{Hom}(V, V)$ kann durch explizite Rechnungen ueberprueft werden. Weiterhin folgt von Satz 4.3.7, dass

$$\Psi_{\mathcal{B}}^{\mathcal{B}}(T \circ S) = [T]_{\mathcal{B}}^{\mathcal{B}} \cdot [S]_{\mathcal{B}}^{\mathcal{B}},$$

das heisst, $\Psi_{\mathcal{B}}^{\mathcal{B}}$ ist kompatibel mit den jeweiligen Ring Multiplikationen. □

6.2. Der duale Vektorraum. Ein sehr wichtiges Beispiel von $\text{Hom}(V, W)$ ist der Fall, wenn $W = K$.

Definition 6.2.1. Es sei V ein K -Vektorraum. Der Dualraum von V ist der Vektorraum

$$V^* = \text{Hom}_K(V, K).$$

Die Elemente von V^* sind lineare Abbildungen $V \rightarrow K$; sie heissen Linearformen.

Beispiele 6.2.2.

- (1) Es sei $n \geq 1$. Dann ist die Spurabbildung

$$\text{Tr} : M_{n \times n}(K) \rightarrow K, \quad A = (a_{ij}) \mapsto a_{11} + \cdots + a_{nn}$$

eine Linearform.

- (2) Es sei V ein endlich-dimensionaler Vektorraum mit Basis $v_1 \dots v_n$. Es sei $\ell : V \rightarrow K$ die Abbildung, die einen Vektor

$$v = a_1 v_1 + \cdots + a_n v_n$$

auf a_1 abbildet. Dann ist ℓ eine Linearform.

- (3) Es sei V der reelle Vektorraum aller stetigen Funktionen $[0, 1] \rightarrow \mathbb{R}$, und es sei $a \in [0, 1]$. Dann ist die Abbildung

$$f \mapsto f(a)$$

eine Linearform.

- (4) Es sei V der reelle Vektorraum aller differenzierbaren Funktionen $(0, 1) \rightarrow \mathbb{R}$, und es sei $a \in (0, 1)$. Dann ist

$$f \mapsto (Df)(a),$$

mit Df die Ableitung von f , eine Linearform.

- (5) Es sei V der reelle Vektorraum aller integrierbaren Funktionen $[0, 1] \rightarrow \mathbb{R}$. Dann ist

$$f \mapsto \int_0^1 f(x) dx$$

eine Linearform.

Beachte 6.2.3. Es sei V ein endlich-dimensionaler Vektorraum mit Basis $\mathcal{B} = (v_1, \dots, v_n)$, und es sei $\ell \in V^*$. Dann ist ℓ durch $\ell(v_1), \dots, \ell(v_n)$ eindeutig bestimmt (Korollar 4.1.8). Insbesondere sind zwei Linearformen $\ell, \lambda \in V^*$ genau dann gleich, wenn

$$\ell(v_i) = \lambda(v_i) \quad \forall 1 \leq i \leq n$$

gilt. Wir werden diese Beobachtung wieder und wieder benutzen.

Definition 6.2.4. Es sei V ein endlich-dimensionaler K -Vektorraum und $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis fuer V . Fuer $1 \leq i \leq n$, definiere eine Linearform $v_i^* \in V^*$ wie folgt:

$$v_i^*(v_j) = \begin{cases} 1 & \text{wenn } i = j \\ 0 & \text{wenn } i \neq j. \end{cases}$$

Explizit bildet die Linearform v_i^* einen Vektor $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ auf den Koeffizienten α_i ab.

Satz 6.2.5. Die Elemente $v_1^*, \dots, v_n^*$ sind eine Basis von V^* . Insbesondere ist V^* ebenfalls n -dimensional.

Proof. Es sei $\ell : V \rightarrow K$ eine Linearform; wir muessen zeigen, dass sich ℓ eindeutig als Linearkombination von $v_1^*, \dots, v_n^*$ schreiben laesst. Definiere die Linearform

$$(27) \quad f = \ell(v_1)v_1^* + \cdots + \ell(v_n)v_n^*.$$

Dann gilt $f(v_i) = \ell(v_i) \forall 1 \leq i \leq n$, und daher gilt $\ell = f$ dank Beobachtung 6.2.3. Daher ist $v_1^*, \dots, v_n^*$ ein Erzeugendensystem von V^* .

Nimm nun an, dass es Skalare $\alpha_1, \dots, \alpha_n \in K$ gibt, so dass

$$(28) \quad \alpha_1 v_1^* + \cdots + \alpha_n v_n^* = 0_{V^*};$$

hier bezeichnet 0_{V^*} die Linearform, die jeden Vektor $v \in V$ auf $0 \in K$ abbildet. Insbesondere koennen wir (28) auf v_i auswerten: es gilt

$$(\alpha_1 v_1^* + \cdots + \alpha_n v_n^*)(v_i) = 0_{V^*}(v_i)$$

Bei der Definition der dualen Basis ist die linke Seite gleich α_i , und die rechte Seite gleich 0. Daher gilt

$$\alpha_1 = \cdots = \alpha_n = 0,$$

das heisst, die Vektoren $v_1^*, \dots, v_n^*$ sind linear unabhangig und daher eine Basis.³⁰ $\square$

Beispiel 6.2.6. Es sei $V = \mathbb{R}[x]^{\leq 2}$ mit Standardbasis e_0, e_1, e_2 mit $e_i = x^i$. Wir betrachten folgende Linearform

$$\ell : V \rightarrow \mathbb{R}. \quad \ell(f(x)) = Df(2),$$

wobei D die Ableitungsabbildung ist. Dann gilt

$$\ell = 0 \cdot e_0^* + 1 \cdot e_1^* + 4 \cdot e_2^*.$$

Mit anderen Worten, wenn wir ein beliebiges Element $y \in V$ als Linearkombination der Basisvektoren schreiben, $y = a + bx + cx^2 = ae_0 + be_1 + ce_2$, dann gilt

$$\ell(y) = 0 \cdot a + 1 \cdot b + 4c.$$

Definition 6.2.7. Die Basis $\mathcal{B}^* = (v_1^*, \dots, v_n^*)$ ist die duale Basis von $\mathcal{B}$.

Es seien nun $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{C} = (w_1, \dots, w_n)$ Basen von V . Wir erinnern uns: die Basiswechselmatrix $A = [\text{id}]_{\mathcal{C}}^{\mathcal{B}}$ ist die Matrix $A = (a_{ij})$, deren Eintrage definiert sind durch die Gleichungen

$$(29) \quad v_j = \sum_{i=1}^n a_{ij} w_i.$$

Was gilt nun fuer die Basiswechselmatrizen $[\text{id}]_{\mathcal{C}^*}^{\mathcal{B}^*}$ und $[\text{id}]_{\mathcal{B}^*}^{\mathcal{C}^*}$?

Beispiel 6.2.8. Es sei $V = \mathbb{R}^2$, und es seien $\mathcal{B} = (v_1, v_2)$ mit $v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und $v_2 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$ und $\mathcal{C}$ die Standardbasis. Dann ist die Basiswechselmatrix gegeben durch

$$[\text{id}]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix},$$

das heisst

$$(30) \quad v_1 = e_1 + 2e_2 \quad \text{und} \quad v_2 = 3e_1 - e_2.$$

Betrachten wir nun die dualen Basen $\mathcal{B}^* = (v_1^*, v_2^*)$ und $\mathcal{C}^* = (e_1^*, e_2^*)$. Schreiben wir $[\text{id}]_{\mathcal{C}^*}^{\mathcal{B}^*} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, so gilt

$$(31) \quad v_1^* = ae_1^* + ce_2^* \quad \text{und} \quad v_2^* = be_1^* + de_2^*.$$

Wenden wir (31) auf (30) an, so erhalten wir

$$[\text{id}]_{\mathcal{C}^*}^{\mathcal{B}^*} = \begin{pmatrix} \frac{1}{7} & \frac{2}{7} \\ \frac{3}{7} & -\frac{1}{7} \end{pmatrix}.$$

Ok. Vielleicht haben wir mit $[\text{id}]_{\mathcal{B}^*}^{\mathcal{C}^*}$ mehr Glueck? Schreiben wir $[\text{id}]_{\mathcal{B}^*}^{\mathcal{C}^*} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, so gilt

$$(32) \quad e_1^* = \alpha v_1^* + \gamma v_2^* \quad \text{und} \quad e_2^* = \beta v_1^* + \delta v_2^*.$$

Wenden wir (32) auf (30) an, so erhalten wir

$$[\text{id}]_{\mathcal{B}^*}^{\mathcal{C}^*} = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}.$$

Faellt Ihnen was auf?

Lecture 19

Satz 6.2.9. Es sei V ein endlich-dimensionaler K -Vektorraum mit Basen $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{C} = (w_1, \dots, w_n)$. Dann gilt

$$[\text{id}]_{\mathcal{B}^*}^{\mathcal{C}^*} = ([\text{id}]_{\mathcal{C}}^{\mathcal{B}})^t.$$

³⁰Anstatt die lineare Unabhangigkeit zu beweisen, haetten wir ebenfalls Korollar 6.1.4 benutzen koennen: $v_1^*, \dots, v_n^*$ ist ein Erzeugendensystem mit n Elementen in einem n -dimensionalen Vektorraum und daher eine Basis.

Proof. Wir argumentieren wie in dem Beispiel. Schreibe $[\text{id}]_{\mathcal{C}}^{\mathcal{B}} = (a_{ij})$ und $[\text{id}]_{\mathcal{B}^*}^{\mathcal{C}^*} = (b_{ij})$, das heisst

$$(33) \quad v_j = \sum_{i=1}^n a_{ij} w_i,$$

$$(34) \quad w_\ell^* = \sum_{k=1}^n b_{k\ell} v_k^*.$$

Wenden wir nun (34) auf (33) an, so erhalten wir

$$\left(\sum_{k=1}^n b_{k\ell} v_k^* \right) (v_j) = w_\ell^* \left(\sum_{i=1}^n a_{ij} w_i \right).$$

Doch nun ist $v_k^*(v_j)$ (bzw. $w_\ell^*(w_j)$) nur dann nicht null, wenn $k = j$ (bzw. wenn $i = \ell$), und daher erhalten wir

$$b_{j\ell} = a_{\ell j} \quad \forall j, \ell.$$

□

Korollar 6.2.10. *Weiterhin gilt*

$$[\text{id}]_{\mathcal{C}^*}^{\mathcal{B}^*} = ([\text{id}]_{\mathcal{B}}^{\mathcal{C}})^t = \left[([\text{id}]_{\mathcal{C}}^{\mathcal{B}})^{-1} \right]^t.$$

Proof. Folgt unmittelbar von Satz 4.5.3 und Lemma 4.7.10 (3). □

Beispiel 6.2.11. Wir ueberpruefen die Rechnung in Beispiel 6.2.8: es gilt

$$([\text{id}]_{\mathcal{C}^*}^{\mathcal{B}^*})^t = \begin{pmatrix} \frac{1}{\frac{2}{7}} & \frac{3}{\frac{7}{1}} \\ \frac{2}{7} & -\frac{1}{7} \end{pmatrix}$$

und weiterhin

$$\begin{pmatrix} \frac{1}{\frac{2}{7}} & \frac{3}{\frac{7}{1}} \\ \frac{2}{7} & -\frac{1}{7} \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{\frac{2}{7}} & \frac{3}{\frac{7}{1}} \\ \frac{2}{7} & -\frac{1}{7} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

6.3. Die duale Abbildung.

Definition 6.3.1. Es seien V, W K -Vektorraeume, und es sei $T : V \rightarrow W$ eine lineare Abbildung. Definiere die duale Abbildung

$$T^* : W^* \rightarrow V^*, \quad \ell \mapsto \ell \circ T.$$

Bemerkung 6.3.2. Explizit ist $T^*(\ell)$ folgendermassen definiert: fuer $v \in V$ gilt

$$T^*(\ell)(v) = \ell(T(v)).$$

Lemma 6.3.3. Die duale Abbildung ist wohldefiniert: $T^*(\ell)$ ist eine Linearform $V \rightarrow K$.

Proof. Es sei $\ell \in V^*$. Da sowohl T und ℓ lineare Abbildungen sind, folge von Lemma 4.1.10, dass $\ell \circ T$ ebenfalls linear ist, d.h. $T^*(\ell) \in V^*$. $\square$

Satz 6.3.4. Die duale Abbildung $T^* : W^* \rightarrow V^*$ ist eine lineare Abbildung.

Proof. Wir ueberpruefen die Axiome:

- (1) Es seien $\alpha \in K$ und $\ell \in W^*$. Dann gilt fuer all $v \in V$

$$T^*(\alpha\ell)(v) = ((\alpha\ell) \circ T)(v) = (\alpha\ell)(T(v)) = \alpha \cdot \ell(T(v)) = \alpha(\ell \circ T)(v) = \alpha T^*(\ell)(v),$$

das heisst $T^*(\alpha\ell) = \alpha T^*(\ell)$.³¹

- (2) Es seien $\ell_1, \ell_2 \in W^*$. Dann gilt fuer all $v \in V$

$$\begin{aligned} T^*(\ell_1 + \ell_2)(v) &= ((\ell_1 + \ell_2) \circ T)(v) \\ &= (\ell_1 + \ell_2)(T(v)) \\ &= \ell_1(T(v)) + \ell_2(T(v)) \\ &= (\ell_1 \circ T)(v) + (\ell_2 \circ T)(v) \\ &= T^*(\ell_1)(v) + T^*(\ell_2)(v), \end{aligned}$$

und daher $T^*(\ell_1 + \ell_2) = T^*(\ell_1) + T^*(\ell_2)$. $\square$

Satz 6.3.5. Es seien U, V, W Vektorraeume ueber K , und es seien $T : U \rightarrow V$ und $S : V \rightarrow W$ lineare Abbildungen. Dann gilt

$$(S \circ T)^* = T^* \circ S^*.$$

Proof. Uebung. $\square$

Beispiele 6.3.6.

- (1) Es gilt $(\text{id}_V)^* = \text{id}_{V^*}$.
(2) Es gilt $(0 : V \rightarrow W)^* = (0 : W^* \rightarrow V^*)$.

Bemerkung 6.3.7. Wir koennen die Ergebnisse dieses Kapitels folgendermassen zusammenfassen: durch Dualisierung erhalten wir aus einem Vektorraum V einen neuen Vektorraum V^* , und aus einer linearen Abbildung $T : V \rightarrow W$ eine neue lineare Abbildung $T^* : W^* \rightarrow V^*$, wobei sich die Richtung der Abbildung umkehrt: Dualisierung ist ein Beispiel eines kontravarianten Funktors.

Wir haben in Definition 6.2.7 gesehen, dass eine Basis von V eine Basis von V^* hervorbringt, naemlich die duale Basis. Es sei nun $T : V \rightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen K -Vektorraeumen, und es seien $\mathcal{B}$ und $\mathcal{C}$ Basen von V und W . Wir koennen dann die Abbildungsmatrix $[T]_{\mathcal{C}}^{\mathcal{B}}$ von T bezueglich $\mathcal{B}$ und $\mathcal{C}$ betrachten. Was koennen wir ueber die Abbildungsmatrix $[T^*]_{\mathcal{B}^*}^{\mathcal{C}^*}$ sagen?

Beispiel 6.3.8. Es sei $V = \mathbb{R}^2$, und es sei $T : V \rightarrow V$ die Abbildung, die bezueglich der Standardbasis $\mathcal{E} = (e_1, e_2)$ durch die Matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}$ gegeben ist, d.h.

$$(35) \quad T(e_1) = e_1 + 3e_2 \quad \text{und} \quad T(e_2) = 2e_1 - e_2.$$

Es sei $B = [T^*]_{\mathcal{E}^*}^{\mathcal{E}^*}$, d.h. wenn $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, dann gilt

$$(36) \quad T^*(e_1^*) = ae_1^* + ce_2^* \quad \text{und} \quad T^*(e_2^*) = be_1^* + de_2^*.$$

³¹Wir benutzen hier, dass zwei Elemente von V^* gleich sind, wenn sie fuer alle $v \in V$ den gleichen Wert annehmen.

Um die Einträge von B zu bestimmen, benutzen wir wieder die definierende Eigenschaft der dualen Basis. Indem wir (36) auf e_1 und e_2 anwenden, erhalten wir

$$\begin{aligned} T^*(e_1^*)(e_1) &= a = (e_1^* \circ T)(e_1) = e_1^*(e_1 + 2e_2) = 1, \\ T^*(e_1^*)(e_2) &= c = (e_1^* \circ T)(e_2) = e_1^*(2e_1 - e_2) = 2, \\ T^*(e_2^*)(e_1) &= b = (e_2^* \circ T)(e_1) = e_2^*(e_1 + 3e_2) = 3, \\ T^*(e_2^*)(e_2) &= d = (e_2^* \circ T)(e_2) = e_2^*(3e_1 - e_2) = -1, \end{aligned}$$

mit anderen Worten $B = [T^*]_{\mathcal{E}^*}^{\mathcal{E}} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}$. Faellt Ihnen was auf?

Theorem 6.3.9. *Es seien V, W endlich-dimensionale K -Vektorraeume, und es sei $\mathcal{B} = (v_1, \dots, v_n)$, bzw. $\mathcal{C} = (w_1, \dots, w_m)$, eine Basis for V , bzw. fuer W . Es sei $T : V \rightarrow W$ eine lineare Abbildung. Dann gilt³²*

$$[T^*]_{\mathcal{B}^*}^{\mathcal{C}^*} = ([T]_{\mathcal{C}}^{\mathcal{B}})^t.$$

Proof. Es sei

$$A = [T]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

das heisst

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

Was ist $T^*(w_k^*)$? Wir benutzen (27): es gilt³³

$$T^*(w_k^*) = [(T^*(w_k^*)) (v_1)] \cdot v_1^* + \dots + [(T^*(w_k^*)) (v_n)] \cdot v_n^*,$$

und von der Definition der dualen Abbildung T^* sehen wir, dass

$$T^*(w_k^*)(v_j) = w_k^*(T(v_j)) = w_k^*\left(\sum_{i=1}^m a_{ij} w_i\right) = a_{kj},$$

das heisst

$$T^*(w_k^*) = a_{k1} v_1^* + \dots + a_{kn} v_n^*.$$

Daher ist die Abbildungsmatrix $[T^*]_{\mathcal{B}^*}^{\mathcal{C}^*}$ gegeben durch

$$[T^*]_{\mathcal{B}^*}^{\mathcal{C}^*} = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix} = A^t$$

□

Bemerkung 6.3.10. *Es sei $A \in M_{m \times n}(K)$, und es sei $T_A : K^n \rightarrow K^m$ die bezueglich der Standardbasen $\mathcal{E}$ und $\mathcal{F}$ von K^n und K^m . Dann ist $[T_A]_{\mathcal{F}^*}^{\mathcal{E}^*} = A^t$. Vielleicht hilft uns das, um einen neuen Beweis zu finden fuer den Satz, dass*

$$\text{Spaltenrang}(A) = \text{Spaltenrang}(A^t)?$$

Wir bemerken, dass $\text{Spaltenrang}(A) = \dim \text{im}(T_A)$ und $\text{Spaltenrang}(A^t) = \dim \text{im}(T_A^)$. Was koennen wir ueber die Dimensionen dieser Raeume sagen?*

Bemerkung 6.3.11. *Es ist vielleicht nicht ganz klar, warum es sich lohnt, Dualraeume zu betrachten. Meine Antwort darauf ist: Dualraeume erklaren, wie transponierte Matrizen auf natuerliche Art und Weise in der Theorie von Vektorraeumen auftreten, naemlich als Abbildungs- und Basiswechselmatrizen von dualen Abbildungen und dualen Basen.*

³²Sie koennen sich das folgendermassen merken: die Abbildungsmatrix von T ist eine $(m \times n)$ -Matrix, daher muss die Abbildungsmatrix von T^* eine $(n \times m)$ -Matrix sein. Deine einfachste Art und Weise, von einer $(m \times n)$ -Matrix eine $(n \times m)$ -Matrix zu erhalten, ist die Transposition.

³³Beachten Sie, dass $T^*(w_k^*) \in V^*$, d.h. wir koennen $T^*(w_k^*)$ an einem Element $v \in V$ auswerten und erhalten einen Skalar: $(T^*(w_k^*))(v) \in K$.

6.4. Annulatoren.

Definition 6.4.1. Es sei V ein Vektorraum und $U \leq V$ ein Unterraum. Der Annulator von U ist die Menge aller Linearformen von V , die alle Elemente von U auf Null abbilden, das heisst

$$U^\circ = \{\ell \in V^* : \ell(u) = 0_V \quad \forall u \in U\}.$$

Mit anderen Worten, ein Element $\ell \in V^*$ gehoert genau dann zu U° , wenn $U \subseteq \ker(\ell)$.

Lemma 6.4.2. Es sei V ein Vektorraum und $U \leq V$. Dann ist U° ein Unterraum von V^* .

Beispiel 6.4.3. Es gilt $\{0_V\}^\circ = V^*$.

Beispiel 6.4.4.

- (1) Es sei $V = \mathbb{R}^3$ und $U = \left\langle \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\rangle$. Was ist $U^\perp$?

Es sei $\mathcal{B} = (e_1, e_2, e_3)$ die Standardbasis von V , so dass $U = \langle e_1 - e_2 + e_3 \rangle$ und $\mathcal{B}^*$ die Dualbasis. Dann ist

$$U^\circ = \{\ell \in V^* : \ell(e_1 - e_2 + e_3) = 0\}.$$

Mit anderen Worten, wenn $\ell = ae_1^* + be_2^* + ce_3^*$, dann gilt

$$\ell(e_1 - e_2 + e_3) = a - b + c = 0,$$

das heisst $b = a + c$, und

$$U^\circ = \{a(e_1^* + e_3^*) + c(e_2^* + e_3^*) : a, c \in \mathbb{R}\}.$$

- (2) Es sei nun $W = \langle e_1 - e_2 + e_3, e_1 + e_2 \rangle \leq V$. Dann ist

$$W^\circ = \{\ell \in V^* : \ell(e_1 - e_2 + e_3) = 0 \quad \text{und} \quad \ell(e_1 + e_2) = 0\}.$$

Schreiben wir $\ell = ae_1^* + be_2^* + ce_3^*$, dann lauten die Bedingungen

$$a - b + c = 0 \quad \text{und} \quad a + b = 0,$$

das heisst $b = -a$ und $c = -2a$, und

$$W^\circ = \{a(e_1^* - e_2^* - 2e_3^*) : a \in \mathbb{R}\}.$$

Was faellt Ihnen auf, wenn Sie die Dimensionen betrachten?

Theorem 6.4.5. Es sei V ein Vektorraum und $U \leq V$. Dann gilt

$$\dim_K(U) + \dim_K(U^\circ) = \dim_K V.$$

Proof. Es sei $n = \dim_K V$ und $\dim_K U = k$. Waehle eine Basis $u_1, \dots, u_k$ von U ; dann koennen wir sie laut Theorem 3.3.19 zu einer Basis $\mathcal{B} = (u_1, \dots, u_k, v_1, \dots, v_{n-k})$ von V erweitern. Es sei $\mathcal{B}^* = (u_1^*, \dots, u_k^*, v_1^*, \dots, v_{n-k}^*)$ die duale Basis. Dann gilt aufgrund der definierenden Eigenschaft von $\mathcal{B}^*$ dass $v_j^*(u_i) = 0$ fuer all $1 \leq i \leq k$ und $1 \leq j \leq n - k$, mit anderen Worten, $v_j^* \in U^\circ \quad \forall 1 \leq j \leq n - k$ und daher $\langle v_1^*, \dots, v_{n-k}^* \rangle \subseteq U^\circ$.

Wir behaupten, dass $v_1^*, \dots, v_{n-k}^*$ eine Basis von U° ist. Natuerlich sind die Vektoren linear unabhengig, es ist daher zu zeigen, dass $\langle v_1^*, \dots, v_{n-k}^* \rangle = U^\circ$. Es sei $\lambda \in U^\circ$. Da $\mathcal{B}^*$ eine Basis von V^* ist, gibt es Skalare $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_{n-k}$, so dass

$$\lambda = \alpha_1 u_1^* + \dots + \alpha_k u_k^* + \beta_1 v_1^* + \dots + \beta_{n-k} v_{n-k}^*.$$

Da $\lambda \in U^\circ$, gilt

$$\lambda(u_i) = \alpha_i = 0 \quad \forall 1 \leq i \leq k,$$

und wir erhalten

$$\lambda = \beta_1 v_1^* + \dots + \beta_{n-k} v_{n-k}^*,$$

was zu beweisen war. □

Hier ist ein alternativer, sehr schoener Beweis von Ande Wu, der Theorem 6.4.5 von Theorem 4.2.9 ableitet:

Proof. Es sei $\iota : U \rightarrow V$ die Inklusionsabbildung und $\iota^* : V^* \rightarrow U^*$ die duale Abbildung. Dann gilt laut Theorem 4.2.9 die Formel

$$(37) \quad \dim(V) = \dim(V^*) = \dim \ker(\iota^*) + \dim \operatorname{im}(\iota^*)$$

Behauptung 1: es gilt $\operatorname{im}(\iota^*) = U^*$

Beweis von Behauptung 1: Per Definition gilt $\operatorname{im}(\iota^*) \subseteq U^*$, da heisst, wir muessen die umgekehrte Inklusion zeigen. Es sei $P : V \rightarrow U$ eine Projektion³⁴, d.h. $P(u) = u$ fuer alle $u \in U$. Es sei $\ell \in U^*$. Dann gilt $\ell \circ P \in V^*$, und eine einfache Rechnung zeigt, dass

$$\iota^*(\ell \circ P) = \ell,$$

das heisst $\ell \in \operatorname{im}(\iota^*)$.

Behauptung 2: es gilt $\ker(\iota^*) = U^\circ$.

Beweis von Behauptung 2: es gilt

$$\begin{aligned} \ell \in \ker(\iota^*) &\Leftrightarrow \iota^*(\ell) = \ell \circ \iota = 0_{V^*} \\ &\Leftrightarrow \forall u \in U : \ell(u) = \ell(\iota(u)) = 0_{V^*}(u) = 0 \\ &\Leftrightarrow \ell \in U^\circ. \end{aligned}$$

Von (37) und den beiden Behauptungen folgt daher, dass

$$\dim(V) = \dim(U^*) + \dim(U^\circ).$$

Doch da $\dim(U^*) = \dim(U)$, folgt das Theorem. $\square$

Satz 6.4.6. *Es seien V, W endlich-dimensionale K -Vektorraeume und $T : V \rightarrow W$ eine lineare Abbildung. Dann gilt*

- (1) $(\operatorname{im} T)^\circ = \ker(T^*);$
- (2) $(\ker T)^\circ = \operatorname{im}(T^*).$

Proof.

- (1) Es sei $\lambda \in \ker(T^*)$. Dann ist per Definition $T^*(\lambda) = \lambda \circ T = 0_{V^*}$, das heisst

$$\lambda(T(v)) = 0 \quad \forall v \in V$$

und daher $\lambda \in (\operatorname{im} T)^\circ$.

Umgekehrt sei $\lambda \in (\operatorname{im} T)^\circ$, das heisst $\lambda(T(v)) = 0 \quad \forall v \in V$. Daraus folgt, dass $\lambda \circ T = 0_{V^*}$ und somit $T^*(\lambda) = 0_{V^*}$, das heisst $\lambda \in \ker(T^*)$.

- (2) Es sei $\lambda \in \operatorname{im}(T^*)$, das heisst, es gibt $\psi \in W^*$ so dass $T^*(\psi) = \lambda$. Also gilt

$$\lambda(v) = T^*(\psi)(v) = \psi(T(v)) = 0 \quad \forall v \in \ker(T),$$

das heisst $\lambda \in (\ker T)^\circ$ und daher $\operatorname{im}(T^*) \subseteq (\ker T)^\circ$.

Anstatt die umgekehrte Inklusion zu zeigen, berechnen wir die Dimension von $\operatorname{im}(T^*)$: es gilt

$$\begin{aligned} \dim \operatorname{im}(T^*) &= \dim W - \dim \ker(T^*) && \text{Theorem 4.2.9} \\ &= \dim W - \dim(\operatorname{im} T)^\circ && (1) \\ &= \dim W - (\dim W - \dim \operatorname{im}(T)) && \text{Theorem 6.4.5} \\ &= \dim \operatorname{im}(T) \\ &= \dim V - \dim \ker(T) && \text{Theorem 4.2.9} \\ &= \dim(\ker T)^\circ && \text{Theorem 6.4.5} \end{aligned}$$

und daher $\operatorname{im}(T^*) = (\ker T)^\circ$. $\square$

Bemerkung 6.4.7. *Insbesondere gilt $\dim \operatorname{im}(T^*) = \dim \operatorname{im}(T)$.*

Korollar 6.4.8. *Es sei $A \in M_{m \times n}(K)$. Dann gilt*

$$\text{Zeilenrang}(A) = \text{Spaltenrang}(A).$$

Proof. Es sei $T_A : K^n \rightarrow K^m$ wie in Definition 4.1.4. Dann gilt

$$\dim \operatorname{im}(T_A) = \text{Spaltenrang}(A) \quad \text{und} \quad \dim \operatorname{im}(T_A^*) = \text{Spaltenrang}(A^t) = \text{Zeilenrang}(A).$$

Doch da $\dim \operatorname{im}(T_A) = \dim \operatorname{im}(T_A^*)$, folgt das Resultat. $\square$

³⁴Wir koennen z.B. eine solche Projektion dadurch konstruieren, indem wir ein Komplement von U betrachten.

Satz 6.4.6 hat noch weitere interessante Konsequenzen:

Korollar 6.4.9. *Es gilt*

- (1) T ist injektiv $\Leftrightarrow T^*$ ist surjektiv;
- (2) T ist surjektiv $\Leftrightarrow T^*$ ist injektiv.

Proof. Wir beweisen (1):

$$\begin{aligned} T \text{ ist injektiv} &\Leftrightarrow \ker(T) = \{0_V\} \\ &\Leftrightarrow (\ker T)^\circ = V^* \\ &\Leftrightarrow \operatorname{im}(T^*) = V^* && \text{Satz 6.4.6} \\ &\Leftrightarrow T^* \text{ ist surjektiv} \end{aligned}$$

□

6.5. **Reflexivitaet.** Was passiert, wenn wir einen Vektorraum zweimal dualisieren?

Definition 6.5.1. Es sei V ein Vektorraum. Dann ist der Bidualraum von V der K -Vektorraum $V^{**} = (V^*)^*$.

Was koennen wir ueber V^{**} sagen?

Bemerkung 6.5.2.

- (1) Wenn V n -dimensional ist, dann ist V^{**} aufgrund von Satz 6.2.5 ebenfalls n -dimensional.
- (2) Per Definition sind Elemente von V^{**} lineare Abbildungen $V^* \rightarrow K$. Koennen wir solche Abbildungen konstruieren?

Beachte 6.5.3. Es sei $v \in V$. Dann definiert die Auswertung an v eine Abbildung $V^* \rightarrow K$: explizit ist sie definiert durch

$$\tau_v : V^* \rightarrow K, \quad \ell \mapsto \ell(v).$$

Eine einfache Rechnung zeigt, dass $\tau_v(v)$ linear ist, das heisst, sie ist ein Element von V^{**} .

Theorem 6.5.4. Es sei V ein endlich-dimensionaler K -Vektorraum. Dann definiert die Abbildung $\tau : V \rightarrow V^{**}$, $V \mapsto \tau_v$ einen Isomorphismus zwischen V und seinem Bidualraum; wir sagen V ist reflexiv.

Bemerkung 6.5.5. Wenn V unendlich-dimensional ist, dann definiert τ immer noch eine Abbildung $V \rightarrow V^{**}$, aber sie ist nicht unbedingt surjektiv: nicht jeder unendlich-dimensionale Vektorraum ist reflexiv.

Proof. Wir zeigen zunaechst, dass τ eine lineare Abbildung ist.

- (1) Es seien $v \in V$ und $\lambda \in K$. Dann ist $\tau_{\lambda v}$ die Abbildung, die eine Linearform auf λv auswertet:

$$\tau_{\lambda v}(\ell) = \ell(\lambda v) = \lambda \ell(v) = \lambda \tau_v(\ell) \quad \forall \ell \in V^*,$$

das heisst $\tau_{\lambda v} = \lambda \tau_v$.

- (2) Es seien $u, v \in V$. Dann gilt

$$\tau_{u+v}(\ell) = \ell(u+v) = \ell(u) + \ell(v) = \tau_u(\ell) + \tau_v(\ell) \quad \forall \ell \in V^*,$$

und daher $\tau_{u+v} = \tau_u + \tau_v$.

Wir zeigen nun, dass die Abbildung ein Isomorphismus ist. Da $\dim V^{**} = \dim V$, ist es ausreichend zu zeigen, dass τ injektiv ist. Nimm an, dass $\tau_v = 0_{V^{**}}$, das heisst

$$\tau_v(\ell) = \ell(v) = 0 \quad \forall \ell \in V^*.$$

Es sei $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V , und schreibe $v = \alpha_1 v_1 + \dots + \alpha_n v_n$. Dann gilt

$$v_i^*(v) = \alpha_i = 0 \quad \forall 1 \leq i \leq n,$$

und daher $v = 0$. Daher ist τ injektiv und ein Isomorphismus. $\square$

Yannis Müller und Marco Vaccaro haben beide einen alternativen Beweis fuer Theorem 6.5.4 gefunden, der ohne die Wahl einer Basis auskommt: sie argumentieren statt dessen mit Hilfe von Annulatoren.

Proof. Nimm an, dass $\tau_v = 0_{V^{**}}$, d.h.

$$\tau_v(\ell) = 0 \quad \forall \ell \in V^* \quad \Leftrightarrow \quad \ell(v) = 0 \quad \forall \ell \in V^*.$$

Dann folgt aus der Linearitaet von ℓ , dass

$$\ell(w) = 0 \quad \forall w \in \langle v \rangle, \quad \forall \ell \in V^*,$$

das heisst $\ell \in \langle v \rangle^\circ \quad \forall \ell \in V^*$ und daher $V^* = \langle v \rangle^\circ$.

Allerdings gilt

$$\dim(V^*) = \dim(\langle v \rangle) + \dim(\langle v \rangle^\circ)$$

und daher $\dim(\langle v \rangle) = 0$, das heisst $v = 0_V$, was zu beweisen war. $\square$

Bemerkung 6.5.6. Die Abbildung $\tau : V \rightarrow V^{**}$ definiert einen kanonischen Isomorphismus zwischen V nach V^{**} : der Isomorphismus haengt nicht von der Wahl von Basen ab! Im Gegensatz dazu sind V und V^* nicht kanonisch isomorph: wir koennen zwar Isomorphismen zwischen den beiden Raeumen defineiren (wie zwischen allen K -Vektorraeumen, die die gleiche Dimension haben), aber diese haengen von der Wahl von Basen ab.

Lemma 6.5.7. *Es sei V ein endlich dimensionaler K -Vektorraum, und es sei $\mathcal{B}$ eine Basis. Es sei $\mathcal{B}^* = (v_1^*, \dots, v_n^*)$ die duale Basis, und $\mathcal{B}^{**} = ((v_1^*)^*, \dots, (v_n^*)^*)$ die duale Basis von $\mathcal{B}^*$. Dann gilt*

$$\tau_{v_i} = (v_i^*)^* \quad \forall i$$

*das heisst, τ bildet die Basis $\mathcal{B}$ auf die biduale Basis $\mathcal{B}^{**}$ ab.*

Proof. Per Definition ist $\mathcal{B}^{**} = ((v_1^*)^*, \dots, (v_n^*)^*)$ diejenige Basis von V^{**} , fuer die gilt

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

Wir muessen also zeigen, dass $\{\tau_{v_1}, \dots, \tau_{v_n}\}$ diese Eigenschaft hat. Aber nun gilt

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

Quod erat demonstrandum. □

Bemerkung 6.5.8. *Es seien V, W endlich dimensionale Vektorraume und $T : V \rightarrow W$ eine lineare Abbildung. Dann sind die Abbildungen τ^V und τ^W mit T kompatibel: wir haben ein kommutative Diagramm*

$$\begin{array}{ccc} V & \xrightarrow{\tau^V} & V^{**} \\ T \downarrow & & \downarrow T^{**} \\ W & \xrightarrow{\tau^W} & W^{**} \end{array}$$

$V \mapsto V^{**}$ ist ein kovarianter Funktor. Nun seien $\mathcal{B}$ und $\mathcal{C}$ jeweilige Basen von V und W . Dann gilt

$$[T^{**}]_{\mathcal{C}^{**}}^{\mathcal{B}^{**}} = \left([T^*]_{\mathcal{B}^*}^{\mathcal{C}^*}\right)^t = \left(\left([T]_{\mathcal{C}}^{\mathcal{B}}\right)^t\right)^t = [T]_{\mathcal{C}}^{\mathcal{B}},$$

7. QUOTIENTENRAEUME

7.1. Definition und erste Eigenschaften. Wir erinnern uns an die Definition von Nebenklassen eines Unterraums:

Bemerkung 7.1.1. Es sei V ein Vektorraum und $U \leq V$. Die Nebenklassen von U in V sind die Äquivalenzklassen folgender Äquivalenzrelation auf V : " $v \sim w$ genau dann, wenn $v - w \in U$ ". Diese bedeutet, dass jedes $v \in V$ in genau einer Nebenklasse liegt, naemlich in der Nebenklasse $[v] = v + U$. Wir nennen v ein repraesentierendes Element dieser Nebenklasse; das bedeutet, zwei Elemente $v, v' \in V$ repraesentieren die selbe Nebenklasse genau dann (d.h. $[v] = [v']$), wenn $v - v' \in U$.

Definition 7.1.2. Wir definieren den Quotientenraum V/U wie folgt: die Elemente von V/U sind die Nebenklassen von U in V . Die Addition und Skalarmultiplikation sind definiert durch

$$[v_1] + [v_2] = [v_1 + v_2] \quad \text{und} \quad \alpha[v] = [\alpha v].$$

Wir muessen zeigen, dass Addition und Multiplikation wohldefiniert sind:

- (1) es seien $v_1, v'_1, v_2, v'_2 \in V$, und nimm an, dass $[v_1] = [v'_1]$ und $[v_2] = [v'_2]$. Wir muessen zeigen, dass

$$[v_1] + [v_2] = [v'_1] + [v'_2] \Leftrightarrow [v_1 + v_2] = [v'_1 + v'_2] \Leftrightarrow v_1 + v_2 - v'_1 - v'_2 \in U.$$

Doch das ist klar, da $v_1 - v'_1 \in U$ und $v_2 - v'_2 \in U$.

- (2) Es seien $v, v' \in V$ so dass $[v] = [v']$, und es sei $\alpha \in K$. Wir muessen zeigen, dass

$$\alpha[v] = \alpha[v'] \Leftrightarrow [\alpha v] = [\alpha v'] \Leftrightarrow \alpha(v - v') \in U.$$

Doch da $v - v' \in U$ und U ein Unterraum ist, gilt $\alpha(v - v') \in U$.

Satz 7.1.3. Mit dieser Definition der Addition und Skalar-Multiplikation ist V/U ein Vektorraum.

Proof. Uebung.³⁵ □

Satz 7.1.4. Es sei $q_U : V \rightarrow V/U$ die Abbildung

$$q_U(v) = [v];$$

sie heisst die kanonische Quotientenabbildung. Dann ist q_U linear, und es gilt³⁶ $\ker(q_U) = U$ und $\text{im}(q_U) = V/U$.

Proof. Wir zeigen zunaechst, dass die Abbildung linear ist:

$$\begin{aligned} q_U(v_1 + \alpha v_2) &= [v_1 + \alpha v_2] \\ &= [v_1] + \alpha[v_2] \quad \text{aufgrund der Definition von Addition und Skalar-Multiplikation} \\ &= q_U(v_1) + \alpha q_U(v_2). \end{aligned}$$

Es sei nun $x \in V/U$. Dann ist x eine Äquivalenzklasse der Äquivalenzrelation in Bemerkung 7.1.1, und es gibt $v \in V$ so dass $x = [v]$, das heisst $x = q_U(v)$. Daher ist q_U surjektiv.

Nimm nun an, dass $v \in \ker(q_U)$. Dann gilt

$$\begin{aligned} q_U(v) = [v] = 0_{V/U} &\Leftrightarrow v \sim 0_V \\ &\Leftrightarrow v - 0_V \in U \\ &\Leftrightarrow v \in U, \end{aligned}$$

und daher $\ker(q_U) \subseteq U$. Umgekehrt sei $u \in U$. Dann gilt $u \sim 0_V$ da $u = u - 0_V \in U$ und daher $q_U(u) = [u] = 0_{V/U}$. Daher gilt $U \subseteq \ker(q_U)$ und daher $\ker(q_U) = U$. □

Korollar 7.1.5. Es sei V ein endlich-dimensionaler Vektorraum und $U \leq V$. Dann ist V/U endlich-dimensional, und es gilt

$$\dim_K V/U = \dim_K V - \dim_K U.$$

Proof. Folgt unmittelbar von Satz 7.1.4 und Theorem 4.2.9. □

Alternativ koennen wir dieses Korollar ebenfalls mit Hilfe eines Komplements des Unterraums U beweisen:

³⁵Was ist die additive Identitaet?

³⁶Mit anderen Worten, jeder Unterraum ist der Kern einer linearen Abbildung!

Satz 7.1.6. Es sei V ein Vektorraum und $U \leq V$. Es sei $W \leq V$ ein Komplement zu U . Dann definiert die Abbildung $w \mapsto [w]$ einen natuerlichen Isomorphismus

$$\gamma : W \xrightarrow{\cong} V/U.$$

Proof. Es sei $\gamma = q_U|_W$. Dann ist γ injektiv: nimm an, dass $w \in \ker(\gamma)$, das heisst $[w] = 0_{V/U}$, mit anderen Worten $w \in U$. Aber $U \cap W = \{0_V\}$, und daher $w = 0_V$.

Weiterhin ist γ surjektiv:³⁷ es sei $x \in V/U$, und es sei $v \in V$ ein repraesentierendes Element von x . Da W ein Komplement von U ist, gilt $V = U + W$, und daher $\exists u \in U, w \in W$ so dass $v = u + w$. Aber dann gilt $[v] = [w]$, und daher $\gamma(w) = [v] = x$. $\square$

Lecture 22

Übung 7.1.7. Beschreibe die inverse Abbildung $\gamma^{-1} : V/U \longrightarrow W$.

Bemerkung 7.1.8. Insbesondere bildet γ eine Basis von W auf einer Basis von V/U ab: wenn $w_1, \dots, w_k$ eine Basis von W ist, dann ist $[w_1], \dots, [w_k]$ eine Basis von V/U , und daher erhalten wir einen alternativen Beweis, dass

$$\dim V/U = \dim W = \dim V - \dim U.$$

Bemerkung 7.1.9. Der Isomorphismus in Satz 7.1.6 haengt von der Wahl eines Komplements ab; er ist also nicht kanonisch.

Weiterhin gibt es eine Korrespondenz zwischen bestimmten Unterraemen von V und Unterraemen von V/U :

Satz 7.1.10. Die Abbildung

$$\begin{aligned} \{W \subseteq V : U \leq W \leq V\} &\rightarrow \{X \leq V/U\} \\ W &\mapsto W/U \end{aligned}$$

ist eine 1 : 1 Korrespondenz zwischen Unterraemen von V , die U enthalten, und Unterraemen von V/U .

Proof. Übung. $\square$

³⁷Sie koennen alternativ mit Hilfe der Dimensionen argumentieren.

7.2. Die Isomorphismensätze. Quotientenräume tauchen in der linearen Algebra (und auch sonst) überall auf. Zum Beispiel können wir Theorem 4.2.9 folgendermassen verfeinern:

Theorem 7.2.1. (Erster Isomorphiesatz) Es seien V, W K -Vektorräume und $T : V \rightarrow W$ eine lineare Abbildung. Definiere eine neue Abbildung

$$\bar{T} : V/\ker(T) \rightarrow \operatorname{im}(T)$$

wie folgt: für ein Element $x \in V/\ker(T)$ wähle ein repräsentierendes Element $v \in V$ und definiere $\bar{T}(x) = T(v)$.³⁸ Dann gilt

- (1) $\bar{T}$ eine wohldefinierte lineare Abbildung;
- (2) $\bar{T} : V/\ker(T) \rightarrow \operatorname{im}(T)$ ist ein Isomorphismus;
- (3) folgendes Diagramm ist kommutativ:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ q_{\ker(T)} \downarrow & & \uparrow \subseteq \\ V/\ker(T) & \xrightarrow{\bar{T}} & \operatorname{im}(T) \end{array}$$

Proof. (1) Es seien v_1, v_2 repräsentierende Elemente von x . Dann gilt $q_{\ker(T)}(v_1) = q_{\ker(T)}(v_2) = x$, das heisst $v_1 - v_2 \in \ker(T)$ und daher $T(v_1) = T(v_2)$. Deshalb ist $\bar{T}$ wohldefiniert.

$\bar{T}$ ist linear: es seien $x = [v_1], y = [v_2]$ und $\alpha \in K$. Dann gilt

$$\begin{aligned} \bar{T}([v_1] + \alpha[v_2]) &= \bar{T}([v_1 + \alpha v_2]) \\ &= T(v_1 + \alpha v_2) \\ &= T(v_1) + \alpha T(v_2) \\ &= \bar{T}([v_1]) + \alpha \bar{T}([v_2]). \end{aligned}$$

- (2) $\bar{T}$ ist injektiv: Es sei $x \in \ker(\bar{T})$, und es sei $v \in V$ ein repräsentierendes Element von x . Dann gilt aufgrund der Definition

$$\bar{T}(x) = T(v) = 0_W,$$

das heisst $v \in \ker(T)$ und daher $x = [v] = 0_{V/\ker(T)}$.

$\bar{T}$ ist surjektiv: Aufgrund der Definition von $\bar{T}$ ist es offensichtlich, dass $\operatorname{im}(\bar{T}) \subseteq \operatorname{im}(T)$. Es sei $w \in \operatorname{im}(T)$. Dann gibt es ein $v \in V$ so dass $T(v) = w$. Aber dies bedeutet, dass

$$\bar{T}([v]) = T(v) = w,$$

und daher $w \in \operatorname{im}(\bar{T})$.

- (3) Die Kommutativität des Diagramms folgt unmittelbar von den Definitionen. □

Theorem 7.2.2. (Zweiter Isomorphiesatz) Es sei V ein Vektorraum und $U, W \leq V$. Es sei

$$i : U \hookrightarrow V \xrightarrow{q_W} V/W, \quad u \mapsto q_W(u)$$

Dann ist $\ker(i) = U \cap W$, und i induziert einen Isomorphismus

$$\bar{i} : U/(U \cap W) \xrightarrow{\cong} (U + W)/W.$$

Proof. Es folgt von Theorem 7.2.1, dass i einen Isomorphismus

$$\bar{i} : U/\ker(i) \xrightarrow{\cong} \operatorname{im}(i)$$

induziert. Wir zeigen zunächst, dass $\ker(i) = U \cap W$.

Es sei $u \in \ker(i)$. Dann gilt $i(u) = 0_{V/W}$, das heisst $u \in W$ und daher $u \in U \cap W$. Umgekehrt gilt $U \cap W \subseteq \ker(i)$, und daher $\ker(i) = U \cap W$.

Um den Beweis abzuschliessen müssen wir zeigen, dass $\operatorname{im}(i) = (U + W)/W$.³⁹ Es ist klar, dass $\operatorname{im}(i) \subseteq (U + W)/W$. Umgekehrt sei $x \in (U + W)/W$. Dann gibt es ein repräsentierendes Element $u \in U$ von x , und es gilt $i(u) = x$. □

³⁸Wir nennen $\bar{T}$ die von T induzierte Abbildung.

³⁹Beachte, dass $(U + W)/W$ ein Unterraum von V/W ist.

Was passiert, wenn wir zwei Unterräume von V betrachten, von dem einer in dem anderen enthalten ist?

Satz 7.2.3. *Es sei V ein Vektorraum und $U \leq W \leq V$ zwei Unterräume. Dann gibt es eine natürliche lineare Abbildung $\varpi_{U,W} : V/U \rightarrow V/W$, gegeben durch $v + U \mapsto v + W$, so dass folgendes Diagramm kommutiert:*

$$\begin{array}{ccc} & V & \\ q_U \swarrow & & \searrow q_W \\ V/U & \xrightarrow{\varpi_{U,W}} & V/W \end{array}$$

Proof. Eine einfache Rechnung zeigt, dass $\varpi_{U,W}$ linear ist. Die Kommutativität des Diagramms ist eine Übung. $\square$

Theorem 7.2.4. (Dritter Isomorphiesatz) *Es sei V ein Vektorraum und $U \leq W \leq V$ zwei Unterräume. Dann ist $\ker(\varpi_{U,W}) = W/U$, und die induzierte Abbildung $\overline{\varpi_{U,W}}$ definiert einen Isomorphismus*

$$\overline{\varpi_{U,W}} : (V/U)/(W/U) \xrightarrow{\cong} V/W.$$

Proof. Wir wenden Theorem 7.2.1 auf die induzierte Abbildung

$$\varpi_{U,W} : V/U \longrightarrow V/W$$

an. Die Abbildung ist surjektiv: Es sei $x \in V/W$ und v ein repräsentierendes Element von x . Dann wird die Nebenklasse $v + U$ auf x abgebildet.

Wir bestimmen nun den Kern von $\varpi_{U,W}$: es ist klar, dass $W/U \subseteq \ker(\varpi_{U,W})$. Es sei nun $x \in V/U$ im Kern von $\varpi_{U,W}$, und es sei v ein repräsentierendes Element von x , d.h. $q_U(v) = x$. Dann gilt aufgrund von Satz 7.2.3

$$\varpi_{U,W}(x) = q_W(v),$$

und wir wissen, dass $q_W(v) = 0_{V/W}$ genau dann, wenn $v \in W$. Es folgt, dass $x = q_U(v) \in W/U$, und daher $\ker(\varpi_{U,W}) \subseteq W/U$.

Es folgt daher von Theorem 7.2.1, dass $\overline{\varpi_{U,W}}$ einen Isomorphismus

$$(V/U)/(W/U) \xrightarrow{\cong} V/W$$

definiert. $\square$

7.3. Weitere Anwendungen. Weiterhin koennen wir den Annulator eines Unterraums mit Hilfe von Quotientenraeumen beschreiben:

Satz 7.3.1. *Es sei V ein K -Vektorraum und $U \leq V$ ein Unterraum. Dann gibt es einen natuerlichen Isomorphismus*

$$\alpha : (V/U)^* \xrightarrow{\cong} U^\circ.$$

Proof. Wir definieren zunaechst die Abbildung α : Es sei $\ell \in (V/U)^*$. Dann ist

$$\ell \circ q_U : V \rightarrow K$$

eine Linearform auf V , und es gilt $U \subseteq \ker(\ell \circ q_U)$, das heisst $\ell \circ q_U \in U^\circ$. Definieren $\alpha(\ell) = \ell \circ q_U$; eine einfache Rechnung zeigt, dass α linear ist. Da $(V/U)^*$ und U° die gleiche Dimension haben, ist es ausreichend zu zeigen, dass α injektiv ist. Dieses ist eine leichte Uebung.

Alternativ koennen Sie die inverse Abbildung konstruieren: es sei $\lambda \in U^\circ$, i.e. $\lambda \in V^*$ so dass $U \leq \ker(\lambda)$. Aufgrund von Theorem 7.2.1 wissen wir, dass λ eine Abbildung

$$\bar{\lambda} : V/\ker(\lambda) \rightarrow K$$

induziert. Da $U \leq \ker(\lambda)$, gibt es laut Satz 7.2.3 eine natuerliche Abbildung

$$\varpi_{U, \ker(\lambda)} : V/U \longrightarrow V/\ker(\lambda).$$

Definiere

$$\alpha^{-1}(\lambda) = \bar{\lambda} \circ \varpi_{U, \ker(\lambda)} : V/U \rightarrow K.$$

Eine einfache Rechnung zeigt, dass α^{-1} die inverse Abbildung von α ist. □

8.1. Ein erstes Beispiel.

Lemma 8.1.1. Es sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2 \times 2}(K)$. Dann ist A genau dann invertierbar, wenn $ad - bc \neq 0$, und in diesem Fall gilt

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Definition 8.1.2. Die Groesse $(ad - bc)$ heisst die Determinante von A , und wir schreiben

$$\det(A) = ad - bc.$$

Beachte 8.1.3. Schreibe $A = (\mathbf{v}_1, \mathbf{v}_2)$, wobei $\mathbf{v}_1, \mathbf{v}_2$ die Spaltenvektoren sind. Die Determinante hat folgende Eigenschaften:

- (1) $\det(I_2) = 1$;
- (2) $\det(\mathbf{v}_1, \mathbf{v}_2) = -\det(\mathbf{v}_2, \mathbf{v}_1)$;
- (3) $\det(\mathbf{v}_1 + \mathbf{u}, \mathbf{v}_2) = \det(\mathbf{v}_1, \mathbf{v}_2) + \det(\mathbf{u}, \mathbf{v}_2)$ und $\det(\mathbf{v}_1, \mathbf{v}_2 + \mathbf{u}) = \det(\mathbf{v}_1, \mathbf{v}_2) + \det(\mathbf{v}_1, \mathbf{u})$;
- (4) $\det(\lambda A) = \lambda^2 \det(A)$ fuer all $\lambda \in K$;
- (5) $\det(\lambda \mathbf{v}_1, \mathbf{v}_2) = \det(\mathbf{v}_1, \lambda \mathbf{v}_2) = \lambda \det(A)$;
- (6) $\det A = \det A^t$.

Diese Eigenschaften sind charakteristisch fuer eine Determinantenfunktion.

Um Determinanten fuer $(n \times n)$ -Matrizen zu definieren, benoetigen wir das Konzept von *Permutationen*

8.2. Permutationen.

Definition 8.2.1. Es sei $n \geq 1$. Eine Permutation der Menge $\{1, \dots, n\}$ ist eine bijektive Abbildung

$$\sigma : \{1, \dots, n\} \longrightarrow \{1, \dots, n\}.$$

Wir schreiben σ als $(\sigma(1), \sigma(2), \dots, \sigma(n))$.

Eine Transposition ist eine Permutation, die genau zwei Elemente vertauscht und alle anderen Elemente auch sich selber abbildet.

Beispiel 8.2.2. Es sei $n = 4$. Dann ist $(1, 3, 2, 4)$ eine Transposition, aber $(2, 3, 1, 4)$ ist keine Transposition.

Satz 8.2.3. Die Permutationen auf der Menge $\{1, \dots, n\}$ bilden eine Gruppe S_n unter Komposition von Funktionen; die Identitaet ist die Identitaetsfunktion, die jedes Element auf sich selber abbildet. Die Gruppe S_n heisst die symmetrische Gruppe vom Grad n ; sie hat $n!$ Elemente.

Proof. Uebung. □

Beachte 8.2.4. Es sei $\tau \in S_n$ eine Transposition. Dann gilt $\tau^2 = \text{id}$, das heisst τ ist sein eigenes Inverses.

Beispiel 8.2.5.

- (1) Es sei $n = 4$ und $\sigma = (3, 1, 2, 4)$. Dann ist $\sigma^{-1} = (2, 3, 1, 4)$.
- (2) Es sei $\tau = (4, 1, 2, 3)$. Dann gilt

$$\sigma \circ \tau = (4, 3, 1, 3)$$

$$\tau \circ \sigma = (2, 4, 1, 3).$$

Satz 8.2.6. Jede Permutation kann als die Verknuepfung von endlich vielen Transpositionen geschrieben werden.

Proof. Es sei $\sigma = (\sigma(1), \dots, \sigma(n))$. Wenn $\sigma(1) \neq 1$, dann sei τ_1 die Transposition, die 1 und $\sigma(1)$ vertauscht; ansonsten sei τ_1 die Identitaet. Beachte, dass $\tau_1 \sigma(1) = 1$.

Wenn $\tau_1 \sigma(2) \neq 2$, dann sei τ_2 die Transposition, die 2 und $\tau_1 \sigma(2)$ vertauscht; ansonsten sei τ_2 die Identitaet. Dann gilt $\tau_2 \tau_1 \sigma(1) = 1$ und $\tau_2 \tau_1 \sigma(2) = 2$. Wiederholen Sie diesen Prozess bis τ_n . Dann gilt

$$\tau_n \tau_{n-1} \cdots \tau_1 \sigma = \text{id} \quad \Rightarrow \quad \sigma = \tau_1 \tau_2 \cdots \tau_n,$$

da $\tau_i^2 = \text{id} \forall i$ aufgrund von Beobachtung 8.2.4. □

Beispiele 8.2.7.

- (1) Es sei $\sigma = (3, 4, 1, 2) \in S_4$. Da $\sigma(1) = 4$, nehmen wir $\tau_1 = (3, 2, 1, 4)$. Dann gilt

$$\tau_1 \sigma = (1, 4, 3, 2).$$

Wir nehmen $\tau_2 = (1, 4, 3, 2)$; dann gilt $\tau_2 \tau_1 \sigma = \text{id}$ und daher

$$\sigma = (3, 2, 1, 4)(1, 4, 3, 2).$$

- (2) Es sei $\sigma = (3, 1, 2, 4, 5) \in S_5$. Dann gilt

$$\begin{aligned}\sigma &= (1, 3, 2, 4, 5)(3, 2, 1, 4, 5) \\ &= (1, 2, 3, 5, 4)(1, 3, 2, 4, 5)(3, 2, 1, 4, 5)(1, 2, 3, 5, 4),\end{aligned}$$

die Zerlegung einer Permutation in Transpositionen ist also nicht eindeutig!

Immerhin ist die Parität der Anzahl von Transpositionen eindeutig bestimmt:

Satz 8.2.8. *Es sei $\sigma \in S_n$ und*

$$\sigma = \tau_1 \cdots \tau_k = \tau'_1 \cdots \tau'_m$$

verschiedene Zerlegungen von σ in Produkte von Transpositionen. Dann gilt

$$k \equiv m \pmod{2}.$$

Proof. Später - siehe Korollar 8.3.15. □

Definition 8.2.9. *Eine Permutation $\sigma \in S_n$ heisst gerade (bzw. ungerade), wenn es sich als Produkt einer geraden (bzw. ungeraden) Anzahl von Transpositionen schreiben lässt. Wir definieren das Vorzeichen einer Permutation σ als*

$$\text{sgn}(\sigma) = \begin{cases} 1 & \text{wenn } \sigma \text{ gerade ist} \\ -1 & \text{wenn } \sigma \text{ ungerade ist.} \end{cases}$$

Beispiele 8.2.10.

- (1) Jede Transposition in S_n ist ungerade.
 (2) Das Element $\sigma \in S_4$ aus Beispiel 8.2.7 (1) ist gerade.

Bemerkung 8.2.11. *Die Menge aller geraden Permutationen in S_n ist eine Untergruppe von S_n ; sie heisst die alternierende Untergruppe A_n .*

Die Gruppe A_5 ist die kleinste nicht auflösbare Gruppe: damit werden Sie sich beschäftigen, wenn Sie im nächsten Jahr Galois Theorie lernen.

8.3. Determinantenfunktionen.

Notation 8.3.1. Es sei $A \in M_{n \times n}(K)$; wir schreiben $A = (\mathbf{v}_1, \dots, \mathbf{v}_n)$, wobei $\mathbf{v}_i$ der i te Spaltenvektor ist.

Definition 8.3.2. Eine Funktion $f : M_{n \times n}(K) \rightarrow K$ ist n -linear (in den Spaltenvektoren), wenn $\forall 1 \leq i \leq n$ f eine lineare Funktion der i ten Spalte ist, wenn die anderen Spalten fixiert werden. Mit anderen Worten, f ist n -linear, wenn

$$\begin{aligned} f(\mathbf{v}_1, \dots, \mathbf{v}_i + \mathbf{u}, \dots, \mathbf{v}_n) &= f(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n) + f(\mathbf{v}_1, \dots, \mathbf{u}, \dots, \mathbf{v}_n), \\ f(\mathbf{v}_1, \dots, \lambda \mathbf{v}_i, \dots, \mathbf{v}_n) &= \lambda f(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n) \end{aligned}$$

fuer alle $A = (\mathbf{v}_1, \dots, \mathbf{v}_n) \in M_{n \times n}(K)$, $\mathbf{u} \in K^n$ und $\lambda \in K$.

Beispiele 8.3.3.

- (1) $\det : M_{2 \times 2}(K) \rightarrow K$ ist bilinear.
- (2) Es sei $A = (a_{ij}) \in M_{n \times n}(K)$. Die Funktion

$$f : A \mapsto a_{11}a_{22} \dots a_{nn}$$

ist n -linear.

Lemma 8.3.4. Es seien $f, g : M_{n \times n}(K) \rightarrow K$ n -linear, und es seien $\alpha, \beta \in K$. Dann ist auch $\alpha f + \beta g$ n -linear.

Proof. Uebung. □

Definition 8.3.5. Eine Funktion $f : M_{n \times n}(K) \rightarrow K$ ist alternierend, wenn folgende Eigenschaft erfuehlt ist: wenn $\mathbf{v}_i = \mathbf{v}_{i+1}$ fuer ein $1 \leq i < n$, dann gilt

$$f(\mathbf{v}_1, \dots, \mathbf{v}_n) = 0.$$

Beispiel 8.3.6. $\det : M_{2 \times 2}(K) \rightarrow K$ ist alternierend.

Lemma 8.3.7. Es sei $f : M_{n \times n}(K) \rightarrow K$ n -linear und alternierend. Dann gilt

- (1) $f(\mathbf{v}_1, \dots, \mathbf{v}_i, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n) = -f(\mathbf{v}_1, \dots, \mathbf{v}_{i+1}, \mathbf{v}_i, \dots, \mathbf{v}_n)$;
- (2) wenn $\mathbf{v}_i = \mathbf{v}_j$ fuer $i \neq j$, dann gilt $f(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j, \dots, \mathbf{v}_n) = 0$;
- (3) $f(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j, \dots, \mathbf{v}_n) = -f(\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n)$.

Proof. (1) Es gilt

$$\begin{aligned} f(\mathbf{v}_1, \dots, \mathbf{v}_i + \mathbf{v}_{i+1}, \mathbf{v}_i + \mathbf{v}_{i+1}, \dots, \mathbf{v}_n) &= f(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n) + f(\mathbf{v}_1, \dots, \mathbf{v}_i, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n) \\ &\quad + f(\mathbf{v}_1, \dots, \mathbf{v}_{i+1}, \mathbf{v}_i, \dots, \mathbf{v}_n) + f(\mathbf{v}_1, \dots, \mathbf{v}_{i+1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n) \\ &= f(\mathbf{v}_1, \dots, \mathbf{v}_i, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n) + f(\mathbf{v}_1, \dots, \mathbf{v}_{i+1}, \mathbf{v}_i, \dots, \mathbf{v}_n) \\ &= 0 \end{aligned}$$

und daher

$$f(\mathbf{v}_1, \dots, \mathbf{v}_i, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n) = -f(\mathbf{v}_1, \dots, \mathbf{v}_{i+1}, \mathbf{v}_i, \dots, \mathbf{v}_n).$$

- (2) Nimm nun an, dass $A = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ eine Matrix ist, fuer die Spalten $\mathbf{v}_i$ und $\mathbf{v}_j$ gleich sind. Wir koennen dann nebeneinanderliegende Spalten so lange vertauschen, bis wir eine Matrix A' erhalten, die zwei gleiche nebeneinanderliegende Spalten hat; aufgrund von (1) wissen wir, dass sich $f(A)$ und $f(A')$ hoechstens durch das Vorzeichen unterscheiden. Aber da f alternierend ist, gilt $f(A') = 0$ und daher $f(A) = 0$.
- (3) Der Beweis ist aehnlich wie (1).

□

Definition 8.3.8. Eine Funktion $D : M_{n \times n}(K) \rightarrow K$ ist eine Determinantenfunktion, wenn sie folgende Eigenschaften hat:

- (1) sie ist n -linear;
- (2) sie ist alternierend;
- (3) $D(\mathbf{1}_n) = 1$.

Frage: Gibt es Determinantenfunktionen ueberhaupt?

Lemma 8.3.9. Die Determinante fuer (2×2) -Matrizen

$$\det : M_{2 \times 2}(K) \rightarrow K$$

ist eine Determinantenfunktion, und sie ist die einzige Determinantenfunktion auf $M_{2 \times 2}(K)$.

Proof. Die Axiome der Determinantenfunktion sind leicht zu ueberpruefen; wir zeigen die Eindeutigkeit. Es sei $D : M_{2 \times 2}(K) \rightarrow K$ eine Determinantenfunktion, und es sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Wir wollen zeigen, dass $D(A) = ad - bc$. Es seien $\mathbf{e}_1, \mathbf{e}_2$ die Standardbasis von K^2 , und wir schreiben $A = (a\mathbf{e}_1 + c\mathbf{e}_2, b\mathbf{e}_1 + d\mathbf{e}_2)$. Dann gilt aufgrund der Axiome

$$\begin{aligned} D(A) &= D(a\mathbf{e}_1 + c\mathbf{e}_2, b\mathbf{e}_1 + d\mathbf{e}_2) \\ &= ab \cdot D(\mathbf{e}_1, \mathbf{e}_1) + ac \cdot D(\mathbf{e}_1, \mathbf{e}_2) + bc \cdot D(\mathbf{e}_2, \mathbf{e}_1) + cd \cdot D(\mathbf{e}_2, \mathbf{e}_2) \\ &= ac \cdot D(\mathbf{e}_1, \mathbf{e}_2) + bc \cdot D(\mathbf{e}_2, \mathbf{e}_1) \\ &= (ad - bc) \cdot D(\mathbf{e}_1, \mathbf{e}_2) \\ &= (ad - bc) \cdot D(I_2) \\ &= ad - bc. \end{aligned}$$

□

Wir wollen nun folgendes zeigen:

- (1) $\forall n \geq 1$ gibt es eine Determinantenfunktion.
- (2) $\forall n \geq 1$ ist die Determinantenfunktion eindeutig bestimmt; das heisst, es gibt *genau eine*.

Wir zeigen zunaechst die Existenz:

Theorem 8.3.10. Es sei $n \geq 1$. Dann gibt es eine Determinantenfunktion $D : M_{n \times n}(K) \rightarrow K$.

Wir brauchen fuer den Beweis folgendes Konzept:

Notation 8.3.11. Sei $A = (a_{ij})_{1 \leq i, j \leq n} \in M_{n \times n}(K)$. Fuer $1 \leq i, j \leq n$ sei $A_{i,j}$ die $(n-1, n-1)$ -Matrix, die wir von A erhalten, indem wir die i te Zeile und j te Spalte herausnehmen.

Beispiele 8.3.12.

- (1) Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Dann ist

$$A_{1,1} = d, \quad A_{1,2} = c, \quad A_{2,1} = b, \quad A_{2,2} = a.$$

- (2) Sei $B = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 0 \\ 0 & 1 & -2 \end{pmatrix}$. Was sind $B_{3,1}$ und $B_{2,3}$?

Um Theorem 8.3.10 zu beweisen, brauchen wir folgenden Satz:

Satz 8.3.13. Es sei $n \geq 2$ und f eine $(n-1)$ -lineare alternierende Funktion. Fuer $1 \leq i \leq n$ definiere die Funktion

$$E_i : M_{n \times n}(K) \rightarrow K, \quad A = (a_{ij}) \mapsto \sum_{j=1}^n (-1)^{i+j} a_{ij} f(A_{i,j}).$$

Dann ist E_i n -linear und alternierend.

Proof. Wir zeigen zunaechst, dass E_i n -linear ist. Es sei $1 \leq j \leq n$, und betrachte die Funktion

$$e_{ij} : A \mapsto a_{ij} f(A_{i,j});$$

wir muessen zeigen, dass sie n -linear ist: dann folgt von Lemma 8.3.4, dass E_i ebenfalls n -linear ist. Betrachte e_{ij} als eine Funktion der k ten Spalte.

- wenn $k \neq j$, dann ist a_{ij} unabhaengig von k , und $f(A_{i,j})$ ist linear in der k ten Spalte, und daher ist e_{ij} linear als eine Funktion der k ten Spalte.
- wenn $k = j$, dann ist $f(A_{i,j})$ unabhaengig von k , und $a_{ij} = a_{ik}$ ist linear in der k ten Spalte, und daher ist e_{ij} linear als eine Funktion der k ten Spalte.

Es folgt daher von Lemma 8.3.4, dass E_i n -linear ist.

Nimm nun an, dass $\mathbf{v}_k = \mathbf{v}_{k+1}$ fuer $1 \leq k < n$. Dann hat fuer all $j \neq k, k+1$ die Matrix $A_{i,j}$ zwei gleiche Spalten, und daher gilt $f(A_{i,j}) = 0$, da f alternierend ist, das heisst

$$E_i(A) = (-1)^{i+k} a_{ik} f(A_{i,k}) + (-1)^{i+k+1} a_{i,k+1} f(A_{i,k+1}).$$

Aber $A_{i,k} = A_{i,k+1}$ und $a_{ik} = a_{i,k+1}$, und daher $E_i(A) = 0$. □

Wir koennen nun Theorem 8.3.10 beweisen.

Proof. Wir zeigen die Existenz durch Induktion nach n . Fuer $n = 1$ sei D die Identitaet, und fuer $n = 2$ die Determinante.

Nimm nun an, dass wir $D_{n-1} : M_{n-1,n-1}(K) \rightarrow K$ definiert haben. Es sei $A = (a_{ij})_{1 \leq i,j \leq n} \in M_{n \times n}(K)$. Waehle nun ein $1 \leq i \leq n$. Definiere

$$D_n^{(i)}(A) = (-1)^{i+1} a_{i1} D_{n-1}(A_{i,1}) + \cdots + (-1)^{i+n} a_{in} D_{n-1}(A_{i,n}).$$

Wir zeigen, dass $D_n^{(i)}$ eine Determinantenfunktion ist.⁴⁰ Es folgt von Satz 8.3.10, dass D_n n -linear und alternierend ist.

Wir muessen nun zeigen, dass $D_n^{(i)}(\mathbf{1}_n) = 1$. Beachte, dass

$$(\mathbf{1}_n)_{ij} = \begin{cases} 0 & \text{wenn } i \neq j \\ 1 & \text{wenn } i = j. \end{cases}$$

Es folgt daher, dass

$$D_n^{(i)}(\mathbf{1}_n) = (-1)^{i+i} D_{n-1}(\mathbf{1}_{n-1}) = 1. \quad \square$$

Beispiel 8.3.14. Fuer $n = 3$ erhalten wir drei Determinantenfunktionen. Es sei $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$.

Dann sind die drei Determinantenfunktionen gegeben durch

$$\begin{aligned} D_3^{(1)} : A &\mapsto a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}, \\ D_3^{(2)} : A &\mapsto -a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{22} \det \begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix} - a_{23} \det \begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix}, \\ D_3^{(3)} : A &\mapsto a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} - a_{32} \det \begin{pmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{pmatrix} + a_{33} \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \end{aligned}$$

As Korollar erhalten wir einen sehr eleganten Beweis von Satz 8.2.8:

Korollar 8.3.15. Es sei $\sigma \in S_n$ und

$$\sigma = \tau_1 \cdots \tau_k = \tau'_1 \cdots \tau'_m$$

verschiedene Zerlegungen von σ in Produkte von Transpositionen. Dann gilt

$$k \equiv m \pmod{2}.$$

Proof. Es sei $D : M_{n \times n}(K) \rightarrow K$ eine Determinantenfunktion, und es sei $\mathbf{e}_1, \dots, \mathbf{e}_n$ die Standardbasis von K^n . Wenn σ durch die Verknuepfung von k Transpositionen erhalten werden kann, dann an die Einheitsmatrix $\mathbf{1}_n$ von der Matrix $(\mathbf{e}_{\sigma(1)}, \dots, \mathbf{e}_{\sigma(n)})$ durch k Vertauschungen von Spalten konstruiert werden. Da das Vertauschen von Spalten das Vorzeichen des Wertes der Determinantenfunktion aendert, erhalten wir

$$D(\mathbf{e}_{\sigma(1)}, \dots, \mathbf{e}_{\sigma(n)}) = (-1)^k D(\mathbf{1}_n) = (-1)^k.$$

Durch das gleiche Argument erhalten wir

$$D(\mathbf{e}_{\sigma(1)}, \dots, \mathbf{e}_{\sigma(n)}) = (-1)^m,$$

und daher gilt $k \equiv m \pmod{2}$. □

Um die Eindeutigkeit der Determinantenfunktion zu zeigen, benutzen wir die gleiche Strategie wie in Lemma 8.3.9 im Fall $n = 2$.

⁴⁰Dieses gilt fuer jedes $1 \leq i \leq n$, d.h. wir bekommen i Determinantenfunktionen! Wir werden spaeter sehen, dass sie alle gleich sind.

Theorem 8.3.16. Es sei $n \geq 1$. Dann gibt es genau eine⁴¹ Determinantenfunktion

$$D : M_{n \times n}(K) \rightarrow K.$$

Explizit ist D durch folgende Formel gegeben: es sei $A = (a_{ij})_{1 \leq i, j \leq n}$. Dann gilt

$$(38) \quad D(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n}.$$

Proof. Es sei $\mathbf{e}_1, \dots, \mathbf{e}_n$ die Standardbasis von K^n ; fuer $A = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ schreiben wir

$$\mathbf{v}_i = a_{1i}\mathbf{e}_1 + \cdots + a_{ni}\mathbf{e}_n.$$

Es sei nun $D : M_{n \times n}(K) \rightarrow K$ eine Determinantenfunktion.⁴² Da D alternierend und n -linear ist, erhalten wir durch Expansion

$$\begin{aligned} D(A) &= D(a_{11}\mathbf{e}_1 + \cdots + a_{n1}\mathbf{e}_n, \dots, a_{1n}\mathbf{e}_1 + \cdots + a_{nn}\mathbf{e}_n) \\ &= \sum_{\sigma \in S_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} \cdot D(\mathbf{e}_{\sigma(1)}, \dots, \mathbf{e}_{\sigma(n)}). \end{aligned}$$

Aber nun gilt

$$D(\mathbf{e}_{\sigma(1)}, \dots, \mathbf{e}_{\sigma(n)}) = \operatorname{sgn}(\sigma) \cdot D(\mathbf{e}_1, \dots, \mathbf{e}_n) = \operatorname{sgn}(\sigma),$$

und das Resultat folgt. □

Bemerkung 8.3.17. Sie fragen sich vielleicht, warum wir die Determinantenfunktion nicht einfach durch die Formel 8.3.16 definiert haben. Im Prinzip ist das moeglich, allerdings haetten wir dann auf anderem Weg zeigen muessen, dass das Vorzeichen einer Permutation wohldefiniert ist.

Lecture 26

Bemerkung 8.3.18. In the Praxis is die Formel (38) nicht wirklich nuetzlich: wie wollen Sie fuer $n \geq 4$ systematisch alle Elemente von S_n auflisten? Fuer explizite Rechnungen empfiehlt sich eher die Zeilenentwicklung der Determinante aus dem Beweis von Theorem 8.3.10: es sei $n \geq 1$ und $1 \leq i \leq n$. Dann gilt fuer $A \in M_{n \times n}(K)$

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}).$$

Insbesondere folgt aus dem Beweis von Theorem 8.3.16, dass folgendes gilt:

Korollar 8.3.19. Es sei $D : M_{n \times n}(K) \rightarrow K$ n -linear und alternierend. Dann gilt

$$D(A) = \det(A) \cdot D(\mathbf{1}_n)$$

fuer jedes $A \in M_{n \times n}(K)$.

Dieses Wissen, dass jede n -lineare und alternierende Funktion ein Vielfaches der Determinante ist, wird sehr nuetzlich sein, wenn wir im naechsten Kapitel Eigenschaften der Determinantenfunktion untersuchen, z.B. die Multiplikativitaet (Satz 8.5.2).

⁴¹Insbesondere sind die drei Formeln in Beispiel 8.3.14 gleich. Rechnen Sie nach!

⁴²Wir wissen dank Theorem 8.3.10, dass eine solche Funktion existiert.

8.4. Erste Eigenschaften. Eine Eigenschaft der Determinante folgt unmittelbar von (38):

Satz 8.4.1. *Es sei $A \in M_{n \times n}(K)$. Dann gilt*

$$\det(A) = \det(A^t).$$

Proof. Es sei $B = (b_{ij}) = A^t$, das heisst $b_{ij} = a_{ji}$. Dann gilt

$$\begin{aligned} \det(B) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) b_{\sigma(1),1} b_{\sigma(2),2} \cdots b_{\sigma(n),n} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}. \end{aligned}$$

Doch wenn $j = \sigma(i)$, dann ist $i = \sigma^{-1}(j)$, und daher $a_{i,\sigma(i)} = a_{\sigma^{-1}(j),j}$. Es folgt, dass

$$\det(B) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma^{-1}(1),1} a_{\sigma^{-1}(2),2} \cdots a_{\sigma^{-1}(n),n}.$$

Doch $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$ (warum?), und daher gilt

$$\det(B) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) a_{\sigma^{-1}(1),1} a_{\sigma^{-1}(2),2} \cdots a_{\sigma^{-1}(n),n} = \det(A).$$

□

Da die Transposition Spalten und Zeilen einer Matrix vertauscht, erhalten wir folgendes sehr ueber-raschendes Korollar:

Korollar 8.4.2. *Die Determinante ist ebenfalls n -linear und alternierend als eine Funktion in the Zeilen-vektoren.*

Bemerkung 8.4.3. *Daher gilt auch das analoge Resultat von Lemma 8.3.7 fuer jede Funktion $M_{n \times n}(K) \rightarrow K$, die n -linear und alternierend in den Zeilenvektoren ist. Weiterhin ist jede n -lineare und alternierende Funktion in den Zeilenvektoren ein Vielfaches von der Determinante.*

Wir verstehen nun, wie sich die Determinante unter den elementaren Zeilenumformungen verhaelt:

Satz 8.4.4. *Es sei $A \in M_{n \times n}(K)$, und es sei B eine Matrix, die wir von A durch die elementare Zeilenumformung X erhalten.*

- (1) *wenn $X = P(r, s)$ fuer $1 \leq r < s \leq n$, dann gilt $\det(B) = -\det(A)$;*
- (2) *wenn $X = M(r, \lambda)$ fuer $1 \leq r \leq n$ und $\lambda \in K^\times$, dann gilt $\det(B) = \lambda \det(A)$;*
- (3) *wenn $X = S(r, s, \lambda)$ fuer $1 \leq r, s \leq n$, $r \neq s$ und $\lambda \in K^\times$, dann gilt $\det(B) = \det(A)$.*

Proof. Uebung.

□

Dieses Resultat ist sehr nuetzlich, um Determinanten in konkreten Beispielen zu berechnen.

Beispiel 8.4.5.

- (1) Es sei $A = \begin{pmatrix} 0 & -9 & 3 \\ 1 & 4 & 0 \\ 2 & 6 & -1 \end{pmatrix}$. Wir hatten bereits in Beispiel 2.3.10 gesehen, wie wir diese Matrix

durch EZUs in die Einheitsmatrix umwandeln:

$$P(1, 2) \circ S(3, 1, \frac{1}{3}) \circ S(3, 2, -\frac{4}{3}) \circ M(3, \frac{-3}{5}) \circ S(2, 3, -2) \circ S(1, 3, -6) \circ S(1, 2, -4) \circ M(1, -\frac{1}{9}).$$

Dies bedeutet, dass

$$\det(\mathbf{1}_3) = (-1) \cdot \left(-\frac{3}{5}\right) \cdot \left(-\frac{1}{9}\right) \cdot \det(A) = -\frac{1}{15},$$

das heisst $\det(A) = -15$. Wir ueberpruefen diese Rechnung:

$$\begin{aligned} \det(A) &= \det \begin{pmatrix} 0 & -9 & 3 \\ 1 & 4 & 0 \\ 2 & 6 & -1 \end{pmatrix} \\ &= 0 \cdot \det \begin{pmatrix} 4 & 0 \\ 6 & -1 \end{pmatrix} - (-9) \cdot \det \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix} + 3 \cdot \det \begin{pmatrix} 1 & 4 \\ 2 & 6 \end{pmatrix} \\ &= -9 - 6 \\ &= -15. \end{aligned}$$

(2) Es sei $B = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 4 & -3 \\ 3 & 6 & -5 \end{pmatrix}$. Dann erhalten wir die Einheitsmatrix durch folgende EZUs:

$S(2, 3, 7/2) \circ S(1, 3, -11/2) \circ S(1, 2, -2) \circ M(3, -2) \circ S(3, 2, -3) \circ M(2, 1/2) \circ S(2, 1, -2)$,
und daher

$$\det(\mathbf{1}_3) = (-2) \cdot \frac{1}{2} \cdot \det(B) = -1,$$

was wir durch die Determinantenentwicklung ueberpruefen koennen.

Wir koennen nun folgendes Resultat beweisen:

Theorem 8.4.6. *Es sei $M \in M_{n \times n}(K)$ von der Form*

$$M = \begin{pmatrix} A & B \\ 0_{s \times r} & C \end{pmatrix} \quad \text{fuer } A \in M_{r \times r}(K), B \in M_{r \times s}(K) \text{ und } C \in M_{s \times s}(K).$$

Dann gilt

$$\det(M) = \det(A) \det(C).$$

Proof. Wir benutzen Korollar 8.3.19. Es seien $A \in M_{r \times r}(K)$ und $B \in M_{r \times s}(K)$. Definiere die Funktion

$$D : M_{s \times s}(K) \longrightarrow K, \\ C \mapsto \det \begin{pmatrix} A & B \\ 0_{s \times r} & C \end{pmatrix}.$$

Dann ist D s -linear und alternierend (in den Zeilen von C), und daher gilt

$$D(C) = \det(C) \cdot D(\mathbf{1}_s).$$

Per Definition gilt $D(\mathbf{1}_s) = \det \begin{pmatrix} A & B \\ 0_{s \times r} & \mathbf{1}_s \end{pmatrix}$. Durch elementare Zeilenumformungen $S(i, j, \lambda)$ koennen

wir die Matrix $\begin{pmatrix} A & B \\ 0_{s \times r} & \mathbf{1}_s \end{pmatrix}$ in die Matrix $\begin{pmatrix} A & 0_{r \times s} \\ 0_{s \times r} & \mathbf{1}_s \end{pmatrix}$ umformen, und es folgt von Satz 8.4.4, dass

$$\det \begin{pmatrix} A & B \\ 0_{s \times r} & \mathbf{1}_s \end{pmatrix} = \det \begin{pmatrix} A & 0_{r \times s} \\ 0_{s \times r} & \mathbf{1}_s \end{pmatrix}.$$

Betrachte nun die Funktion

$$\Delta : M_{r \times r}(K) \longrightarrow K, \\ A \mapsto \det \begin{pmatrix} A & 0_{r \times s} \\ 0_{s \times r} & \mathbf{1}_s \end{pmatrix}.$$

Dann ist Δ r -linear und alternierend (in den Zeilen oder Spalten von A), und daher gilt

$$\Delta(A) = \det(A) \cdot \Delta(\mathbf{1}_r) = \det(A) \cdot \det(\mathbf{1}_{r+s}) = \det(A).$$

Es folgt, dass

$$\begin{aligned} \det \begin{pmatrix} A & B \\ 0_{s \times r} & C \end{pmatrix} &= \det(C) \cdot D(\mathbf{1}_s) \\ &= \det(C) \cdot \det \begin{pmatrix} A & 0_{r \times s} \\ 0_{s \times r} & \mathbf{1}_s \end{pmatrix} \\ &= \det(C) \Delta(A) \\ &= \det(C) \det(A). \end{aligned}$$

□

Wir koennen nun ohne viel Aufwand folgendes Korollar beweisen; richtige Loesungen habe ich von Jiarui Kang, Teniver Birkhauser, Jakob Göhring und Marco Vaccaro erhalten.

Korollar 1. *Es sei $M \in M_{n \times n}(K)$ von der Form*

$$M = \begin{pmatrix} 0_{r \times s} & C \\ A & B \end{pmatrix} \quad \text{fuer } A \in M_{s \times s}(K), B \in M_{s \times r}(K) \text{ und } C \in M_{r \times r}(K).$$

Dann gilt

$$\det(M) = (-1)^{rs} \det(A) \det(C).$$

Proof. Wir zeigen, dass sich die Matrix mit $r(s+r-1)$ Zeilenvertauschungen in die Form

$$M' = \begin{pmatrix} A & B \\ 0_{r \times s} & C \end{pmatrix}$$

bringen lässt. Da jede Vertauschung das Vorzeichen ändert und

$$r(s+r-1) \equiv rs \pmod{2},$$

folgt das Resultat von Theorem 8.4.6.

Es sei U die Verknüpfung der $r+s-1$ EZUs

$$U = P(1, 2) \circ P(2, 3) \circ \cdots \circ P(s+r-1, s+r).$$

Wenn wir dann die Operation U r -mal auf M anwenden, erhalten wir M' . □

Bemerkung 8.4.7. Es ist nicht richtig, dass wir einfach die Transposition auf M anwenden können, um eine Matrix der Blockform $\begin{pmatrix} \star & \star & \star \\ 0 & \star \end{pmatrix}$ zu erhalten! Es gilt

$$M^t = \begin{pmatrix} 0_{s \times r} & B^t \\ C^t & t \end{pmatrix} :$$

die Transposition ist die Spiegelung in der Hauptdiagonalen.

Beispiele 8.4.8. (1) Es sei $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 2 & 1 \\ 0 & 3 & -2 \end{pmatrix}$. Dann gilt

$$\begin{aligned} \det(A) &= 1 \cdot \det \begin{pmatrix} 2 & 1 \\ 3 & -2 \end{pmatrix} \\ &= -7. \end{aligned}$$

(2) Es sei $B = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 2 & -1 & 1 & 1 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$. Dann gilt

$$\begin{aligned} \det(B) &= \det \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \cdot \det \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix} \\ &= (-5) \cdot 4 \\ &= 20. \end{aligned}$$

(3) Es sei $C = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 5 \\ 0 & 0 & 2 \end{pmatrix}$. Dann gilt

$$\begin{aligned} \det(C) &= 1 \cdot (-1) \cdot 2 \\ &= -2. \end{aligned}$$

Das letzte Beispiel lässt sich folgendermassen verallgemeinern:

Korollar 8.4.9. Es sei $A = (a_{ij}) \in M_{n \times n}(K)$ eine obere Dreiecksmatrix. Dann gilt

$$\det(A) = a_{11} \cdots a_{nn}.$$

8.5. Determinanten und Invertierbarkeit.

Beachte 8.5.1. Es sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2 \times 2}(K)$. Dann ist

$$\det(A) = ad - bc,$$

und A ist genau dann invertierbar, wenn $\det(A) \neq 0$; in diesem Fall gilt

$$A^{-1} = (ad - bc)^{-1} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

und eine leichte Rechnung zeigt, dass $\det(A^{-1}) = \det(A)^{-1}$.

Frage. Lassen sich diese Resultate auf $(n \times n)$ -Matrizen verallgemeinern?

Satz 8.5.2. Es sei $n \geq 1$ und $A, B \in M_{n \times n}(K)$. Dann gilt

$$\det(AB) = \det(A) \det(B);$$

mit anderen Worten, die Determinante ist multiplikativ.

Proof. Wir benutzen wieder Korollar 8.3.19. Definiere die Funktion

$$\begin{aligned} D : M_{n \times n}(K) &\longrightarrow K, \\ B &\mapsto \det(AB). \end{aligned}$$

Behauptung. Die Funktion D ist n -linear und alternierend als eine Funktion in den Spalten.⁴³

Beweis der Behauptung. Schreibe $A = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ und $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$. Dann sind die Spalten der Matrix AB gegeben durch

$$AB = (A\mathbf{b}_1, \dots, A\mathbf{b}_n).$$

Daher gilt

$$\begin{aligned} D(\mathbf{b}_1, \dots, \mathbf{b}_i + \mathbf{u}, \dots, \mathbf{b}_n) &= \det(A\mathbf{b}_1, \dots, A(\mathbf{b}_i + \mathbf{u}), \dots, A\mathbf{b}_n) \\ &= \det(A\mathbf{b}_1, \dots, A\mathbf{b}_i + A\mathbf{u}, \dots, A\mathbf{b}_n) \\ &= \det(A\mathbf{b}_1, \dots, A\mathbf{b}_i, \dots, A\mathbf{b}_n) + \det(A\mathbf{b}_1, \dots, A\mathbf{u}, \dots, A\mathbf{b}_n) \\ &= D(\mathbf{b}_1, \dots, \mathbf{b}_i, \dots, \mathbf{b}_n) + D(\mathbf{b}_1, \dots, \mathbf{u}, \dots, \mathbf{b}_n), \end{aligned}$$

wobei die dritte Gleichheit von der n -Linearität von $\det$ folgt. Daher ist D n -linear.

Nimm nun an, dass $\mathbf{b}_i = \mathbf{b}_{i+1}$ fuer ein $1 \leq i < n$. Dann ist $A\mathbf{b}_i = A\mathbf{b}_{i+1}$, und daher

$$D(\mathbf{b}_1, \dots, \mathbf{b}_n) = \det(A\mathbf{b}_1, \dots, A\mathbf{b}_n) = 0,$$

da $\det$ alternierend ist. Das beweist die Behauptung.

Wir wissen daher von Korollar 8.3.19, dass

$$D(B) = \det(B) \cdot D(\mathbf{1}_n) \quad \forall B \in M_{n \times n}(K).$$

Doch $D(\mathbf{1}_n) = \det(A)$, und daher

$$\det(AB) = \det(A) \cdot \det(B).$$

□

Korollar 8.5.3. Wenn $A \in M_{n \times n}(K)$ invertierbar ist, dann ist $\det(A) \neq 0$, und

$$\det(A^{-1}) = \det(A)^{-1}.$$

Proof. Wenn A invertierbar ist, dann gibt es eine Matrix $A^{-1} \in M_{n \times n}(K)$, so dass $AA^{-1} = \mathbf{1}_n$. Dann gilt

$$\det(A) \det(A^{-1}) = \det(AA^{-1}) = \det(\mathbf{1}_n) = 1.$$

□

Bemerkung 8.5.4. Korollar 8.5.3 wirft zwei neue Fragen auf:

- (1) Gilt auch die Umkehrung, d.h. wenn $\det(A) \neq 0$, folgt es dann, dass A invertierbar ist?
- (2) Wenn A invertierbar ist, wie findet man A^{-1} ?

Wir werden uns gleich mit diesen Fragen beschaeftigen. Zunaechst ernten wir aber ein paar Konsequenzen des Korollars.

Korollar 8.5.5. Aehnliche Matrizen haben die gleiche Determinante.

Proof. Es seien $A, B, C \in M_{n \times n}(K)$ mit C invertierbar, so dass $B = C^{-1}AC$. Dann gilt

$$\det(B) = \det(C^{-1}AC) = \det(C)^{-1} \det(A) \det(C) = \det(A).$$

□

Korollar 8.5.6. Nimm an, dass K ein unendlicher Koerper ist. Dann gibt es unendlich viele Aequivalenzklassen aehnlicher Matrizen.⁴⁴

⁴³Sie koennen alternativ auch die Funktion $A \mapsto \det(AB)$ betrachten; diese ist dann n -linear und alternierend in den Zeilen.

⁴⁴Vergl. mit Korollar 4.7.4.

Proof. Übung. □

Bemerkung 8.5.7. *Korollar 8.5.5 ist aus zweierlei Gruenden sehr wichtig: erstens, weil es die Berechnung von Determinanten viel leichter macht (wie wir gleich sehen werden), und zweitens, da es uns erlaubt, die Determinante einer linearen Abbildung $V \rightarrow V$ zu definieren. Wir kommen im naechsten Kapitel darauf zurueck.*

Beispiele 8.5.8. (1) Es sei $A = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$ und $B = \begin{pmatrix} -3 & -10 \\ 2 & 5 \end{pmatrix}$. Dann gilt

$$B = C^{-1}AC \quad \text{mit } C = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

und daher $\det(A) = \det(B)$, was wir durch explizite Rechnung ueberpruefen.

(2) Es sei $A = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix}$. Dann gilt

$$A = C^{-1} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \cdot C \quad \text{mit } C = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

und daher

$$\det(A) = \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} = 4.$$

Um die Umkehrung von Korollar 8.5.3 zu beweisen, erinnern wir uns an die explizite Formel der Determinante im Beweis von Satz 8.3.13: Es sei $A = (a_{ij})_{1 \leq i, j \leq n} \in M_{n \times n}(K)$, und es sei $1 \leq i \leq n$. Dann gilt

$$(39) \quad \det(A) = (-1)^{i+1}a_{i1} \det(A_{i,1}) + \cdots + (-1)^{i+n}a_{in} \det(A_{i,n}),$$

wobei $A_{i,k}$ die $(n-1) \times (n-1)$ -Matrix ist, die wir von A erhalten, indem wir die i te Zeile und die k te Spalte entfernen.

Definition 8.5.9. *Fuer $1 \leq i, j \leq n$ schreiben wir*

$$(40) \quad c_{i,j} = (-1)^{i+j} \det(A_{i,j}).$$

Die Kofaktormatrix von A ist die Matrix $C = (c_{i,j}) \in M_{n \times n}(K)$. Die adjunkte Matrix $\text{adj}(A)$ von A ist die Transposition der Kofaktormatrix.⁴⁵

$$\text{adj}(A) = C^t.$$

Beispiele 8.5.10. (1) Es sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Dann ist

$$\text{adj}(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

(2) Es sei $B = \begin{pmatrix} 4 & 5 & 2 \\ -1 & 0 & 3 \\ 1 & 1 & 2 \end{pmatrix}$. Dann gilt

$$\text{adj}(B) = \begin{pmatrix} -3 & -8 & 15 \\ 5 & 6 & -14 \\ -1 & 1 & 5 \end{pmatrix}.$$

Bemerkung 8.5.11. *Von Lemma 8.3.9 wissen wir, dass*

$$A \cdot \text{adj}(A) = \text{adj}(A) \cdot A = \det(A) \cdot \mathbf{1}_2.$$

Berechnen Sie $B \cdot \text{adj}(B)$, $\text{adj}(B) \cdot B$ und $\det(B)$. Was faellt Ihnen auf?

⁴⁵Die adjunkte Matrix einer (1×1) -Matrix ist $(1)!$

Vermutung 8.5.12. Es sei $A = (a_{ij}) \in M_{n \times n}(K)$ und $\text{adj}(A) = (b_{ij})$ die adjunkte Matrix; per Definition gilt $b_{ij} = c_{ji}$, wobei $c_{i,j}$ durch (40) gegeben ist. Dann gilt

$$(41) \quad A \cdot \text{adj}(A) = \text{adj}(A) \cdot A = \det(A) \cdot \mathbf{1}_n.$$

Mit anderen Worten, wenn $\det(A) \neq 0$, dann ist A invertierbar, und es gilt

$$A^{-1} = \det(A)^{-1} \cdot \text{adj}(A).$$

Um diese Vermutung zu beweisen, muessen wir das Produkt $A \cdot \text{adj}(A)$ ausrechnen. Schreibe

$$M = (m_{ik}) = A \cdot \text{adj}(A).$$

Wir muessen zeigen, dass

$$m_{ik} = \begin{cases} \det(A) & \text{wenn } i = k \\ 0 & \text{wenn } i \neq k \end{cases}$$

Bemerkung 8.5.13. Es gilt

$$\begin{aligned} m_{ik} &= \sum_{j=1}^n a_{ij} b_{jk} \\ &= \sum_{j=1}^n a_{ij} c_{kj}. \end{aligned}$$

Lemma 8.5.14. Es sei $1 \leq i \leq n$. Dann gilt

$$m_{ii} = a_{i1}c_{i,1} + \cdots + a_{in}c_{i,n} = \det(A).$$

Proof. Das ist Formel (39). □

Lemma 8.5.15. Es seien $1 \leq i, k \leq n$ und $i \neq k$. Dann gilt

$$m_{ki} = a_{k1}c_{i,1} + \cdots + a_{kn}c_{i,n} = 0.$$

Proof. Die Idee ist es, A durch eine Matrix A' zu ersetzen, so dass $m'_{ii} = m_{ik}$ und $\det(A') = 0$; das Resultat folgt dann unmittelbar von Lemma 8.5.14.

Es sei A' die Matrix, die wir von A erhalten, indem wir die i te Zeile durch die k te Zeile ersetzen, und es sei C' die Kofaktormatrix von A' . Dann hat A' zwei gleiche Zeilen, und daher gilt $\det(A') = 0$. Andererseits koennen wir Lemma 8.5.14 auf A' anwenden und erhalten

$$(42) \quad \det(A') = 0 = a'_{i1}c'_{i,1} + \cdots + a'_{in}c'_{i,n}.$$

Nun folgt von der Definition von A' , dass $a'_{ij} = a_{kj}$ fuer alle $1 \leq j \leq n$. Ausserdem gilt

$$c'_{i,j} = \det(A'_{ij}) = \det(A_{ij}) = c_{i,j}$$

fuer alle $1 \leq j \leq n$, da A' und A sich ja nur in der i ten Zeile unterscheiden. Daher koennen wir (42) schreiben als

$$\det(A') = 0 = a_{k1}c_{i,1} + \cdots + a_{kn}c_{i,n},$$

was zu beweisen war. □

Wir koennen Lemmata 8.5.14 und 8.5.15 folgendermassen zusammenfassen:

Satz 8.5.16. Es gilt

$$A \cdot \text{adj}(A) = \det(A) \cdot \mathbf{1}_n.$$

Proof. Es seien $1 \leq i, k \leq n$. Dann ist der Eintrag der Matrix $A \cdot \text{adj}(A)$ an der Stelle (i, k) gegeben durch

$$a_{k1}c_{i,1} + \cdots + a_{kn}c_{i,n} = 0.$$

□

Um Vermutung 8.5.12 zu beweisen, muessen wir noch zeigen, dass

$$\text{adj}(A) \cdot A = \det(A) \cdot \mathbf{1}_n.$$

Lemma 8.5.17. Es gilt

$$\text{adj}(A) \cdot A = \det(A) \mathbf{1}_n.$$

Proof. Wir wenden Theorem 8.5.16 auf die transponierte Matrix A^t an:

$$A^t \cdot \text{adj}(A^t) = \det(A^t) \mathbf{1}_n.$$

Aber $\det(A^t) = \det(A)$ aufgrund von Satz 8.4.1, und daher

$$A^t \cdot \text{adj}(A^t) = \det(A) \mathbf{1}_n.$$

Wir wenden nun die Transposition auf diese Gleichung an und erhalten

$$(\text{adj}(A^t))^t \cdot A = \det(A) \cdot \mathbf{1}_n.$$

Aber eine einfache Rechnung zeigt, dass $\text{adj}(A^t) = (\text{adj}(A))^t$. Das Lemma folgt. $\square$

Damit haben wir die Vermutung bewiesen:

Theorem 8.5.18. *Eine Matrix $A \in M_{n \times n}(K)$ ist genau dann invertierbar, wenn $\det(A) \neq 0$. In diesem Fall ist die inverse Matrix gegeben durch*

$$A^{-1} = \det(A)^{-1} \cdot \text{adj}(A).$$

Proof. Folgt unmittelbar von Lemma 8.5.17. $\square$

Bemerkung 8.5.19. *Nimm an, wir haben ein lineares Gleichungssystem*

$$S : A \cdot x = b,$$

wobei A eine invertierbare $(n \times n)$ -Matrix ist. In diesem Fall ist die eindeutige Lösung $L(S)$ gegeben durch

$$A^{-1} \cdot b = \det(A)^{-1} \cdot \text{adj}(A) \cdot b.$$

In der Praxis ist es aber sehr mühsam, die adjunkte Matrix von A zu berechnen; es ist meistens wesentlich einfacher, das lineare Gleichungssystem durch Reduktion in reduzierte Zeilenstufenform umzuwandeln.

8.6. Die Determinante eines Endomorphismus. Es sei V ein endlich-dimensionaler Vektorraum, und es sei $T : V \rightarrow V$ eine lineare Abbildung.

Definition 8.6.1. *Es sei $\mathcal{B}$ eine Basis von V . Wir definieren*

$$\det(T) = \det[T]_{\mathcal{B}}^{\mathcal{B}}.$$

Wir müssen überprüfen, dass diese Definition sinnvoll ist, d.h. dass sie unabhängig von der Wahl der Basis ist.

Lemma 8.6.2. *Es sei $\mathcal{C}$ eine weitere Basis von V . Dann gilt*

$$\det[T]_{\mathcal{B}}^{\mathcal{B}} = \det[T]_{\mathcal{C}}^{\mathcal{C}};$$

mit anderen Worten, die Determinante eines Endomorphismus ist wohl-definiert.

Proof. Von Theorem 4.5.5 wissen wir, dass

$$[T]_{\mathcal{C}}^{\mathcal{C}} = [\text{id}]_{\mathcal{C}}^{\mathcal{B}} \cdot [T]_{\mathcal{B}}^{\mathcal{B}} \cdot [\text{id}_V]_{\mathcal{B}}^{\mathcal{C}}.$$

Weiterhin gilt, dass $[\text{id}_V]_{\mathcal{B}}^{\mathcal{C}} = ([\text{id}_V]_{\mathcal{C}}^{\mathcal{B}})^{-1}$. Daher gilt

$$\begin{aligned} \det[T]_{\mathcal{C}}^{\mathcal{C}} &= \det \left([\text{id}]_{\mathcal{C}}^{\mathcal{B}} \cdot [T]_{\mathcal{B}}^{\mathcal{B}} \cdot ([\text{id}_V]_{\mathcal{B}}^{\mathcal{C}})^{-1} \right) \\ &= \det[\text{id}]_{\mathcal{C}}^{\mathcal{B}} \cdot \det[T]_{\mathcal{B}}^{\mathcal{B}} \cdot \det([\text{id}_V]_{\mathcal{B}}^{\mathcal{C}})^{-1} \\ &= \det[T]_{\mathcal{B}}^{\mathcal{B}}. \end{aligned}$$

$\square$

Bemerkung 8.6.3. *Für unterschiedlich Basen $\mathcal{B}$ und $\mathcal{C}$ von V könnten wir natürlich auch $\det[T]_{\mathcal{B}}^{\mathcal{C}}$ betrachten. Dieser Wert ist nicht wohldefiniert!⁴⁶*

Satz 8.6.4. *Es sei V ein endlich-dimensionaler K -Vektorraum. Dann hat die Funktion*

$$\det : \text{Hom}_K(V, V) \rightarrow K, \quad T \mapsto \det(T)$$

folgende Eigenschaften:

⁴⁶Warum nicht?

(1) Es gilt

$$\det(S \circ T) = \det(S) \cdot \det(T) \quad \forall S, T \in \text{Hom}_K(V, V).$$

(2) T ist genau dann ein Isomorphismus, wenn $\det(T) \neq 0$. In diesem Fall gilt $\det(T^{-1}) = \det(T)^{-1}$.

(3) es gilt $\det(\mathbf{1}_V) = 1$ und $\det(0_V) = 0$.

Proof. (i) und (ii) folgen von Satz 6.1.5, Korollar 8.5.3 und Theorem 8.5.18. (iii) ist eine Übung. $\square$

Beispiel 8.6.5. Es sei V der Vektorraum aller Polynome ueber $\mathbb{Q}$ vom Grad ≤ 2 und $D : V \rightarrow V$ die Ableitungsabbildung. Dann ist die Abbildungsmatrix von D bezueglich der Basis $1, x, x^2$ gegeben durch $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, und diese Matrix hat Determinante 0. Dies war zu erwarten: da alle Konstanten unter D auf Null abgebildet werden, ist die Abbildung nicht invertierbar.

9. POLYNOME

Lecture 27

Zur Erinnerung: es sei K ein Koerper und $K[x]$ der Ring aller Polynome mit Koeffizienten in K . Fuer ein Element

$$f(x) = a_n x^n + \cdots + a_1 x + a_0 \in K[x],$$

$f \neq 0$ definieren wir den Grad $\deg f$ von $f(x)$ als die groesste ganze Zahl $m \geq 0$, so dass $a_m \neq 0$.

Beachte 9.0.1. Fuer $f(x), g(x) \in K[x]$, $f, g \neq 0$ gilt

$$\deg(fg) = \deg(f) + \deg(g).$$

Satz 9.0.2 (Division mit Rest). *Es sei K ein Koerper, und es seien $f(x), g(x) \in K[x]$, $g \neq 0$. Dann gibt es eindeutig bestimmte Polynome $q(x), r(x) \in K[x]$ with either $r = 0$ or $\deg(r) < \deg(g)$, such that*

$$f = qg + r.$$

Proof. Sollten Sie aus der Schule kennen. $\square$

Beispiel 9.0.3. Es seien

$$f(x) = x^3 - x^2 + x + 1 \quad \text{und} \quad g(x) = x^2 + 1.$$

Dann gilt

$$f(x) = (x - 1) \cdot g(x) + 2.$$

Korollar 9.0.4. *Es sei $f(x) \in K[x]$, $f \neq 0$, und es sei $\lambda \in K$ so dass $f(\lambda) = 0$. Dann gibt es $q(x) \in K[x]$, so dass*

$$f(x) = (x - \lambda)q(x);$$

die Nullstelle λ kann von $f(x)$ abgespalten werden.

Proof. Es sei $g(x) = x - \lambda$. Aufgrund von Satz 9.0.2 gibt es eindeutig bestimmte Polynome $q(x), r(x) \in K[x]$ with either $r = 0$ or $\deg(r) < 1$ (d.h. $r(x) \in K - \{0\}$), so dass

$$(43) \quad f(x) = g(x)q(x) + r(x).$$

Durch Auswertung von Gleichung (43) an $x = 0$ sehen wir, dass $r(0) = 0$ und daher $r = 0$. $\square$

Beispiel 9.0.5. Sei $f(x) = x^5 + 2x^4 - x^2 + 4 \in \mathbb{R}[x]$. Dann gilt $f(-2) = 0$, und

$$f(x) = (x^4 - x + 2)(x + 2).$$

Korollar 9.0.6. *Es sei $f(x) \in K[x]$, $\deg(f) = n > 0$. Dann hat $f(x)$ hoechstens n Nullstellen (nicht notwendigerweise verschieden) in K .*

Beispiel 9.0.7. Es sei $f(x) = (x + 1)(x^2 + 1)$. Als ein Polynom in $\mathbb{R}[x]$ hat f genau eine Nullstelle, aber es hat drei Nullstellen als ein Element von $\mathbb{C}[x]$.

Theorem 9.0.8 (Fundamentale Satz der Algebra). *Es sei $f(x) \in \mathbb{C}[x]$ vom Grad $n > 0$. Dann hat $f(x)$ genau n Nullstellen⁴⁷ in $\mathbb{C}$, das heisst, es gibt $\lambda_1, \dots, \lambda_n \in \mathbb{C}$, nicht notwendigerweise verschieden, so dass*

$$f(x) = (x - \lambda_1) \cdots (x - \lambda_n).$$

⁴⁷Ein Koerper mit dieser Eigenschaft heisst *algebraisch abgeschlossen*.

Beispiel 9.0.9. Es sei $m \geq 1$ und $f(x) = x^m - 1 \in \mathbb{C}[x]$. Es sei $\zeta_m = e^{\frac{2\pi i}{m}}$. Dann gilt

$$f(x) = \prod_{i=0}^{m-1} (x - \zeta_m^i).$$

Bemerkung 9.0.10. Man kann zeigen, dass jeder K rper in einem algebraisch abgeschlossenen K rper enthalten ist.

Definition 9.0.11. Es sei $f(x) \in K[x]$, $f \neq 0$, und es sei $\lambda \in K$ so dass $f(\lambda) = 0$. Die Ordnung (oder Vielfachheit) der Nullstelle λ in $f(x)$ ist die ganze Zahl $n \geq 0$, f r die es ein $q(x) \in K[x]$ gibt, so dass

$$f(x) = (x - \lambda)^n \cdot q(x) \quad \text{und} \quad q(\lambda) \neq 0.$$

Wenn die Ordnung der Nullstelle 1 ist, dann ist λ eine einfache Nullstelle von $f(x)$.

Beispiele 9.0.12. (1) Es sei $f(x) = (x-1)^2(x+2)$. Die Nullstell $x = 1$ hat Ordnung 2, und $x = -2$ ist eine einfache Nullstelle.

(2) Es sei $p > 2$ eine Primzahl und $\mathbb{F}_p$ der K rper mit p Elementen. Es sei $g(x) = x^p - 1 \in \mathbb{F}_p[x]$. Dann gilt

$$g(x) = (x - 1)^p,$$

das heisst, $x = 1$ ist eine Nullstelle von $g(x)$ der Ordnung $p!$

Bemerkung 9.0.13. (1) Man kann mit Polynomdivision viel Spass haben. Ein Polynom $f(x) \in K[x]$ ist unzerlegbar, wenn es keine Polynome $g(x), h(x) \in K[x]$ gibt mit $\partial(g), \partial(h) > 0$, so dass $f(x) = g(x)h(x)$. Es folgt dann von Satz 9.0.2, dass sich jedes Polynom vom Grad > 0 eindeutig⁴⁸ in ein Produkt von unzerlegbaren Polynomen schreiben l sst. Unzerlegbare Polynome spielen also die Rolle der Primzahlen in dem Ring $K[x]$! Diese Beobachtung ist von sehr grosser Wichtigkeit in der algebraischen Geometrie und der algebraischen Zahlentheorie.

(2) Eine andere sehr interessante Frage ist, wie man die Nullstellen eines Polynoms finden kann. F r Polynome vom Grad ≤ 4 gibt es explizite Formeln f r die Nullstellen, aber bei Polynomen vom h heren Grad gilt dies nicht mehr. Dieses ist eines der Hauptresultate der Galois Theorie. Diese Theorie ordnet jedem Polynom eine endliche Gruppe zu, und die Eigenschaften der Gruppe entscheiden, ob das Polynom durch Wurzeln l sbar ist oder nicht.

10. EIGENWERTE UND EIGENVEKTOREN

10.1. Definitionen und erste Eigenschaften.

Definition 10.1.1. Es sei V ein endlich-dimensionaler Vektorraum  ber K und $T : V \rightarrow V$ eine lineare Abbildung.

- (1) Ein Element $\lambda \in K$ ist ein Eigenwert⁴⁹ von T , wenn es einen Vektor $v \in V$ gibt, $v \neq 0_V$, so dass $Tv = \lambda \cdot v$.
- (2) Ein Vektor $v \in V$, $v \neq 0_V$ mit der Eigenschaft $Tv = \lambda v$ ist ein Eigenvektor von T mit Eigenwert λ .

Wenn $A \in M_{n \times n}(K)$, dann sind die Eigenwerte und -vektoren die Eigenwerte und -vektoren der Abbildung $T_A : K^n \rightarrow K^n$.

Bemerkung 10.1.2. Wenn v ein Eigenvektor von T ist, dann ist auch αv ein Eigenvektor (mit dem gleichen Eigenwert) f r alle $\alpha \in K$, $\alpha \neq 0$. Ist die Menge aller Eigenvektoren mit dem gleichen Eigenwert ein Unterraum von V ?

Beispiele 10.1.3.

- (1) In Beispiel 8.6.5 ist die konstante Funktion 1 ein Eigenvektor mit Eigenwert 0.
- (2) Es sei $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Durch Betrachtung der Matrix sehen wir, dass $\lambda = 3$ und $\lambda = -1$ Eigenwerte von A sind, da

$$A \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad A \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Sind das alle Eigenwerte von A ?

⁴⁸abgesehen von Multiplikation mit Elementen in $K - \{0\}$.

⁴⁹engl: eigenvalue

- (3) Es sei $B = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}$. Wenn λ ein Eigenwert von B ist, dann gibt es einen Vektor $v = \begin{pmatrix} x \\ y \end{pmatrix}$, $v \neq 0$, so dass $B \cdot v = \lambda v$, das heisst

$$\begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow \left[\begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wir erhalten daraus das lineare Gleichungssystem,

$$\begin{aligned} (1 - \lambda)x - 2y &= 0 \\ x + (4 - \lambda)y &= 0. \end{aligned}$$

Indem wir die zweite Gleichung mit $-(1 - \lambda)$ multiplizieren und zur ersten Gleichung addieren, erhalten wir

$$(-(1 - \lambda)(4 - \lambda) - 2)y = 0 \rightarrow y = 0 \vee \lambda^2 - 5\lambda + 6 = 0.$$

- Wenn $y = 0$, dann folgt, dass $x = 0$.
- Wenn $\lambda^2 - 5\lambda + 6 = 0$, dann gilt $\lambda = 2$ oder $\lambda = 3$. Fuer jeden dieser Werte loesen wir das lineare Gleichungssystem:

$$\begin{aligned} \lambda = 2 &\Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \alpha \cdot \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \\ \lambda = 3 &\Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \beta \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \end{aligned}$$

Mit anderen Worten: 2 und 3 sind die Eigenwerte von B , mit Eigenvektoren $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Wir haben diese Eigenwerte durch Berechnung der Nullstellen des Polynoms $\lambda^2 - 5\lambda + 6$ bestimmt. Laesst sich diese Beobachtung verallgemeinern?

Satz 10.1.4. Es sei $T : V \rightarrow V$ linear. Dann gilt: $\lambda \in K$ ist genau dann ein Eigenvalue von T , wenn $\ker(T - \lambda \mathbf{1}_V) \neq \{0_V\}$.

Lecture 28

Proof.

$$\begin{aligned} \lambda \in K \text{ ist ein Eigenvalue von } T &\Leftrightarrow \exists v \in V, v \neq 0_V \text{ so dass } T.v = \lambda v \\ &\Leftrightarrow \exists v \in V, v \neq 0_V \text{ so dass } (T - \lambda \mathbf{1}_V).v = 0_V \\ &\Leftrightarrow v \in \ker(T - \lambda \mathbf{1}_V). \end{aligned}$$

□

Bemerkung 10.1.5. Insbesondere ist 0 genau dann ein Eigenwert von T , wenn T nicht invertierbar ist.

Korollar 10.1.6. Folgende Aussagen sind aequivalent:

- (1) $\lambda \in K$ ist ein Eigenwert von T ;
- (2) $\ker(T - \lambda \mathbf{1}_V) \neq \{0_V\}$;
- (3) $T - \lambda \mathbf{1}_V$ ist kein Isomorphismus;
- (4) $\det(T - \lambda \mathbf{1}_V) = 0$.

Proof. (i) $\Leftrightarrow$ (ii) ist Satz 10.1.4, und (iii) $\Leftrightarrow$ (iv) ist Satz 8.6.4 (ii). Die Aequivalenz (ii) $\Leftrightarrow$ (iii) ist Theorem 4.2.9. □

Bemerkung 10.1.7. Korollar 10.1.6 (iv) gibt uns ein sehr nuetzliches Kriterium, um die Eigenwerte einer Matrix bzw. linearen Abbildung zu bestimmen!

Beispiele 10.1.8.

- (1) Zurueck zum Beispiel 10.1.3 (iii): die Eigenwerte der Matrix B sind genau die Wurzeln des Polynoms

$$\begin{aligned} \det \begin{pmatrix} 1 - \lambda & -2 \\ 1 & 4 - \lambda \end{pmatrix} &= (1 - \lambda)(4 - \lambda) - (-2) \\ &= \lambda^2 - 5\lambda + 6. \end{aligned}$$

- (2) Wir koennen ausserdem die Frage in Beispiel 10.1.3 (ii) beantworten: $\lambda = 3$ und $\lambda = 1$ sind die einzigen Eigenwerte von A , da sie genau die Wurzeln des Polynoms

$$\det \begin{pmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix} = \lambda^2 - 2\lambda - 3$$

sind.

- (3) Es sei V der Vektorraum aller Fibonacci-Folgen und $S : V \rightarrow V$ die Verschiebungsabbildung (siehe Satz 1.1.15). Die Abbildungsmatrix von S bezueglich der Basis $\mathcal{B} = \{\mathcal{F}_{0,1}, \mathcal{F}_{1,0}\}$ ist

$$[S]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

und es gilt

$$\det \left([S]_{\mathcal{B}}^{\mathcal{B}} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) = \lambda^2 - \lambda - 1,$$

das heisst, die Eigenwerte dieser Matrix sind $\frac{1 \pm \sqrt{5}}{2}$. Vergleichen Sie das mit dem Beweis von Theorem 1.1.19.

10.2. Das charakteristische Polynom.

Definition 10.2.1. Es sei $A \in M_{n \times n}(K)$. Dann ist

$$\chi_A(x) = \det(A - x \cdot \mathbf{1}_n) \in K[x]$$

das charakteristische Polynom von A .

Beispiel 10.2.2. Es sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Dann ist

$$\chi_A(x) = x^2 - (a+d)x + (ad-bc).$$

Insbesondere gilt

$$\chi_{\mathbf{1}_2}(x) = x^2 - 2x + 1 = (x-1)^2 :$$

der Eigenwert $\lambda = 1$ ist eine doppelte Nullstelle des charakteristischen Polynoms.

Definition 10.2.3. Fuer eine lineare Abbildung $T : V \rightarrow V$ definieren wir das charakteristische Polynom von T als

$$\chi_T(x) = \det([T]_{\mathcal{B}}^{\mathcal{B}} - x \cdot \mathbf{1}_n),$$

wobei $\mathcal{B}$ eine beliebige Basis von V ist.

Beachte 10.2.4. Wenn $\mathcal{B}$ und $\mathcal{C}$ zwei verschiedene Basen von V sind und $A = [\mathbf{1}_V]_{\mathcal{B}}^{\mathcal{C}}$ die Basiswechselmatrix, dann gilt

$$[T]_{\mathcal{B}}^{\mathcal{B}} = C^{-1}[T]_{\mathcal{C}}^{\mathcal{C}}C$$

und daher

$$\begin{aligned} \det([T]_{\mathcal{B}}^{\mathcal{B}} - x \cdot \mathbf{1}_n) &= \det(A^{-1}[T]_{\mathcal{C}}^{\mathcal{C}}A - x \cdot \mathbf{1}_n) \\ &= \det[A^{-1}([T]_{\mathcal{C}}^{\mathcal{C}}A - xA\mathbf{1}_n \cdot A^{-1})A] \\ &= \det(A^{-1}) \cdot \det([T]_{\mathcal{C}}^{\mathcal{C}} - x\mathbf{1}_n) \cdot \det(A) \\ &= \det([T]_{\mathcal{C}}^{\mathcal{C}} - x\mathbf{1}_n), \end{aligned}$$

das heisst, das charakteristische Polynom von T ist wohl-definiert.

Wir koennen Korollar 10.1.6 folgendermassen zusammenfassen:

Theorem 10.2.5. Es sei $T : V \rightarrow V$ linear. Dann ist $\lambda \in K$ genau dann ein Eigenwert von T , wenn $\chi_T(\lambda) = 0$. Mit anderen Worten:

$$\{\text{Eigenwerte von } T\} = \{\lambda \in K : \chi_T(\lambda) = 0\}.$$

Was koennen wir ausserdem ueber das charakteristische Polynom sagen?

Lemma 10.2.6.

(1) Es sei $A = (a_{ij}) \in M_{n \times n}(K)$ eine obere Dreiecksmatrix. Dann gilt

$$\chi_A(x) = \prod_{i=1}^n (a_{ii} - x),$$

und die Eigenwerte sind $a_{11}, \dots, a_{nn}$.

(2) Es sei $M \in M_{n \times n}(K)$ von der Form

$$M = \begin{pmatrix} A & B \\ 0_{s \times r} & C \end{pmatrix} \quad \text{für } A \in M_{r \times r}(K), B \in M_{r \times s}(K) \text{ und } C \in M_{s \times s}(K).$$

Dann gilt

$$\chi_M(x) = \chi_A(x) \cdot \chi_C(x).$$

Proof. Übung. □

Wir koennen auch einiges ueber die Koeffizienten des charakteristischen Polynoms sagen: wir erinnern uns an die Spurabbildung (Beispiel 4.1.3 (v))

$$\text{Tr} : M_{n \times n}(K) \rightarrow K, \quad A = (a_{ij}) \mapsto a_{11} + \dots + a_{nn}.$$

Lemma 10.2.7. Es gilt

$$\text{Tr}(AB) = \text{Tr}(BA) \quad \forall A, B \in M_{n \times n}(K).$$

Proof. Explizite Rechnung. □

Definition 10.2.8. Es sei $T : V \rightarrow V$ linear. Wir definiere die Spur von T as

$$\text{Tr}(T) = \text{Tr}[T]_{\mathcal{B}}^{\mathcal{B}}$$

fuer eine beliebige Basis $\mathcal{B}$ von V .

Satz 10.2.9. $\text{Tr}(T)$ ist wohl-definiert.

Proof. Es seien $\mathcal{B}$ und $\mathcal{C}$ verschiedene Basen von V , und es sei $C = [\text{id}_V]_{\mathcal{B}}^{\mathcal{C}}$, so dass

$$[T]_{\mathcal{B}}^{\mathcal{B}} = C^{-1} \cdot [T]_{\mathcal{C}}^{\mathcal{C}} \cdot C.$$

Eine Anwendung von Lemma 10.2.7 zeigt, dass

$$\text{Tr}([T]_{\mathcal{B}}^{\mathcal{B}}) = \text{Tr}(C^{-1} \cdot [T]_{\mathcal{C}}^{\mathcal{C}} \cdot C),$$

d.h. $\text{Tr}(T)$ ist wohldefiniert. □

Satz 10.2.10. Es sei $A \in M_{n \times n}$. Dann gilt

$$\chi_A(x) = (-1)^n x^n + (-1)^{n-1} \text{Tr}(A) x^{n-1} + \dots + \det(A).$$

Proof. Schreibe $A = (a_{ij})_{1 \leq i, j \leq n}$. Dann gilt

$$\chi_A(x) = \det \begin{pmatrix} a_{11} - x & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} - x & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} - x & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} - x \end{pmatrix}$$

Wir benutzen nun (38), um die Determinante zu berechnen. Es sei $\sigma \in S_n$. Wenn σ nicht die Identitaet ist, dann bildet σ hoechstens $n - 2$ Elemente aus der Menge $\{1, 2, \dots, n\}$ auf sich selber ab. Daher gilt

$$\chi_A(x) = \prod_{i=1}^n (a_{ii} - x) + g(x),$$

wobei $g(x)$ aus einer Summe von Produkten besteht, die hoechstens $n - 2$ der diagonalen Eintraege $a_{ii} - x$ enthalten. Mit anderen Worten, $g(x)$ ist ein Polynom vom Grad $\leq n - 2$. Daher gilt

$$\chi_A(x) = (-1)^n x^n + (-1)^{n-1} (a_{11} + a_{22} + \dots + a_{nn}) x^{n-1} + g(x).$$

Um den konstanten Koeffizienten von $\chi_T(x)$ zu berechnen, berechnen wir $\chi_A(0) = \det(A)$. □

Bemerkung 10.2.11. (1) Um den Koeffizienten von x^{n-1} auszurechnen koennten wir auch ueber Induktion nach n argumentieren (Uebung).

- (2) Felix Xu ist aufgefallen, dass wir von Satz 10.2.10 einen sehr eleganten Beweis dafür erhalten, dass die Spur einer linearen Abbildung wohldefiniert ist: wir wissen von 10.2.4, dass $\chi_T(x)$ unabhängig von der Wahl einer Basis von V ist; insbesondere gilt dies für den Koeffizienten von x^{n-1} . Nun wissen wir aber von Satz 10.2.10, dass dieser Koeffizient gegeben ist durch

$$(-1)^{n-1} \cdot \text{Tr}([T]_{\mathcal{B}}^{\mathcal{B}})$$

wobei $\mathcal{B}$ eine beliebige Basis von V ist; mit anderen Worten, der Wert $\text{Tr}([T]_{\mathcal{B}}^{\mathcal{B}})$ ist unabhängig von der Wahl der Basis. Dieses gibt einen neuen (und schöneren) Beweis von Satz 10.2.9.

Korollar 10.2.12. Es sei V ein n -dimensionaler Vektorraum und $T : V \rightarrow V$ linear. Dann hat T höchstens n Eigenwerte.

Proof. Klar, da $\chi_T(x)$ ein Polynom vom Grad n ist und daher höchstens n Nullstellen haben kann. $\square$

Bemerkung 10.2.13. Wenn λ ein Eigenwert von T ist, dann kann es sein, dass λ eine Nullstelle höherer Ordnung von $\chi_T(x)$ ist (e.g. für die Identitätsabbildung id_n für $n > 1$).

10.3. Diagonalisierung. Es sei $T : V \rightarrow V$ eine lineare Abbildung. Wir wollen nun die Idee von Eigenvektoren mit unserer Theorie des Basiswechsels zu verknüpfen: das Ziel ist es, um eine Basis $\mathcal{B}$ von V zu finden, so dass die Abbildungsmatrix $[T]_{\mathcal{B}}^{\mathcal{B}}$ eine besonders einfache Form hat.

Beispiel 10.3.1. Wir betrachten nochmal das Beispiel der Fibonacci-Folgen. Wir hatten bereits gesehen, dass $\varphi = \frac{1+\sqrt{5}}{2}$ und $\psi = \frac{1-\sqrt{5}}{2}$ die Eigenwerte der Abbildung S sind, mit dazu gehörenden Eigenfolgen $\mathcal{F}_{1,\varphi}$ und $\mathcal{F}_{1,\psi}$. Eine einfache Rechnung zeigt, dass die beiden Eigenfolgen linear unabhängig und daher eine Basis des Vektorraums aller Fibonacci-Folgen sind. Die Darstellungsmatrix von S bezüglich dieser Basis ist $\begin{pmatrix} \varphi & 0 \\ 0 & \psi \end{pmatrix}$: sie ist diagonal.

Satz 10.3.2. Es sei $T : V \rightarrow V$ eine lineare Abbildung, und $\lambda_1, \dots, \lambda_m$ verschiedene Eigenwerte von T . $\forall i$ sei v_i ein Eigenvektor mit Eigenwert λ_i . Dann sind $v_1, \dots, v_m$ linear unabhängig.

Proof. Nimm an, dass die Vektoren linear abhängig sind. Es sei $1 \leq k \leq m$ der kleinste Index, so dass die Menge $\{v_1, \dots, v_k\}$ linear abhängig ist. Dann gibt es Skalare $\alpha_1, \dots, \alpha_k \in K$, alle $\neq 0$, so dass

$$\alpha_1 v_1 + \dots + \alpha_k v_k = 0_V.$$

Wende die Abbildung $T - \lambda_k \mathbf{1}_V$ auf diese Gleichung an. Dann erhalten wir

$$\begin{aligned} (T - \lambda_k \mathbf{1}_V)(\alpha_1 v_1 + \dots + \alpha_k v_k) &= \alpha_1 (T - \lambda_k \mathbf{1}_V)v_1 + \dots + \alpha_k (T - \lambda_k \mathbf{1}_V)v_k \\ &= \alpha_1 (\lambda_1 - \lambda_k)v_1 + \dots + \alpha_{k-1} (\lambda_{k-1} - \lambda_k)v_{k-1} + \alpha_k (\lambda_k - \lambda_k)v_k \\ &= \alpha_1 (\lambda_1 - \lambda_k)v_1 + \dots + \alpha_{k-1} (\lambda_{k-1} - \lambda_k)v_{k-1} \\ &= 0_V. \end{aligned}$$

Da $\lambda_i \neq \lambda_j$ für $i \neq j$, gilt $\alpha_i (\lambda_i - \lambda_k) \neq 0$ für $1 \leq i < k$, mit anderen Worten, die Vektoren $v_1, \dots, v_{k-1}$ sind linear abhängig. Dies ist ein Widerspruch zu der Wahl von k . $\square$

Korollar 10.3.3. Es sei V ein n -dimensionaler K -Vektorraum und $T : V \rightarrow V$ eine lineare Abbildung. Nimm an, dass T genau n verschiedene Eigenwerte hat.⁵⁰ Dann besitzt V eine Basis, die aus Eigenvektoren von T besteht; wir sagen, T ist diagonalisierbar.

Proof. Es seien $\lambda_1, \dots, \lambda_n$ die Eigenwerte von T mit Eigenvektoren $v_1, \dots, v_n$. Dann sind $v_1, \dots, v_n$ linear unabhängig aufgrund von Satz 10.3.2 und daher eine Basis (Satz 3.3.23). $\square$

Definition 10.3.4. Eine lineare Abbildung $T : V \rightarrow V$ ist diagonalisierbar, wenn V eine Basis besitzt, die aus Eigenvektoren von T besteht; in diesem Fall ist die Abbildungsmatrix von T bezüglich dieser Basis diagonal.

Eine Matrix $A \in M_{n \times n}(K)$ ist diagonalisierbar, wenn die lineare Abbildung T_A diagonalisierbar ist.

Bemerkung 10.3.5. Eine Matrix $A \in M_{n \times n}(K)$ ist genau dann diagonalisierbar, wenn es eine invertierbare Matrix $B \in M_{n \times n}(K)$ gibt, so dass die Matrix $B^{-1}AB$ diagonal ist.

⁵⁰Wir haben bereits gesehen, dass das vorkommen kann, aber nicht immer der Fall ist!

Beispiel 10.3.6. (1) Die Matrix $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ in Beispiel 10.1.3 (2) ist diagonalisierbar, da sie zwei Eigenwerte hat, naemlich $\lambda_1 = 3$ und $\lambda_2 = -1$. Explizit koennen wir A folgendermassen diagonalisieren: die zu $\lambda_{1,2}$ gehoerigen Eigenvektoren sind $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, und sie bilden eine Basis $\mathcal{B} = \{v_1, v_2\}$ von $\mathbb{R}^2$. Die Basiswechselmatrix von der Standardbasis $\mathcal{E} = \{e_1, e_2\}$ nach $\mathcal{B}$ ist gegeben durch

$$C = [\text{id}]_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Dann gilt

$$C^{-1}AC = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}.$$

(2) Die Matrix $B = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}$ ist ebenfalls diagonalisierbar (Uebung).

Lemma 10.3.7. Wenn A diagonalisierbar ist, mit Eigenwerten $\lambda_1, \dots, \lambda_n$, dann gilt aufgrund von Lemma 10.2.6 (i)

$$\chi_A(x) = (x - \lambda_1) \cdots (x - \lambda_n).$$

Frage: Sind alle Matrizen/linearen Abbildungen diagonalisierbar?

Lecture 29

Beispiele 10.3.8.

- (1) Es ist nicht notwendig, dass die Eigenwerte einer Matrix verschieden sind, damit die Matrix diagonalisierbar ist: die Matrix $\mathbf{1}_2$ ist diagonal (und daher trivialerweise diagonalisierbar), hat aber nur einen Eigenwert $\lambda = 1$.
- (2) Bisweilen haengt es vom Grundkoerper ab, ob eine Matrix diagonalisierbar ist: es sei $M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_{2 \times 2}(\mathbb{R})$. Dann hat das charakteristische Polynom

$$\chi_M(x) = x^2 + 1$$

keine Nullstellen in $\mathbb{R}$, das heisst, M ist ueber $\mathbb{R}$ nicht diagonalisierbar. Betrachten wir M aber als eine Matrix in $M_{2 \times 2}(\mathbb{C})$, dann gilt

$$\chi_M(c) = (x - i)(x + i),$$

und M ist diagonalisierbar.

- (3) Es sei $D = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Ist D diagonalisierbar? Wir berechnen das charakteristische Polynom:

$$\chi_D(x) = \det \begin{pmatrix} 1-x & 1 \\ 0 & 1-x \end{pmatrix} = (1-x)^2,$$

das heisst, $\lambda = 1$ ist der einzige Eigenwert dieser Matrix⁵¹. Wir berechnen den dazugehoerigen Eigenvektor (oder sind es mehrere?):

$$\begin{aligned} D \cdot \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x \\ y \end{pmatrix} \\ \Rightarrow x + y &= x \quad \text{und} \quad y = y, \end{aligned}$$

das heisst, die Eigenvektoren mit Eigenvalue 1 sind von der Form $\alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ mit $\alpha \neq 0$. Mit anderen Worten, D ist nicht diagonalisierbar. Was ist passiert?

⁵¹Beachte, dass $\chi_D(x) = \chi_{\mathbf{1}_2}(x)$!

Beispiel 10.3.9. Beispiel 10.3.8 (iv) lässt sich folgendermassen verallgemeinern: es sei $n \geq 1$ und $\lambda \in K$. Definiere die Jordan'sche Blockmatrix $J_n(\lambda) \in M_{n \times n}(K)$,

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}.$$

Dann gilt

$$\chi_{J_n(\lambda)}(x) = (\lambda - x)^n,$$

das heisst, λ ist der einzige Eigenwert. Andererseits ist e_1 der einzige Eigenvektor mit Eigenwert λ : fuer $n > 1$ ist $J_n(\lambda)$ also nicht diagonalisierbar.

Wir koennen unsere Beobachtungen folgendermassen zusammenfassen: Die Identitaetsmatrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ und die Matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ haben beide charakteristisches Polynom $(1 - x)^2$; allerdings ist die erste diagonalisierbar und die zweite nicht diagonalisierbar. Mit anderen Worten, das charakteristische Polynom alleine kann *nicht* entscheiden, ob eine Matrix diagonalisierbar ist. Wichtig ist die Dimension des zu einem Eigenwert gehoerigen Unterraums, der von den Eigenvektoren aufgespannt wird.

10.4. Eigenraeume. Wir gehen von jetzt an davon aus, dass der Koerper K algebraisch abgeschlossen ist.

Definition 10.4.1. Es sei $T : V \rightarrow V$ eine lineare Abbildung, und es sei λ ein Eigenwert von T . Der Eigenraum von λ ist gegeben durch

$$E_\lambda = \{v \in V : T.v = \lambda v\}.$$

Mit anderen Worten, E_λ ist die lineare Huelle der zu λ gehoerigen Eigenvektoren.

Lemma 10.4.2. E_λ ist ein Unterraum von V .

Proof. Es folgt von der Definition, dass

$$E_\lambda = \ker(T - \lambda \cdot \text{id}_V),$$

und wir haben in Lemma 4.2.3 gezeigt, dass der Kern einer linearen Abbildung ein Unterraum ist.

Alternativ koennen Sie auch die Unterraum Axiome ueberpruefen. $\square$

Beispiele 10.4.3. (1) Es sei $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. Eine einfache Rechnung zeigt, dass

$$\begin{aligned} \chi_A(x) &= -x^3 + 3x + 2 \\ &= -(x - 2)(x + 1)^2, \end{aligned}$$

das heisst, die Eigenwerte sind $\lambda_1 = -1$ und $\lambda_2 = 2$. Um E_{λ_1} zu bestimmen, loesen wir das lineare Gleichungssystem

$$(A - \lambda_1 \mathbf{1}_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Wir erhalten

$$E_{\lambda_1} = \left\langle \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \right\rangle,$$

das heisst, E_{λ_1} ist 2-dimensional.

(2) In Beispiel 10.3.8 (iii) hat der zu $\lambda = 1$ gehoerige Eigenraum Dimension 1.

Wir definieren zunaechst die direkte Summe von Unterraemen eines Vektorraums:

Definition 10.4.4. Es sei V ein Vektorraum, und es seien $U_1, \dots, U_k \leq V$. Es sei $W = U_1 + \dots + U_k$. Dann ist W die direkte Summe von $U_1, \dots, U_k$ wenn jedes Element $w \in W$ in genau einer Weise als Summe

$$w = u_1 + \dots + u_k \quad \text{mit } u_i \in U_i \forall 1 \leq i \leq k$$

geschrieben werden kann. In diesem Fall schreiben wir

$$W = U_1 \oplus \dots \oplus U_k.$$

Bemerkung 10.4.5. Diese Definition verallgemeinert die Idee des Komplements eines Unterraums (vergl. Definition 3.4.7): wenn $U \leq V$, dann ist $W \leq V$ genau dann ein Komplement, wenn sich jedes Element $v \in V$ in genau einer Art und Weise als $u + w$ schreiben lässt mit $u \in U$ und $w \in W$; mit anderen Worten genau dann, wenn $V = U \oplus W$.

Lemma 10.4.6. In dem Kontext von Definition 10.4.4 ist W genau dann die direkte Summe von $U_1, \dots, U_k$, wenn die Gleichung

$$0_V = u_1 + \dots + u_k \quad \text{mit } u_i \in U_i \forall i$$

nur die Lösung

$$u_1 = u_2 = \dots = u_k = 0_V$$

hat.

Proof. Übung. □

Beispiele 10.4.7.

- (1) $\mathbb{R}^3$ ist die direkte Summe von $U_1 = \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\rangle$, $U_2 = \left\langle \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \right\rangle$ und $U_3 = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle$
- (2) $\mathbb{R}^2$ ist die Summe, aber nicht die direkte Summe von $U'_1 = \left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle$ und von $U_2 = \left\langle \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \end{pmatrix} \right\rangle$.

Beachte 10.4.8. Wenn $W = U_1 \oplus \dots \oplus U_k$ ist, dann gilt

$$\dim(W) = \dim(U_1) + \dots + \dim(U_k).$$

Weiterhin gilt: wenn $\mathcal{B}_i$ eine Basis von U_i ist $\forall 1 \leq i \leq k$, dann ist

$$\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$$

eine Basis von W .

Zurück zu den Eigenräumen:

Satz 10.4.9. Es sei $T : V \rightarrow V$ eine lineare Abbildung, und es seien $\lambda_1, \dots, \lambda_k$ die Eigenwerte von T . Dann ist der Unterraum

$$W = E_{\lambda_1} + \dots + E_{\lambda_k}$$

die direkte Summe $E_{\lambda_1} \oplus \dots \oplus E_{\lambda_k}$.

Proof. Wir benutzen Lemma 10.4.6: es seien $v_1, \dots, v_k$ mit $v_i \in E_{\lambda_i}$, nicht alle 0_V . Es seien $i_1, \dots, i_r$ die Indizes, für die $v_{i_j} \neq 0_V$. Da $\lambda_i \neq \lambda_j$ für $i \neq j$, sind $v_{i_1}, \dots, v_{i_r}$ linear unabhängig aufgrund von Satz 10.3.2; daher gilt⁵²

$$v_{i_1} + \dots + v_{i_r} \neq 0_V.$$

□

Korollar 10.4.10. Es sei $T : V \rightarrow V$ eine lineare Abbildung, und es seien $\lambda_1, \dots, \lambda_k$ die Eigenwerte von T . Dann ist T genau dann diagonalisierbar, wenn

$$(44) \quad \dim(V) = \dim E_{\lambda_1} + \dots + \dim E_{\lambda_k}.$$

⁵²Sie können alternativ auch mit Hilfe von Basen von Eigenräumen argumentieren.

10.5. Algebraische und geometrische Vielfachheit. Es sei K algebraisch abgeschlossen.

Bemerkung 10.5.1. Nimm an, dass V Dimension n hat. Dann ist $\deg(\chi_T(x)) = n$, und es gibt $a_1, \dots, a_k \geq 1$, so dass

$$\chi_T(x) = (\lambda_1 - x)^{a_1} \cdots (\lambda_k - x)^{a_k},$$

und insbesondere

$$(45) \quad n = a_1 + \cdots + a_n.$$

Wir koennen also einem Eigenwert λ von T zwei weitere Groessen zuordnen: die Dimension des dazugehoerigen Eigenraums E_λ und die Vielfachheit der Nullstelle $x = \lambda$ des charakteristischen Polynoms. Wie haengen diese zusammen?

Definition 10.5.2. Es sei $T : V \rightarrow V$ eine lineare Abbildung, und es sei λ ein Eigenwert von T .

- (1) Die geometrische Vielfachheit g_λ von λ ist $\dim_K E_\lambda$.
- (2) Die algebraische Vielfachheit a_λ von λ ist die Vielfachheit von λ als eine Nullstelle von $\chi_T(x)$.

Beispiele 10.5.3.

- (1) In Beispiel 10.4.3 (i) gilt $g_{\lambda_1} = a_{\lambda_1} = 2$.
- (2) In Beispiel 10.3.8 (4) gilt fuer $\lambda = 1$: $g_\lambda = 1$ und $a_\lambda = 2$.
- (3) Wir betrachten den Eigenwert $\lambda = 1$ der Einheitsmatrix $\mathbf{1}_2$: fuer ihn gilt $g_\lambda = a_\lambda = 2$.
- (4) Fuer den Jordanblock $J_n(\lambda)$ gilt

$$g_\lambda = 1 \quad \text{und} \quad a_\lambda = n.$$

Faellt Ihnen was auf?

Satz 10.5.4. Es sei $T : V \rightarrow V$ eine lineare Abbildung, und es sei λ ein Eigenwert von T . Dann gilt $g_\lambda \leq a_\lambda$.

Proof. Es sei $k = g_\lambda$. Waehle eine Basis $v_1, \dots, v_k$ des Eigenraums E_λ . Dann koennen wir $\{v_1, \dots, v_k\}$ zu einer Basis $\mathcal{B} = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ von V erweitern (Theorem 3.3.19). Dann ist die Abbildungsmatrix $[T]_{\mathcal{B}}^{\mathcal{B}}$ von der Form

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} \lambda \cdot \mathbf{1}_k & A \\ 0_{(n-k) \times k} & B \end{pmatrix} \quad \text{mit } A \in M_{k \times (n-k)}(K) \text{ und } B \in M_{(n-k) \times (n-k)}(K),$$

da T jeden Vektor in E_λ mit λ multipliziert. Daher gilt

$$[T]_{\mathcal{B}}^{\mathcal{B}} - x\mathbf{1}_n = \begin{pmatrix} (\lambda - x) \cdot \mathbf{1}_k & A \\ 0_{(n-k) \times k} & B - x\mathbf{1}_{n-k} \end{pmatrix},$$

und wir folgern von Lemma 10.2.6 (ii), dass

$$\begin{aligned} \chi_T(x) &= \det([T]_{\mathcal{B}}^{\mathcal{B}} - x\mathbf{1}_n) \\ &= \det((\lambda - x) \cdot \mathbf{1}_k) \cdot \det(B - x\mathbf{1}_{n-k}) \\ &= (\lambda - x)^k \cdot \chi_B(x) \end{aligned}$$

mit $\chi_B(x) \in K[x]$. Daher gilt $k \leq a_\lambda$. □

Korollar 10.5.5. Es sei $T : V \rightarrow V$ eine lineare Abbildung, und es seien $\lambda_1, \dots, \lambda_k$ die Eigenwerte von T . Dann ist T genau dann diagonalisierbar, wenn

$$g_{\lambda_i} = a_{\lambda_i} \quad \forall 1 \leq i \leq k.$$

Proof. Wir wissen von Korollar 10.4.10, dass T genau dann diagonalisierbar, wenn

$$(46) \quad \dim(V) = \dim E_{\lambda_1} + \cdots + \dim E_{\lambda_k}.$$

Von (45) erhalten wir, dass

$$g_{\lambda_1} + \cdots + g_{\lambda_k} \leq a_{\lambda_1} + \cdots + a_{\lambda_k} = \dim(V).$$

Da $g_{\lambda_i} \leq a_{\lambda_i} \quad \forall i$, ist die Summe $g_{\lambda_1} + \cdots + g_{\lambda_k}$ genau dann gleich $\dim(V)$, wenn $g_{\lambda_i} = a_{\lambda_i}$ fuer alle $1 \leq i \leq k$. □

Wir koennen unsere Resultate ueber diagonalisierbare Matrizen folgendermassen zusammenfassen:

Theorem 10.5.6. Es sei V ein n -dimensionaler Vektorraum und $T : V \rightarrow V$ eine lineare Abbildung. Dann sind folgende Aussagen aequivalent:

- (1) T ist diagonalisierbar;
 (2) fuer jeden Eigenwert λ von T gilt $a_\lambda = g_\lambda$;
 (3) es seien $\lambda_1, \dots, \lambda_k$ die Eigenwerte von T . Dann gilt

$$\chi_T(x) = \prod_{i=1}^k (\lambda_i - x)^{g_{\lambda_i}};$$

(4) $V = \bigoplus_{i=1}^k E_{\lambda_i}$.

Proof. Die Aequivalenz (1) $\Leftrightarrow$ (2) ist Korollar 10.5.5.
 Aufgrund der Definition der Eigenwerte gilt

$$\chi_T(x) = \prod_{i=1}^k (\lambda_i - x)^{a_{\lambda_i}}.$$

Da diese Faktorisierung eindeutig ist, gilt (2) $\Leftrightarrow$ (3).
 Die Aequivalenz (1) $\Leftrightarrow$ (4) ist Korollar 10.4.10. □

Frage. Nimm an, dass T nicht diagonalisierbar ist. Kann man dann trotzdem eine Basis von V finden, so dass die Abbildungsmatrix von T eine besonders einfache Form hat? Die Antwort ist ja, wenn das charakteristische Polynom $\chi_T(c)$ in Linearfaktoren zerfaellt: in diesem Fall kann man eine Basis finden, so dass die Abbildungsmatrix von T aus Jordan'schen Block-matrizen besteht; man nennt diese die *Jordansche Normalenform* von T . Wir werden in einem spaeteren Kapitel darauf zurueckkommen zurueckkommen.

11. DAS MINIMALE POLYNOM

Lecture 31

11.1. Definition und erste Eigenschaften.

Bemerkung 11.1.1. Fuer einen endlich-dimensionalen Vektorraum V und $T : V \rightarrow V$ eine lineare Abbildung definieren wir T^k , $k \geq 1$, als die k -fache Verknuepfung von T mit sich selbst:

$$T^k = T \circ \dots \circ T.$$

Weiterhin sei $T^0 = \text{id}_V$.

Dieses entspricht der folgenden Definition fuer Matrizen: es sei $A \in M_{n \times n}(K)$. Dann koennen wir fuer alle $k \geq 1$ das k -fache Produkt von A mit sich selber betrachten:

$$A^k = A \cdots A.$$

Wir definieren $A^0 = \mathbf{1}_n$.

Definition 11.1.2. Es sei $T \in \text{End}_K(V)$ und

$$g(x) = a_d x^d + a_{d-1} x^{d-1} + \dots + a_1 x + a_0 \in K[x].$$

Dann definieren wir

$$g(T) = a_d T^d + a_{d-1} T^{d-1} + \dots + a_1 T + a_0 \text{id}_V \in \text{End}_K(V).$$

Entsprechend definieren wir $g(A)$ fuer $A \in M_{n \times n}(K)$ als

$$g(A) = a_d A^d + a_{d-1} A^{d-1} + \dots + a_1 A + a_0 \mathbf{1}_n \in M_{n \times n}(K).$$

Bemerkung 11.1.3. Die Auswertung eines Polynoms an einer Matrix oder eines Endomorphismus respektiert die Korrespondenz zwischen Matrizen und Endomorphismen: wenn $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ fuer eine Basis $\mathcal{B}$ von V , dann gilt $g(A) = [g(T)]_{\mathcal{B}}^{\mathcal{B}}$ (Satz 6.1.5). Umgekehrt sei $A \in M_{n \times n}(K)$ und $T_A : K^n \rightarrow K^n$ der Endomorphismus bezueglich der Standardbasis $\mathcal{E}$. Dann gilt $T_{g(A)} = g(T_A)$.

Beispiele 11.1.4.

(1) Es sei $A = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$, und es sei $f(x) = x^2 - x + 3$. Dann gilt

$$\begin{aligned} f(A) &= \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}^2 - \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 4 \\ -2 & -1 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}. \end{aligned}$$

(2) Es sei $g(x) = x^n$ und $N_n \in M_{n \times n}(K)$ definiert durch

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

Dann gilt $g(N_n) = 0_{n \times n}$ ⁵³, aber $N_n^i \neq 0_{n \times n} \forall 0 \leq i < n$.

Satz 11.1.5. *Fuer alle $T \in \text{End}_K(V)$ ⁵⁴ gibt es $g(x) \in K[x]$, $g(x) \neq 0$, so dass $g(T) = 0_V$.*

Proof. $\text{End}_K(V) \cong M_{n \times n}(K)$ ist ein K -Vektorraum der Dimension n^2 ; daher sind die linearen Abbildungen $T^0, T^1, \dots, T^{n^2}$ linear abhaengig. Mit anderen Worten, es gibt $a_0, \dots, a_{n^2} \in K$, nicht alle 0, so dass

$$a_{n^2} T^{n^2} + \dots + a_1 T + a_0 \text{id}_V = 0_V.$$

□

Bemerkung 11.1.6. *Wenn $g(T) = 0_V$, dann gilt natuerlich auch $(\alpha g)(T) = 0_V$ fuer alle $\alpha \in K$.*

Beispiele 11.1.7.

- (1) Es sei $n \geq 1$. Dann gilt $g(\mathbf{1}_n) = 0_{n \times n}$ fuer $g(x) = x - 1$.
- (2) Nimm an, dass $T : V \rightarrow V$ diagonalisierbar ist, mit Eigenwerten $\lambda_1, \dots, \lambda_k$ mit (algebraischen oder geometrischen) Vielfachheiten $a_1, \dots, a_k$. Dann gibt es eine Basis von V , so dass die Abbildungsmatrix von T die Form hat

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} \lambda_1 \mathbf{1}_{a_1} & 0 & \dots & 0 \\ 0 & \lambda_2 \mathbf{1}_{a_2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_k \mathbf{1}_{a_k} \end{pmatrix}$$

Dann gilt

$$\chi_T([T]_{\mathcal{B}}^{\mathcal{B}}) = \begin{pmatrix} \chi_T(\lambda_1) \mathbf{1}_{a_1} & 0 & \dots & 0 \\ 0 & \chi_T(\lambda_2) \mathbf{1}_{a_2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \chi_T(\lambda_k) \mathbf{1}_{a_k} \end{pmatrix} = 0_{n \times n}.$$

Definition 11.1.8. *Es sei $T : V \rightarrow V$ linear. Das minimale Polynom von T ist das monische⁵⁵ Polynom $m_T(x) \in K[x]$ kleinsten Grades, so dass $m_T(T) = 0_V$.*

Wir muessen zunaechst zeigen, dass diese Definition sinnvoll ist:

Lemma 11.1.9. *Es seien $m(x), m'(x) \in K[x]$ beide monisch vom kleinsten Grad $d \geq 1$, so dass $m(T) = m'(T) = 0_V$. Dann gilt $m(x) = m'(x)$.*

Proof. Nimm an, dass $m(x) \neq m'(x)$. Dann ist das Polynom $g(x) = m(x) - m'(x)$ vom Grad $< d$, und es gilt $g(T) = 0$. Das ist ein Widerspruch zur Definition des minimalen Polynoms. □

⁵³Eine Matrix $A \in M_{n \times n}(K)$, fuer die es ein $m \geq 0$ gibt, so dass $A^m = 0_{n \times n}$, heisst *nilpotent*.

⁵⁴und entsprechend fuer alle $A \in M_{n \times n}(K)$

⁵⁵d.h. Leitkoeffizient ist gleich 1

Satz 11.1.10. Es sei $T : V \rightarrow V$ linear, und es sei $g(x) \in K[x]$, $g \neq 0$, so dass $g(T) = 0_V$. Dann gilt $m_T(x) \mid g(x)$.

Proof. Mithilfe von Polynomdivision finden wir Polynom $q(x), r(x) \in K[x]$, mit $r(x) = 0$ oder $\partial(r) < \partial(m)$, so dass

$$(47) \quad g(x) = q(x) \cdot m(x) + r(x).$$

- Wenn $r(x) = 0$, dann gilt $g(x) = q(x)m(x)$ und daher $m(x) \mid g(x)$.
- Wenn $r(x) \neq 0$, dann werten wir (47) an T aus und erhalten $r(T) = 0_V$. Doch dies gibt einen Widerspruch zu der Minimalität von $m_T(x)$.

□

Beispiele 11.1.11.

Lecture 32

- (1) Es sei $A = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$ mit charakteristischem Polynom $\chi_A(x) = (\lambda - x)(\mu - x)$. Was ist das minimale Polynom?

(a) wenn $\lambda = \mu$, dann gilt $m_A(x) = x - \lambda$.

(b) wenn $\lambda \neq \mu$, dann gilt $m_A(x) = \chi_A(x)$.

Wir sehen hier eine Besonderheit des Ringes $M_{2 \times 2}(K)$: er hat *Nullteiler*. im zweiten Fall gilt naemlich

$$m_A(A) = (\lambda \cdot \mathbf{I}_2 - A)(\mu \cdot \mathbf{I}_2 - A) = 0_{2 \times 2},$$

aber keine der Matrizen

$$\lambda \cdot \mathbf{I}_2 - A \quad \text{und} \quad \mu \cdot \mathbf{I}_2 - A$$

ist die Nullmatrix.

- (2) Wir koennen das erste Beispiel folgendermassen verallgemeinern. Nimm an, dass $T : V \rightarrow V$ diagonalisierbar ist, mit Eigenwerten $\lambda_1, \dots, \lambda_k$, $\lambda_i \neq \lambda_j$ fuer $i \neq j$. Wir haben bereits in Beispiel 11.1.7 gesehen, dass $\chi_T(T) = 0_V$. Was ist das minimale Polynom von T ?
- (3) Es sei $A = J_n(\lambda)$ fuer $n \geq 1$ und $\lambda \in K$. Dann folgt von Beispiel 11.1.4, dass

$$m_A(x) = (-1)^n \chi_A(x) = (x - \lambda)^n.$$

Vergleichen wir dieses mit dem minimalen Polynom der diagonalen Matrix $B = \lambda \mathbf{1}_n$: fuer sie gilt $m_B(x) = x - \lambda$. Die beiden Matrizen haben also unterschiedlichen minimalen Polynome aber das gleiche charakteristische Polynom. Mit anderen Worten, das minimale Polynom ‘sieht’ also in einem gewissen Sinne die 1en ueber den Diagonalen von A .

Definition 11.1.12. Es sei K algebraisch abgeschlossen, und es seien $f(x), g(x) \in K[x]$ gegeben durch

$$f(x) = (x - \lambda_1)^{a_1} \cdots (x - \lambda_m)^{a_m},$$

$$g(x) = (x - \lambda_1)^{b_1} \cdots (x - \lambda_n)^{b_n}$$

wobei $a_i, b_i \geq 0$ fuer all 1. Das kleinste gemeinsame Vielfache von $f(x)$ und $g(x)$ ist definiert als

$$\text{lcm}(f, g) = (x - \lambda_1)^{\max\{a_1, b_1\}} \cdots (x - \lambda_n)^{\max\{a_n, b_n\}}.$$

Beispiel 11.1.13. Es seien

$$f(x) = x(x - 3)^4(x + 1)^2, \quad g(x) = (x - 3)(x + 1)^3(x - 2)^2.$$

Dann ist

$$\text{lcm}(f, g) = x(x - 3)^4(x + 1)^3(x - 2)^2.$$

Folgender Satz wird in Kapitel 12 sehr wichtig:

Satz 11.1.14. Es sei K algebraisch abgeschlossen und $C \in M_{n \times n}(K)$. Nimm an, dass

$$C = \begin{pmatrix} A & 0_{m \times \ell} \\ 0_{\ell \times m} & B \end{pmatrix}$$

mit $A \in M_{m \times m}(K)$ und $B \in M_{\ell \times \ell}(K)$. Dann ist $m_C(x)$ das kleinste gemeinsame Vielfache von $m_A(x)$ und $m_B(x)$:

$$m_C(x) = \text{lcm}(m_A(x), m_B(x)).$$

Proof. Wir beachten zunachst, dass fuer all $f(x) \in K[x]$ gilt

$$f(C) = \begin{pmatrix} f(A) & 0 \\ 0 & f(B) \end{pmatrix};$$

daher gilt $f(C) = 0$ genau dann, wenn $f(A) = 0$ und $f(B) = 0$. Es sei nun $f(x) = \text{lcm}(m_A(x), m_B(x))$. Da $m_A(x)$ und $m_B(x)$ das Polynom $f(x)$ teilen, gilt $f(A) = 0$ und $f(B) = 0$ und daher $f(C) = 0$ was bedeutet, dass $m_C(x) | f(x)$. Wenn $f(x) \neq m_C(x)$, dann koennen nicht $m_A(x)$ und $m_B(x)$ das Polynom $m_C(x)$ teilen; sagen wir $m_A(x) \nmid m_C(x)$. Aber dann gilt $m_C(A) \neq 0$ und daher $m_C(C) \neq 0$, was ein Widerspruch ist. $\square$

Korollar 11.1.15. *Es sei K algebraisch abgeschlossen und $C \in M_{n \times n}(K)$. Nimm an, dass*

$$C = \begin{pmatrix} A_1 & 0 & 0 & 0 & \dots \\ 0 & A_2 & 0 & 0 & \dots \\ & & \dots & & \\ 0 & 0 & 0 & 0 & A_k \end{pmatrix}$$

mit $A_i \in M_{n_i \times n_i}(K)$. Dann gilt

$$m_C(x) = \text{lcm}\{m_{A_1}(x), \dots, m_{A_k}(x)\}.$$

Proof. Induktion. $\square$

Beispiele 11.1.16.

(1) Es sei

$$A = \begin{pmatrix} J_3(\lambda) & 0 & 0 \\ 0 & J_3(\lambda) & 0 \\ 0 & 0 & J_2(\lambda) \end{pmatrix}.$$

Dann gilt $m_A(x) = (x - \lambda)^3$.

(2) Es sei $\lambda \neq 0$ und

$$B = \begin{pmatrix} J_5(\lambda) & 0 & 0 \\ 0 & J_6(\lambda^{-1}) & 0 \\ 0 & 0 & J_4(\lambda^{-1}) \end{pmatrix}.$$

Dann gilt

$$m_B(x) = \begin{cases} (x - \lambda)^5 (x - \lambda^{-1})^6 & \text{wenn } \lambda \notin \{\pm 1\} \\ (x - \lambda)^6 & \text{wenn } \lambda \in \{\pm 1\}. \end{cases}$$

Bemerkung 11.1.17. Satz 11.1.14 gilt auch dann, wenn K nicht algebraisch abgeschlossen ist und sich die minimalen Polynome nicht in Linearfaktoren zerlegen lassen. Dazu braucht man das Konzept der unzerlegbaren Polynome, um definieren zu koennen, was das kleinste gemeinsame Vielfache von zwei Polynomen ist.

11.2. Der Satz von Cayley–Hamilton. Wir wollen uns nun mit dem Problem beschaeftigen, wie wir das minimale Polynom einer Matrix $A \in M_{n \times n}(K)$ bestimmen koennen. Aufgrund von Satz 11.1.10 waere es ein guter Anfang, irgendein Polynom $g(x) \in K[x]$ zu finden, so dass $g(A) = 0$. In Beispiel 11.1.7 haben wir gezeigt, dass $\chi_A(A) = 0_{n \times n}$, wenn A diagonalisierbar ist. Gilt das allgemein?

Wir wollen nun folgenden Satz beweisen:

Theorem 11.2.1 (Cayley–Hamilton). *Es sei $T \in \text{End}_K(V)$. Dann gilt $\chi_T(T) = 0_{\text{End}_K(V)}$.*

Korollar 11.2.2. $\forall T \in \text{End}_K(V)$ gilt $m_T(x) | \chi_T(x)$.

Proof. Folgt unmittelbar von Satz 11.1.10 und Theorem 11.2.1. $\square$

Lemma 11.2.3. Es sei $f(x) = x^k + a_{k-1}x^{k-1} + \dots + a_1x + a_0$, und es sei $A \in M_{k \times k}(K)$ gegeben durch⁵⁶

$$A = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 & -a_0 \\ 1 & 0 & 0 & \dots & 0 & 0 & -a_1 \\ 0 & 1 & 0 & \dots & 0 & 0 & -a_2 \\ 0 & 0 & 1 & \dots & 0 & 0 & -a_3 \\ & \dots & & \dots & & & \\ 0 & 0 & 0 & \dots & 1 & 0 & -a_{k-2} \\ 0 & 0 & 0 & \dots & 0 & 1 & -a_{k-1} \end{pmatrix}.$$

Dann ist $\chi_A(x) = (-1)^k f(x)$.

Proof. Übung. □

Lecture 33

Lemma 11.2.4. Es sei $T : V \rightarrow V$ eine lineare Abbildung. Es sei $W \subset K^n$ ein Unterraum, so dass $T(W) \subseteq W$ (ein solcher Unterraum heisst T -invariant), und es sei $T' = T|_W$. Dann gilt

$$\chi_{T'}(x) \mid \chi_T(x).$$

Proof. Es sei $w_1, \dots, w_k$ eine Basis von W , die wir zu einer Basis $\mathcal{B} = \{w_1, \dots, w_k, v_1, \dots, v_{n-k}\}$ von V erweitern. Dann ist die Abbildungsmatrix $[T]_{\mathcal{B}}^{\mathcal{B}}$ von der Form

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} A & B \\ 0_{(n-k) \times k} & C \end{pmatrix},$$

wobei $A \in M_{k \times k}(K)$ die Abbildungsmatrix von T' bezueglich der gewaehlten Basis von W ist. Dann gilt

$$\begin{aligned} \chi_T(x) &= \det([T]_{\mathcal{B}}^{\mathcal{B}} - x\mathbf{1}_n) \\ &= \det \begin{pmatrix} A - x\mathbf{1}_k & B \\ 0_{(n-k) \times k} & C - x\mathbf{1}_{n-k} \end{pmatrix} \\ &= \det(A - x\mathbf{1}_k) \cdot \det(C - x\mathbf{1}_{n-k}) \\ &= \chi_{T'}(x) \cdot \det(C - x\mathbf{1}_{n-k}). \end{aligned}$$

Daraus folgt, dass $\chi_{T'}(x) \mid \chi_T(x)$. □

Satz 11.2.5. Es sei $T : V \rightarrow V$ eine lineare Abbildung, und es sei $w \in V$. Es sei $W = \langle w, Tw, T^2w, \dots \rangle$. Dann gilt $T(W) \subseteq W$, und wenn $T' = T|_W$, dann gilt

$$\chi_{T'}(T').w = 0_V.$$

Proof. Es sei $k \geq 1$ maximal, so dass $\mathcal{B} = \{w, Tw, \dots, T^{k-1}w\}$ linear unabhaengig und daher eine Basis von W ist. Dann gilt $T^k w \in W$, das heisst es gibt $a_1, \dots, a_{k-1} \in K$, so dass

$$(48) \quad T^k w + a_{k-1} \cdot T^{k-1} w + \dots + a_1 \cdot Tw + a_0 \cdot w = 0_V.$$

Dann ist $[T]_{\mathcal{B}}^{\mathcal{B}}$ die Begleitmatrix des Polynoms

$$f(x) = x^k + a_{k-1}x^{k-1} + \dots + a_1x + a_0,$$

und es folgt von Lemma 11.2.3, dass $\chi_{T'}(x) = f(x)$. Daher ist $\chi_{T'}(T).w$ gleich (48), und der Satz folgt. □

Wir erhalten nun Theorem 11.2.1 als einfache Konsequenz:

Proof. Da $\chi_{T'}(x) \mid \chi_T(x)$ aufgrund von Lemma 11.2.4 und $w \in V$ beliebig war, gilt

$$\chi_T(T).w = 0_V \quad \forall w \in V.$$

□

⁵⁶Wir nennen A die *Begleitmatrix* von $f(x)$.

11.3. Ein alternativer Beweis von Cayley–Hamilton.

Bemerkung 11.3.1. Es sei $A \in M_{n \times n}(K)$. Indem wir Satz 10.3.11 auf die Matrix $A - x\mathbf{1}_n$ anwenden, erhalten wir

$$(49) \quad (A - x\mathbf{1}_n)\text{adj}(A - x\mathbf{1}_n) = \chi_A(x) \cdot \mathbf{1}_n.$$

Warum koennen wir A nicht direkt in diese Gleichung einsetzen und erhalten $\chi_A(A) = 0_{n \times n}$? Erinnern Sie sich, dass $\text{adj}(A - x\mathbf{1}_n)$ eine $(n \times n)$ -Matrix ist mit Koeffizienten in $k[x]$. Wir koennen nicht eine Matrix in eine andere einsetzen!

Bemerkung 11.3.2. Es sei $C = (c_{ij}(x)) \in M_{n \times n}(K[x])$ eine Matrix, so dass jeder Eintrag $c_{ij}(x)$ in Polynom in x vom Grad $\leq k$ ist. Dann gibt es Matrizen $B_0, \dots, B_k \in M_{n \times n}(K)$, so dass

$$C = B_k \cdot x^k + B_{k-1} \cdot x^{k-1} + \dots + B_1 \cdot x + B_0.$$

Beispiel 11.3.3. Es sei $A \in M_{2 \times 2}(\mathbb{R}[x])$ gegeben durch

$$A = \begin{pmatrix} 2x^2 + x & -2x - 3 \\ x^2 - 1 & -x^2 \end{pmatrix}.$$

Dann ist

$$\Psi(A) = \begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix} \cdot x^2 + \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix} \cdot x + \begin{pmatrix} 0 & -3 \\ -1 & 0 \end{pmatrix}.$$

Bemerkung 11.3.4. Wir koennen also $\text{adj}(A - x\mathbf{1}_n)$ schreiben als

$$(50) \quad \text{adj}(A - x\mathbf{1}_n) = B_{n-1}x^{n-1} + \dots + B_1x + B_0$$

mit $B_0, B_1, \dots, B_{k-1} \in M_{n \times n}(K)$. Koennnten wir hier nicht einfach A einsetzen? Die Antwort ist NEIN, weil wir nicht wissen, ob A mit den B_i kommutiert! Es gilt ja

$$B_{n-1}x^{n-1} + \dots + B_1x + B_0 = x^{n-1} \cdot B_{n-1} + \dots + x \cdot B_1 + B_0,$$

aber wir wissen nicht, ob

$$B_{n-1}A^{n-1} + \dots + B_1A + B_0 = A^{n-1} \cdot B_{n-1} + \dots + A \cdot B_1 + B_0!$$

Das Einsetzen von A in die Gleichung ist also a priori⁵⁷ nicht wohldefiniert.

Beweis von Theorem 11.2.1: Wir schreiben $\text{adj}(A - x\mathbf{1}_n) = (p_{ij}(x))_{1 \leq i, j \leq n}$, wobei $\forall i, j$ $p_{ij}(x) \in K[x]$ ein Polynom vom Grad $< n$ ist. Aufgrund von Bemerkung 11.3.2 koennen wir daher $\text{adj}(A - x\mathbf{1}_n)$ schreiben als

$$\text{adj}(A - x\mathbf{1}_n) = B_{n-1}x^{n-1} + \dots + B_1x + B_0$$

mit $B_0, B_1, \dots, B_{k-1} \in M_{n \times n}(K)$. Indem wir dies in (49) einsetzen, erhalten wir

$$(51) \quad (A - x\mathbf{1}_n)(B_{n-1}x^{n-1} + \dots + B_1x + B_0) = \chi_A(x) \cdot \mathbf{1}_n.$$

Wenn wir $\chi_A(x)$ schreiben als

$$\chi_A(x) = (-1)^n(x^n + \dots + a_1x + a_0),$$

dann koenne wir in (51) $\forall 0 \leq i \leq n$ die Koeffizienten von x^i vergleichen und erhalten folgendes System von Matrixgleichungen

$$\begin{aligned} AB_0 &= (-1)^n a_0 \mathbf{1}_n, \\ -B_0 + AB_1 &= (-1)^n a_1 \mathbf{1}_n, \\ &\vdots \\ -B_{n-2} + AB_{n-1} &= (-1)^n a_{n-1} \mathbf{1}_n, \\ -B_{n-1} &= (-1)^n \mathbf{1}_n. \end{aligned}$$

⁵⁷siehe Bemerkung 11.3.5

Indem wir diese Gleichungen von links mit $\mathbf{1}_n, A, A^2, \dots, A^n$ multiplizieren, erhalten wir

$$\begin{aligned} AB_0 &= (-1)^n a_0 \mathbf{1}_n, \\ -AB_0 + A^2 B_1 &= (-1)^n a_1 A, \\ &\vdots \\ -A^{n-1} B_{n-2} + A^n B_{n-1} &= (-1)^n a_{n-1} A^{n-1}, \\ -A^n B_{n-1} &= (-1)^n A^n \end{aligned}$$

Wenn wir alle diese Gleichungen addieren, dann ergibt die linke Seite $0_{n \times n}$, während die rechte Seite den Wert $\chi_A(A)$ ergibt. Daher gilt $\chi_A(A) = 0_{n \times n}$, was zu beweisen war.

Bemerkung 11.3.5. *Nimo Liu hat sehr scharfsinnig erkannt, wie sich obiger ‘falsche’ Beweis retten lässt: wie wir gesehen haben, erhalten wir durch Koeffizientenvergleich das System von Matrixgleichungen*

$$\begin{aligned} AB_0 &= (-1)^n a_0 \mathbf{1}_n, \\ -B_0 + AB_1 &= (-1)^n a_1 \mathbf{1}_n, \\ &\vdots \\ -B_{n-2} + AB_{n-1} &= (-1)^n a_{n-1} \mathbf{1}_n, \\ -B_{n-1} &= (-1)^n \mathbf{1}_n. \end{aligned}$$

Da

$$(A - x\mathbf{1}_n)\text{adj}(A - x\mathbf{1}_n) = \text{adj}(A - x\mathbf{1}_n)(A - x\mathbf{1}_n),$$

erhalten wir noch ein zweites System von Matrixgleichungen

$$\begin{aligned} AB_0 &= (-1)^n a_0 \mathbf{1}_n, \\ -B_0 + B_1 A &= (-1)^n a_1 \mathbf{1}_n, \\ &\vdots \\ -B_{n-2} + B_{n-1} A &= (-1)^n a_{n-1} \mathbf{1}_n, \\ -B_{n-1} &= (-1)^n \mathbf{1}_n. \end{aligned}$$

Wir folgern daraus, dass $AB_i = B_i A$ für alle $1 \leq i < n$, d.h. das Einsetzen von A in (50) ist letztendlich doch wohldefiniert!

12. DIE JORDANSCHEN NORMALENFORM EINER MATRIX

Annahme: Für den Rest dieses Kapitels sei K algebraisch abgeschlossen.

12.1. Definition und Theorem. Wir erinnern uns an die Definition der Jordanschen Blocks: es sei $n \geq 1$ und $\lambda \in K$. Dann ist der Jordanblock der Länge n und Eigenwert λ die Matrix $J_n(\lambda) \in M_{n \times n}(K)$,

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \lambda \end{pmatrix}$$

Folgende Eigenschaften des Jordanblocks haben wir bereits bewiesen:

Lemma 12.1.1. λ ist der einzige Eigenwert von $J_n(\lambda)$ mit $g_\lambda = 1$ und $a_\lambda = n$. Weiterhin gilt $E_\lambda = \langle e_1 \rangle$. Das minimale Polynom von $J_n(\lambda)$ ist

$$m_{J_n(\lambda)}(x) = (x - \lambda)^n.$$

Theorem 12.1.2. [Jordansche Normalenform] Es sei V ein n -dimensionaler Vektorraum ueber K und $T : V \rightarrow V$ linear. Dann gibt es eine Basis $\mathcal{B}$ von V , so dass

$$(52) \quad [T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} J_{n_1}(\alpha_1) & & & & \\ & J_{n_2}(\alpha_2) & & & \\ & & J_{n_3}(\alpha_3) & & \\ & & & \ddots & \\ & & & & J_{n_k}(\alpha_k) \end{pmatrix}$$

fuer $n_1, \dots, n_k \geq 1$ so dass $n_1 + \dots + n_k = n$ und $\alpha_1, \dots, \alpha_k \in K$ ⁵⁸. Weiterhin ist diese Darstellungsmatrix von T eindeutig, abgesehen von der moeglichen Vertauschung der Jordanbloecke.

Wir koennen natuerlich das Theorem auch in Matrixform ausdruecken:

Theorem 12.1.3. Es sei $A \in M_{n \times n}(K)$. Dann gibt es eine invertierbare Matrix $B \in M_{n \times n}(K)$, so dass $B^{-1}AB$ die Form (53) hat.

Bemerkung 12.1.4. Diagonalisierbare Matrizen sind ein Spezialfall von Theorem 12.1.2: in diesem Fall haben alle Jordanbloecke Laenge 1!

12.2. Eigenschaften der Jordanschen Normalenform. Wir wollen nun verstehen, was uns die Jordanform ueber die Eigenwerte und deren algebraischen und geometrischen Vielfachheiten verraeht; dieses wird uns spaeter dabei helfen, Theorem 12.1.2 zu beweisen. Wir sehen uns zunaechst ein Beispiel an:

Beispiel 12.2.1. Es seien $\lambda_1, \lambda_2 \in K$, voneinander verschieden. Wir betrachten die Matrix $B \in M_{9 \times 9}(K)$, gegeben durch

$$B = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & \lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & \lambda_1 & 1 & 0 & 0 & 0 & 0 \\ & & & & \lambda_1 & 0 & 0 & 0 & 0 \\ & & & & & \lambda_2 & 1 & 0 & 0 \\ & & & & & & \lambda_2 & 0 & 0 \\ & & & & & & & \lambda_2 & 0 \\ & & & & & & & & \lambda_2 \end{pmatrix}$$

Mit anderen Worten, B ist zusammengesetzt aus Jordanbloecken $J_3(\lambda_1)$, $J_2(\lambda_1)$, $J_2(\lambda_2)$ und zweimal $J_1(\lambda_2)$. Es folgt, dass

- $\chi_B(x) = (\lambda_1 - x)^5(\lambda_2 - x)^4$;
- die Eigenwerte von B sind λ_1 und λ_2 , mit Eigenraeumen
$$E_{\lambda_1} = \langle e_1, e_4 \rangle \quad E_{\lambda_2} = \langle e_6, e_8, e_9 \rangle;$$
- es gilt $a_{\lambda_1} = 5$ und $a_{\lambda_2} = 4$;
- es gilt $g_{\lambda_1} = 2$ und $g_{\lambda_2} = 3$;
- von Korollar 11.1.15 folgt, dass das minimale Polynom ist gegeben durch

$$m_B(x) = (x - \lambda_1)^3(x - \lambda_2)^2.$$

Um die Beobachtungen aus diesem Beispiel zu verallgemeinern, brauchen wir folgendes Lemma:

Lemma 12.2.2. Es sei $C \in M_{n \times n}(K)$ von der Form

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

mit $A \in M_{\ell \times \ell}(K)$ und $B \in M_{(n-\ell) \times (n-\ell)}(K)$. Definiere die Unterraume $U = \langle e_1, \dots, e_{\ell} \rangle$ und $W = \langle e_{\ell+1}, \dots, e_n \rangle$. Es sei $v = u + w \in V$, $v \neq 0$, mit $u \in U$ und $w \in W$. Dann ist v genau dann ein Eigenvektor von T_C mit Eigenwert λ , wenn $T_A(u) = \lambda u$ und $T_B(w) = \lambda w$. Weiterhin gilt⁵⁹

$$E_{\lambda}(T_C) = E_{\lambda}(T_A) \oplus E_{\lambda}(T_B)$$

⁵⁸nicht notwendigerweise verschieden!

⁵⁹Hier schreiben wir fuer eine lineare Abbildung $T : V \rightarrow V$ und $\lambda \in K$

$$E_{\lambda}(T) = \ker(T - \lambda \mathbf{1}_V)$$

und $g_{\lambda}(T) = \dim E_{\lambda}(T)$.

und daher $g_\lambda(T_C) = g_\lambda(T_A) + g_\lambda(T_B)$.

Proof. Übung. □

Theorem 12.2.3. *Es sei $T : V \rightarrow V$ linear. Nimm an, dass es eine Basis $\mathcal{B}$ von V gibt, so dass*

$$(53) \quad [T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} J_{n_1}(\alpha_1) & & & \\ & J_{n_2}(\alpha_2) & & \\ & & J_{n_3}(\alpha_3) & \\ & & & \ddots \\ & & & & J_{n_k}(\alpha_k) \end{pmatrix}$$

Es sei λ ein Eigenwert von T . Dann ist die geometrische Vielfachheit von λ die Anzahl der Jordanblöcke mit Eigenwert λ , d.h.

$$g_\lambda = \#\{i : 1 \leq i \leq k, \alpha_i = \lambda\}.$$

Das minimale Polynom von T ist gegeben durch

$$m_T(x) = \prod_{\lambda \text{ Eigenwert}} (x - \lambda)^{s(\lambda)},$$

wobei $s(\lambda)$ die Länge des grössten Jordanblocks mit Eigenwert λ ist, d.h.

$$s(\lambda) = \max\{n_j : 1 \leq j \leq k, \alpha_j = \lambda\}.$$

Proof. Schreibe die Basis $\mathcal{B}$ als

$$\mathcal{B} = \{b_1^{(1)}, \dots, b_{n_1}^{(1)}, b_1^{(2)}, \dots, b_{n_2}^{(2)}, \dots, b_1^{(k)}, \dots, b_{n_k}^{(k)}\},$$

und es sei $W_i = \langle b_1^{(i)}, \dots, b_{n_i}^{(i)} \rangle$, d.h. W_i ist der Unterraum von V mit Basis

$$\mathcal{B}^{(i)} = \{b_1^{(i)}, \dots, b_{n_i}^{(i)}\}$$

Dann gilt

$$V = W_1 \oplus \dots \oplus W_k,$$

und die Abbildungsmatrix von $T|_{W_i}$ bezüglich der Basis $\mathcal{B}^{(i)}$ ist gleich dem Jordanblock $J_{n_i}(\alpha_i)$. Dann folgt von Lemma 12.2.2, dass

$$g_\lambda(T) = g_\lambda(T|_{W_1}) + \dots + g_\lambda(T|_{W_k}).$$

Doch

$$g_\lambda(T|_{W_i}) = \begin{cases} 0 & \text{wenn } \lambda \neq \alpha_i \\ 1 & \text{wenn } \lambda = \alpha_i \end{cases}$$

Daraus folgt, dass

$$g_\lambda = \#\{i : 1 \leq i \leq k, \alpha_i = \lambda\}.$$

Laut Korollar 11.1.15 ist

$$m_T(x) = \text{lcm}\{m_{J_{n_1}}(\alpha_1), \dots, m_{J_{n_k}}(\alpha_k)\}.$$

Nun gilt

$$m_{J_{n_i}}(\alpha_i) = (x - \alpha_i)^{n_i}.$$

Mit anderen Worten, wenn λ ein Eigenwert von T ist, dann ist der Exponent des Faktors $(x - \lambda)$ in $m_T(x)$ gegeben durch

$$s(\lambda) = \max\{n_j : 1 \leq j \leq k, \alpha_j = \lambda\},$$

und daher gilt

$$m_T(x) = \prod_{\lambda \text{ Eigenwert}} (x - \lambda)^{s(\lambda)}.$$

□

12.3. Verallgemeinerte Eigenraeume.

Es sei V ein n -dimensionaler K -Vektorraum.

Definition 12.3.1. Es sei $T : V \rightarrow V$ linear, und es sei λ ein Eigenwert von T . Der verallgemeinerte Eigenraum von λ ist

$$\tilde{E}_\lambda = \bigcup_{j=1}^{\infty} \ker(T - \lambda \mathbf{1}_V)^j.$$

Lemma 12.3.2. Es sei $S : V \rightarrow V$ linear, und es sei $v \in V$, $v \neq 0$. Nimm an, dass $S^k v = 0_V$ fuer ein $k \geq 1$, aber dass $S^{k-1} v \neq 0_V$. Dann sind

$$v, Sv, \dots, S^{k-1}v$$

linear unabhaengig.

Proof. Nimm an, dass es Skalare $a_0, \dots, a_{k-1}$ gibt, so dass

$$(54) \quad a_0 v + a_1 Sv + \dots + a_{k-1} S^{k-1} v = 0_V.$$

Indem wir S^{k-1} auf die Gleichung (54) anwenden, erhalten wir $a_0 S^{k-1} v = 0_V$ und daher $a_0 = 0$. Dann wenden wir S^{k-2} auf (54) an und erhalten $a_1 S^{k-1} v = 0_V$ und daher $a_1 = 0$. Ueber Induktion erhalten wir daher $a_i = 0 \forall i$. $\square$

Lemma 12.3.3. Es gilt $\tilde{E}_\lambda = \ker(T - \lambda \mathbf{1}_V)^n$.

Proof. Es sei $S = T - \lambda \mathbf{1}_V$. Nimm an, dass es $v \in V$ und $m \geq 1$ gibt, so dass $S^m v = 0$ und $S^{m-1} v \neq 0_V$. Dann sind laut Lemma 12.3.2 die Vektoren $m, Sv, \dots, S^{m-1}v$ linear unabhaengig. Daher gilt $m \leq n$. $\square$

Definition 12.3.4. Es sei $v \in \tilde{E}_\lambda$, $v \neq 0_V$, und es sei $k \geq 1$ minimal, so dass $(T - \lambda \mathbf{1}_V)^k v = 0_V$. Dann ist die Menge

$$\{v, (T - \lambda \mathbf{1}_V)v, \dots, (T - \lambda \mathbf{1}_V)^{k-1}v\}$$

die Jordankette von v (der Laenge k).

Beachte 12.3.5. $(T - \lambda \mathbf{1}_V)^{k-1}v$ ist ein Eigenvektor von T mit Eigenwert λ .

Bemerkung 12.3.6. Jeder Eigenvektor von T bildet eine Jordankette der Laenge 1.

Beispiele 12.3.7.

(1) Es sei $A = J_n(\lambda)$ und $T_A : K^n \rightarrow K^n$ die dazugehoerige Abbildung. Was ist $\tilde{E}_\lambda$?

(2) Es seien $\mu, \lambda \in K$, $\mu \neq \lambda$, und $A = \begin{pmatrix} J_2(\lambda) & 0 & 0 & 0 & 0 \\ 0 & J_3(\lambda) & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & J_2(\mu) & 0 \\ 0 & 0 & 0 & 0 & \mu \end{pmatrix}$. Wir betrachten die

Abbildung $T_A : K^9 \rightarrow K^9$. Dann ist

$$\tilde{E}_\lambda = \langle e_1, \dots, e_6 \rangle \quad \text{und} \quad \tilde{E}_\mu = \langle e_7, e_8, e_9 \rangle.$$

Die Jordankette von e_2 ist $\{e_2, e_1\}$, von e_5 erhalten wir $\{e_5, e_4, e_3\}$.

(3) Es sei B wie in Beispiel 12.2.1. Dann hat T_B die verallgemeinerten Eigenraeume

$$\tilde{E}_{\lambda_1} = \langle e_1, \dots, e_5 \rangle \quad \text{und} \quad \tilde{E}_{\lambda_2} = \langle e_6, \dots, e_9 \rangle.$$

Was faellt auf?

Lemma 12.3.8. Fuer jeden Eigenwert λ ist $\tilde{E}_\lambda$ ein T -invarianter Unterraum von V .

Proof. Da $\tilde{E}_\lambda = \ker(T - \lambda \mathbf{1}_V)^n$, folgt unmittelbar, dass $\tilde{E}_\lambda \leq V$. Wir muessen daher nur zeigen, dass

$$T(\tilde{E}_\lambda) \subseteq \tilde{E}_\lambda.$$

Nimm an, dass $(T - \lambda \mathbf{1}_V)^n v = 0_V$. Dann gilt

$$(T - \lambda \mathbf{1}_V)^n T v = T(T - \lambda \mathbf{1}_V)^n v = 0_V,$$

und daher $T v \in \tilde{E}_\lambda$. $\square$

Als einfache Konsequenz erhalten wir folgenden Satz:

Satz 12.3.9. λ ist der einzige Eigenwert von $T|_{\tilde{E}_\lambda}$.

Proof. Es sei μ ein Eigenwert von $T|_{\tilde{E}_\lambda}$ mit Eigenvektor v . Es sei $k \geq 1$ minimal so dass $(T - \lambda \text{id}_V)^k v = 0$, und es sei $w = (T - \lambda \text{id}_V)^{k-1} v$. Dann gilt

$$(T - \lambda \text{id}_V)w = (T - \lambda \text{id}_V)^k v = 0_V = (\mu - \lambda)w$$

und daher $\mu = \lambda$. □

Korollar 12.3.10. Es sei $T_\lambda = T|_{\tilde{E}_\lambda}$. Dann gibt es $k_\lambda \leq a_\lambda(T)$ so dass

$$\chi_{T_\lambda}(x) = (\lambda - x)^{k_\lambda}.$$

Proof. Da K algebraisch abgeschlossen ist und λ der einzige Eigenwert von T_λ ist, gilt $\chi_{T_\lambda}(x) = (\lambda - x)^{k_\lambda}$ fuer ein $m_\lambda \geq 1$. Laut Lemma 11.2.4 gilt $\chi_{T_\lambda}(x)|\chi_T(x)$, und daher $k_\lambda \leq a_\lambda(T)$. □

Lemma 12.3.11. Es seien μ, λ verschiedene Eigenwerte von T . Dann gilt

$$\tilde{E}_\lambda \cap \tilde{E}_\mu = \{0_V\}.$$

Mit anderen Worten, der Unterraum $\tilde{E}_\lambda + \tilde{E}_\mu$ ist die direkte Summe $\tilde{E}_\lambda \oplus \tilde{E}_\mu$.

Proof. Es sei $U = \tilde{E}_\lambda \cap \tilde{E}_\mu$, und nimm an, dass $U \neq \{0_V\}$. Dann ist U ein T -invarianter Unterraum (warum?), und daher gilt $\chi_{T|U}(x)|(x - \mu)^{k_\mu}$ und $\chi_{T|U}(x)|(x - \lambda)^{k_\lambda}$. Da $\partial \chi_{T|U}(x) = \dim U$ und $\lambda \neq \mu$, erhalten wir einen Widerspruch. □

Wir koennen dieses Lemma verallgemeinern:

Lemma 12.3.12. Es seien $\lambda_1, \dots, \lambda_k$ voneinander verschiedene Eigenwerte von T . Dann ist der Unterraum $W = \tilde{E}_{\lambda_1} + \dots + \tilde{E}_{\lambda_k}$ die direkte Summe

$$W = \tilde{E}_{\lambda_1} \oplus \dots \oplus \tilde{E}_{\lambda_k}.$$

Proof. Induktion nach k . □

Bemerkung 12.3.13. Vergleichen sie dieses Lemma mit Satz 10.4.9.

Wie moechten nun zeigen, dass V die direkte Summe seiner verallgemeinerten Eigenraeume ist – *unabhaengig davon, ob T diagonalisierbar ist oder nicht*. Als Vorbereitung brauchen wir folgendes Lemma:

Lemma 12.3.14.

- (1) Es sei $T : V \rightarrow V$ linear, und es sei λ ein Eigenwert von T . Es sei $g(x) = (x - \lambda)^n$. Dann ist $\text{im}(g(T))$ ein T -invarianter Unterraum von V , und es gilt

$$V = \tilde{E}_\lambda \oplus \text{im}(g(T)).$$

- (2) Es sei μ ein von λ verschiedener Eigenwert von T . Dann gilt

$$\tilde{E}_\mu \subseteq \text{im}(g(T)).$$

Proof. (1) Die Invarianz von $\text{im}(g(T))$ ist eine Uebung. Von Theorem 4.2.9 wissen wir, dass

$$\dim V = \dim \ker(g(T)) + \dim \text{im}(g(T));$$

es ist daher ausreichend zu zeigen, dass $\ker(g(T)) \cap \text{im}(g(T)) = \{0\}$. Es sei $w \in \text{im}(g(T))$, d.h. es gibt $v \in V$ so dass $g(T)v = w$. Wenn $w \in \tilde{E}_\lambda$, dann gilt

$$(T - \lambda \text{id}_V)^n w = (T - \lambda \text{id}_V)^{2n} v = 0_V,$$

das heisst $v \in \tilde{E}_\lambda$. Nun wissen wir aber von Lemma 12.3.3, dass $\tilde{E}_\lambda = \ker(T - \lambda \text{id}_V)^n$, und daher gilt

$$w = (T - \lambda \text{id}_V)^n v = 0_V.$$

(2) Es sei $S = (T - \lambda \text{id}_V)^n$. Eine einfache Rechnung zeigt, dass $\tilde{E}_\mu$ ein S -invarianter Unterraum von V ist. Es sei $S' = S|_{\tilde{E}_\mu}$. Dann ist

$$\ker(S') = \tilde{E}_\lambda \cap \tilde{E}_\mu = \{0\} \quad (\text{aufgrund von Lemma 12.3.11}),$$

das heisst, S' ist ein Isomorphismus. Insbesondere gilt $\tilde{E}_\mu = \text{im}(S') \subseteq \text{im}(g(T))$, quod erat demonstrandum. □

Theorem 12.3.15. *Es sei $T : V \rightarrow V$ linear. Dann gilt*

$$V = \bigoplus_{\eta \text{ Eigenwert}} \tilde{E}_\eta :$$

V ist die direkte Summer seiner verallgemeinerten Eigenraume.⁶⁰

Proof. Wir benutzen Induktion ueber $n = \dim V$. Es sei λ ein Eigenwert von T . Dann wissen wir von Lemma 12.3.14, dass

$$V = \tilde{E}_\lambda \oplus \operatorname{im}((T - \lambda \mathbf{1}_V)^n),$$

und dass $U = \operatorname{im}((T - \lambda \mathbf{1}_V)^n)$ ein T -invarianter Unterraum ist, von Dimension $< n$. Daher gilt

$$U = \bigoplus_{\mu \text{ Eigenwert von } T|_U} \tilde{E}_\mu$$

und daher

$$\begin{aligned} V &= \tilde{E}_\lambda \oplus \left(\bigoplus_{\mu \text{ Eigenwert von } T|_U} \tilde{E}_\mu \right) \\ &= \bigoplus_{\eta \text{ Eigenwert von } T} \tilde{E}_\eta. \end{aligned}$$

□

Wir koennen dieses Theorem folgendermassen formulieren:

Korollar 12.3.16. *Es sei $T : V \rightarrow V$ linear, und es seien $\lambda_1, \dots, \lambda_k$ die Eigenwerte von T . Dann gibt es $n_1, \dots, n_k \geq 1$, so dass*

$$V = \ker((T - \lambda_1 \mathbf{1}_V)^{n_1}) \oplus \dots \oplus \ker((T - \lambda_k \mathbf{1}_V)^{n_k}).$$

Bemerkung 12.3.17. *Um Theorem 12.1.2 zu beweisen, reicht es also zu zeigen, dass jeder verallgemeinerte Eigenraum $\tilde{E}_\lambda$ eine Jordanbasis besitzt.*

12.4. Beweis der JNF. Wir betrachten zunaechst den Spezialfall, dass $S : V \rightarrow V$ nilpotent ist.

Theorem 12.4.1. *Es sei V ein n -dimensionaler Vektorraum ueber K und $S : V \rightarrow V$ nilpotent. Dann gibt es $k \geq 1$, $n_1, \dots, n_k \geq 1$, so dass $n_1 + \dots + n_k = n$, und eine Basis $\mathcal{B}$ von V , die aus Jordanketten besteht, d.h. von der Form*

$$(55) \quad S^{n_1}u_1, \dots, Su_1, u_1, S^{n_2-1}u_2, \dots, u_2, \dots, S^{n_k-1}u_k, \dots, u_k,$$

wobei $S^{n_i}u_i = 0_V \ \forall i$. Weiterhin sind $n_1, \dots, n_k$ eindeutig bestimmt.

Bemerkung 12.4.2. *Die Basis (55) von V heisst Jordanbasis.*

Bemerkung 12.4.3. *Die Abbildungsmatrix von S bezueglic der Basis $\mathcal{B}$ ist in Jordanscher Normalenform:*

$$[S]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} J_{n_1}(0) & & & \\ & J_{n_2}(0) & & \\ & & J_{n_3}(0) & \\ & & \vdots & \\ & & & J_{n_k}(0) \end{pmatrix}.$$

Da die n_i eindeutig bestimmt sind, gilt das gleiche fuer die Jordansche Normalenform, abgesehen von einer Vertauschung der Bloecke.

Proof. Wir beweisen Theorem 12.4.1 ueber Induktion nach $n = \dim(V)$. Fuer $n = 1$ ist das Resultat klar. Nimm nun an, dass $n > 1$, und dass das Resultat wahr ist fuer alle Vektorraeume der Dimension $< n$. Beachte zunaechst, dass $S(V) \leq V$ (warum?). Wenn $S(V) = \{0_V\}$, dann ist S die Nullabbildung, und das Theorem haelt trivialerweise. Wir koennen daher annehmen, dass

$$\{0_V\} \subsetneq S(V) \subsetneq V.$$

⁶⁰Dieses gilt nicht, wenn $\chi_T(x)$ sich nicht in Linearfaktoren zerlegen laesst. Daher die Annahme, dass K algebraisch abgeschlossen ist.

Wir wenden nun die Induktionshypothese auf den S -invarianten Unterraum $S(V)$ an: es gibt also Elemente $v_1, \dots, v_\ell$ und $b_1, \dots, b_\ell \geq 1$, so dass $b_1 + \dots + b_\ell = \dim S(V)$ und

$$(56) \quad \mathcal{B}' = \{S^{b_1-1}v_1, \dots, v_1, \dots, S^{b_\ell-1}v_\ell, \dots, v_\ell\}$$

eine Basis von $S(V)$ ist, mit $S^{b_i}v_i = 0_V \forall i$.

Für $1 \leq i \leq \ell$ wähle $u_i \in V$, so dass $S(u_i) = v_i$. Dann gilt $S^{b_i}u_i = S^{b_i-1}v_i \in \ker(S) \forall 1 \leq i \leq \ell$. Beachte, dass die Vektoren $\{S^{b_1}u_1, \dots, S^{b_\ell}u_\ell\}$ Teil der Basis $\mathcal{B}'$ und daher linear unabhängig sind; wir können sie daher zu einer Basis

$$\{S^{b_1}u_1, \dots, S^{b_\ell}u_\ell, w_1, \dots, w_m\}$$

von $\ker(S)$ erweitern.

Behauptung. Die Vektoren

$$\{S^{b_1}u_1, Su_1, \dots, u_1, S^{b_2}u_2, \dots, u_2, \dots, S^{b_\ell}u_\ell, \dots, u_\ell, w_1, \dots, w_m\}$$

sind eine Basis $\mathcal{B}$ von V .⁶¹

Beweis der Behauptung. Die lineare Unabhängigkeit ist eine Übung. Um zu zeigen, dass die Vektoren auch ein Erzeugendensystem sind, berechnen wir die Dimension von V . Wir wissen bereits, dass $\dim(\ker(S)) = \ell + m$ und dass

$$\operatorname{rk}(S) = \dim \operatorname{im}(S) = \dim S(V) = b_1 + \dots + b_\ell.$$

Daher folgt von Theorem 4.2.9, dass

$$\begin{aligned} \dim V &= \dim \ker(S) + \operatorname{rk}(S) \\ &= (\ell + m) + (b_1 + \dots + b_\ell) \\ &= m + (b_1 + 1) + \dots + (b_\ell + 1). \end{aligned}$$

Doch dies ist genau die Anzahl der Vektoren in $\mathcal{B}$; es folgt daher von Satz 3.3.23, dass $\mathcal{B}$ eine Basis ist.

Aufgrund der Induktionshypothese sind $b_1, \dots, b_\ell$ eindeutig bestimmt. Ebenfalls ist

$$m = \dim(\ker(S)) - \ell$$

eindeutig bestimmt. Doch diese Zahlen bestimmen eindeutig die n_i in dem Theorem.⁶² □

Korollar 12.4.4. Es sei $T : V \rightarrow V$ eine lineare Abbildung und λ ein Eigenwert mit dazugehörigem verallgemeinerten Eigenraum $U = \tilde{E}_\lambda$. Dann hat U eine Basis, in der die Abbildungsmatrix von der Beschränkung $T|_U$ in Jordanscher Normalenform ist.

Proof. Es sei $S = (T - \lambda)\mathbf{1}|_U$. Dann ist $S : U \rightarrow U$ nilpotent, und es gibt daher eine Basis $\mathcal{B}$ von U , so dass $[S]_{\mathcal{B}}^{\mathcal{B}}$ in Jordanscher Normalenform ist:

$$[S]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} J_{n_1}(0) & & & \\ & J_{n_2}(0) & & \\ & & J_{n_3}(0) & \\ & & \vdots & \\ & & & J_{n_k}(0) \end{pmatrix}.$$

Da die Abbildungsmatrix der Identität in jeder Basis die Identitätsmatrix ist, folgt, dass

$$[T]_{\mathcal{B}}^{\mathcal{B}} = [S]_{\mathcal{B}}^{\mathcal{B}} + \lambda \mathbf{1}_n = \begin{pmatrix} J_{n_1}(\lambda) & & & \\ & J_{n_2}(\lambda) & & \\ & & J_{n_3}(\lambda) & \\ & & \vdots & \\ & & & J_{n_k}(\lambda) \end{pmatrix}.$$

□

Das beendet den Beweis von Theorem 12.1.2, eines der wichtigsten Resultate der linearen Algebra.

⁶¹Wenn sie die Jordanketten andersrum anordnen, d.h. $u_1, \dots, S^{b_1}u_1$ etc., dann erhalten Sie Einsen unterhalb der Hauptdiagonalen anstatt darüber.

⁶²Sie könnten auch mit einem anderen Unterraum $U \leq V$ arbeiten, um den Induktionsschritt durchzuführen. Dann ist es aber wahrscheinlich schwieriger, einen Basis von V zu konstruieren, die aus maximalen Jordanketten besteht.

Bemerkung 12.4.5. Theorem 12.1.2 sagt uns, dass zwei Matrizen $A, B \in M_{n \times n}(K)$ ueber einem algebraisch abgeschlossenen Koerper K genau dann aehnlich sind, wenn sie (von einer Vertauschung der Jordanbloecke abgesehen) die gleiche Jordansche Normalenform haben. Insbesondere ist die Anzahl der Aequivalenzklassen aehnlicher $(n \times n)$ -Matrizen mit einem einzigen Eigenwert λ die Anzahl aller Partitionen von n .

Zum Zusammenhang zwischen verschiedenen Jrodanbasen haben wir folgendes Resultat von Lauren Früchtl.

Theorem 12.4.6. Sei $J := J_n(\lambda) \in M_{n \times n}(\mathbb{K})$ ein Jordanblock und $B \in GL_n(\mathbb{K})$. Genau dann wenn B die unten angegebene Form annimmt, gilt $B^{-1}JB = J$

$$\begin{bmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_{n-1} & \lambda_n \\ 0 & \lambda_1 & \ddots & \lambda_{n-2} & \lambda_{n-1} \\ 0 & 0 & \lambda_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ 0 & \dots & 0 & 0 & \lambda_1 \end{bmatrix}$$

mit $\lambda_i \in \mathbb{K}$

Proof. Behauptung: Thm. $\iff JB = BJ$.

Beweis der Behauptung: " $\Leftarrow$ " Da $B^{-1}JB = J$ rechts Multiplikation mit $B \implies JB = BJ$.

" $\Rightarrow$ " links Multiplikation mit B^{-1} .

Wir betrachten nun $(BJ)_{ik} = (JB)_{ik}$ und versuchen Bedingungen für B zu finden. Wir wissen durch die Definition von Jordanblöcken das folgende:

$$(57) \quad J_{jk} = \begin{cases} \lambda & \text{if } j = k \\ 1 & \text{if } j + 1 = k \\ 0 & \text{sonst} \end{cases}$$

Sei $k \neq 1$ und $i \neq n$

$$(58) \quad (BJ)_{ik} = \sum_{j=1}^n B_{ij} J_{jk} = \lambda B_{ik} + B_{i(k-1)}$$

$$(59) \quad (JB)_{ik} = \sum_{j=1}^n J_{ij} B_{jk} = \lambda B_{ik} + B_{(i+1)k}$$

Also muss $B_{i(k-1)} = B_{(i+1)k}$ gelten. Demnach sind die Werte auf den "diagonalen" gleich. Wir müssen noch die Fälle $i = n$ und $k = 1$ betrachten (nach analoger Argumentation):

$$(60) \quad (BJ)_{i1} = \lambda B_{i1} \text{ und } (JB)_{nk} = \lambda B_{nk}$$

Aus diesen Fällen folgt:

$$(61) \quad (BJ)_{11} = \lambda B_{11} \text{ aber } (JB)_{11} = \lambda B_{11} + B_{21} \implies B_{21} = 0$$

Nach Ausführung dieser Argumentation folgt, dass B eine obere Dreiecksmatrix ist. D.h. $B^{-1}JB = J$ gilt genau dann, wenn B die oben genannte Form annimmt. $\square$

Korollar 12.4.7. Sei $J \in M_{n \times n}(\mathbb{K})$ in JNF. Wir finden für jeden Jordanblock eine Matrix B_i wie im Theorem 0.1. Die folgende Matrix erfüllt $B^{-1}JB = J$:

$$(62) \quad \begin{bmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & \dots & 0 & B_k \end{bmatrix}$$

Proof. Klar, da die Blockmatrizen unabhängig voneinander sind und jeweils die Bedingung erfüllen. $\square$

12.5. Berechnung der Jordanschen Normalenform.

Beispiel 12.5.1. Es sei $A \in M_{4 \times 4}(\mathbb{C})$,

$$A = \begin{pmatrix} 1 & 1 & 6 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 2 \end{pmatrix}.$$

Wir berechnen das charakteristische Polynom:

$$\chi_A(x) = (1 - \lambda)^3(2 - \lambda),$$

das heisst, die Eigenwerte sind $\lambda_1 = 1$ und $\lambda_2 = 2$ mit algebraischen Vielfachheiten $a_{\lambda_1} = 3$ und $a_{\lambda_2} = 1$. Es folgt unmittelbar, dass $g_{\lambda_2} = 1$. Um die geometrische Vielfachheit von λ_1 zu bestimmen, berechnen wir den dazugehoerigen Eigenraum:

$$E_{\lambda_1} = \langle e_1 \rangle,$$

das heisst $g_{\lambda_1} = 1$. Es folgt daher von Theorem 12.2.3 dass es fuer λ_1 und λ_2 jeweils nur einen Jordanschen Block gibt, deren Laenge n_i durch die jeweilige algebraische Vielfachheit beschaenkt ist. Daher gilt $n_1 = 3$ und $n_2 = 1$, und die Jordansche Normalenform von A ist gegeben durch

$$J(A) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Beispiel 12.5.2. Es sei $B \in M_{4 \times 4}(\mathbb{C})$,

$$B = \begin{pmatrix} 5 & -1 & 0 & 0 \\ 9 & -1 & 0 & 0 \\ 0 & 0 & 7 & -2 \\ 0 & 0 & 12 & -3 \end{pmatrix}.$$

Dann gilt

$$\chi_B(x) = (x - 2)^2(x - 3)(x - 1),$$

das heisst, die Eigenwerte sind $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$ mit algebraischen Vielfachheiten $a_{\lambda_1} = 1$, $a_{\lambda_2} = 2$, $a_{\lambda_3} = 1$. Die geometrischen Vielfachheiten von λ_1 und λ_3 sind gleich den algebraischen Vielfachheiten:

$$g_{\lambda_i} = a_{\lambda_i} \quad \text{fuer } i = 1, 3,$$

und $g_{\lambda_2} \in \{1, 2\}$. Wir bestimmen den zu λ_2 gehoerigen Eigenraum:

$$E_{\lambda_2} = \left\langle \begin{pmatrix} 1 \\ 3 \\ 0 \\ 0 \end{pmatrix} \right\rangle,$$

das heisst $g_{\lambda_2} = 1$ und es gibt laut Theorem 12.2.3 genau einen Jordanblock mit Eigenwert λ_2 . Die Jordansche Normalenform von B ist daher

$$J(B) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Wir gibt daher eine Matrix $C \in \text{GL}_4(\mathbb{C})$, so dass

$$J(B) = C^{-1}BC.$$

Um diese Matrix zu finden, bestimmen wir zunaechst die Eigenraeume von λ_1 und λ_3 :

$$E_{\lambda_1} = \left\langle \begin{pmatrix} 0 \\ 0 \\ -1 \\ 3 \end{pmatrix} \right\rangle \quad \text{und} \quad E_{\lambda_3} = \left\langle \begin{pmatrix} 0 \\ 0 \\ -1 \\ 2 \end{pmatrix} \right\rangle.$$

Jeder dieser Vektoren bildet eine Jordankette der Laenge 1. Wir muessen nun eine Jordankette $\{v, (B - 2\mathbf{1}_4)v\}$ in $\tilde{E}_{\lambda_2}$ finden, so dass

$$(B - 2\mathbf{1}_4)v = \begin{pmatrix} 1 \\ 3 \\ 0 \\ 0 \end{pmatrix}.$$

Wir loesen dieses lineare Gleichungssystem und erhalten

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

Dann ist

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

eine Jordanbasis der Abbildung T_B , das heisst, fuer

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

gilt $C^{-1}BC = J(B)$.

Beispiel 12.5.3. Es sei $D \in M_{4 \times 4}(\mathbb{C})$,

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

Dann gilt

$$\chi_D(x) = (1 - x)^4,$$

das heisst, $\lambda = 1$ ist der einzige Eigenwert von D , und es gilt $a_\lambda = 4$. Wir bestimmen den Eigenraum:

$$E_\lambda = \langle e_1, e_1 + e_2 + e_3 + e_4 \rangle,$$

das heisst $g_\lambda = 2$ und es gibt zwei Jordanbloেকে, entweder der Laengen 1 und 3, oder beide der Laenge 2. Im ersten Fall gibt es einen Eigenvektor, der nicht zu einer Jordankette der Laenge 2 gehoert, und im zweiten Fall gehoert jeder Eigenvektor zu einer Jordankette der Laenge 2. Um zu entscheiden, welche der beiden Moeglichkeiten zutrifft, rechnen wir nach, ob der Vektor e_1 Teil einer Jordankette der Laenge 2 ist: in diesem Fall gibt es $v \in \mathbb{C}^4$, $v \neq 0_V$, so dass

$$(D - \mathbf{1}_4)v = e_1.$$

Eine einfache Rechnung zeigt, dass diese Gleichung keine Loesung hat. Daher gibt es zwei Jordanbloেকে der Laengen 1 und 3, und $y = e_1 + e_2 + e_3 + e_4$ gehoert zu einer Jordankette der Laenge 3. Wir rechnen nach:

$$w = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} \text{ ist eine Loesung der Gleichung } (D - \mathbf{1}_4)w = y,$$

$$u = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \text{ ist eine Loesung der Gleichung } (D - \mathbf{1}_4)u = w,$$

das heisst $\{u, (D - \mathbf{1}_4)u, (D - \mathbf{1}_4)^2 u\}$ ist eine Jordankette der Laenge 3, und eine Jordanbasis ist gegeben durch $e_1, u, (D - \mathbf{1}_4)u, (D - \mathbf{1}_4)^2 u$.

Alternativ koennen wir auch einfach die Potenzen der nilpotenten Abbildung $D - \mathbf{1}_4$ berechnen: wir sehen, dass $(D - \mathbf{1}_4)^2 \neq 0_{4 \times 4}$, aber $(D - \mathbf{1}_4)^3 = 0_{4 \times 4}$. Daher hat der grosste Jordanblock die Laenge 3.

Beispiel 12.5.4. Es sei

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

Das charakteristische Polynom ist gegeben durch $\chi_A(x) = -x^5$, das heisst, $\lambda = 0$ ist der einzige Eigenwert und A ist nilpotent. Wir berechnen den Eigenraum:

$$E_\lambda = \langle e_2 - e_4, e_1 - e_3 \rangle;$$

es folgt daher von Theorem 12.2.3, dass A zwei Jordanblöcke hat. Die möglichen Längen sind 2 und 3, oder 1 und 4. Eine einfache Rechnung zeigt, dass $A^3 = 0_{5 \times 5}$, daher ist $\begin{pmatrix} J_2(0) & 0 \\ 0 & J_3(0) \end{pmatrix}$ die JNF von A .

Beispiel 12.5.5. Es seien $A, B \in M_{3 \times 3}(K)$ mit dem gleichen charakteristischen Polynom $\chi(x)$ und dem gleichen minimalen Polynom $m(x)$.

Behauptung. A und B sind ähnlich.

Beweis der Behauptung. Wir analysieren die verschiedenen Fälle, die auftreten können:

- Wenn

$$\chi(x) = (\lambda_1 - x)(\lambda_2 - x)(\lambda_3 - x)$$

für drei verschiedene Eigenwerte $\lambda_1, \lambda_2, \lambda_3$, dann haben A und B beide die Jordansche Normalenform

$$J = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

- Nimm an, dass

$$\chi(x) = (\lambda_1 - x)^2(\lambda_2 - x)$$

mit zwei verschiedenen Eigenwerten λ_1, λ_2 . Dann gilt

$$m(x) = (x - \lambda_1)^j(x - \lambda_2),$$

wobei $j \in \{1, 2\}$. Wenn $j = 1$, dann haben laut Theorem 12.2.3 alle Jordanblöcke die Länge 1, das heisst, A und B sind ähnlich zu

$$J = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$

Wenn $j = 2$, dann gehört λ_1 zu einem Jordanblock der Länge 2, d.h. A und B sind ähnlich zu

$$J = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$

- Wenn

$$\chi(x) = (\lambda - x)^3,$$

dann gilt

$$m(x) = (\lambda - x)^j \quad \text{für } j \in \{1, 2, 3\},$$

und es folgt von Theorem 12.2.3, dass

$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \quad \text{oder} \quad J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \quad \text{oder} \quad J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix},$$

abhängig davon, ob $j = 1$ oder $j = 2$ oder $j = 3$.

Bemerkung 12.5.6. Für (4×4) Matrizen ist die Verallgemeinerung des vorheriges Beispiels falsch: Die Matrizen

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

haben das gleiche charakteristische und minimale Polynom, aber sie sind nicht ähnlich.

13.1. Normierte RAEUME.

Definition 13.1.1. Es sei $K = \mathbb{R}$ oder $\mathbb{C}$, und es sei V ein K -Vektorraum. Eine Norm auf V ist eine Funktion

$$\| \cdot \| : V \rightarrow \mathbb{R}_{\geq 0}$$

mit folgenden Eigenschaften:

- $\|u + v\| \leq \|u\| + \|v\| \quad \forall u, v \in V$ (Dreiecksungleichung);
- $\|\alpha v\| = |\alpha| \cdot \|v\|$ fuer alle $\alpha \in K, v \in V$;
- wenn $\|v\| = 0$, dann gilt $v = 0_V$ (Nicht-Degeneriertheit).

Beispiele 13.1.2. Es sei $V = \mathbb{R}^2$ und $v = \begin{pmatrix} a \\ b \end{pmatrix} \in V$.

(1) Die Funktion

$$\|v\| = \sqrt{a^2 + b^2}$$

ist eine Norm.

(2) Die Funktion

$$\|v\|' = \max\{|a|, |b|\}$$

ist eine Norm.

(3) Die Funktion

$$\|v\|'' = |a| + |b|$$

ist ebenfalls eine Norm.

(4) Andererseits ist die Funktion $v \mapsto \min\{|a|, |b|\}$ keine Norm.

Bemerkung 13.1.3. Wir werden im naechsten Kapitel sehen, dass sich die Norm $\| \cdot \|$ in Beispiel 13.1.2 von den Normen $\| \cdot \|'$ und $\| \cdot \|''$ unterscheidet: sie entsteht aus einem inneren Produkt.

Übung 13.1.4. Es sei $p \geq 1$. Fuer $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ definiere

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}.$$

Dann ist $\| \cdot \|_p$ eine Norm; sie heisst die p -Norm.

Beispiel 13.1.5. Es sei V der Vektorraum aller stetigen Funktionen $f : [0, 1] \rightarrow \mathbb{R}$. Definiere

$$\|f\|_{\max} = \max\{|f(x)| : x \in [0, 1]\}.$$

Dann ist $\| \cdot \|_{\max}$ eine Norm: sie heisst die *Maximumsnorm*.

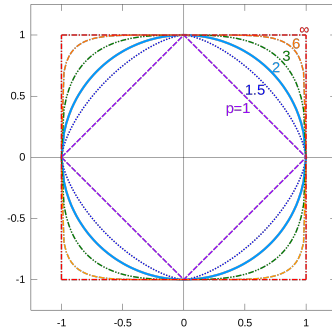
Definition 13.1.6.

- (1) Ein Vektor $v \in V$ ist ein Einheitsvektor wenn $\|v\| = 1$.
- (2) Die Distanz zwischen zwei Vektoren $v, w \in V$ ist gegeben durch

$$d(v, w) = \|v - w\|.$$

Bemerkung 13.1.7. Es sei $v \in V, v \neq 0_V$. Dann ist $\frac{1}{\|v\|}v$ ein Einheitsvektor; wir nennen ihn die Normalisierung von v .

Bemerkung 13.1.8. Die Menge der Einheitsvektoren in einem Vektorraum kann je nach Norm voellig verschieden aussehen. Folgendes Diagramm zeigt die Menge der Einheitsvektoren in $\mathbb{R}^2$ bezueglich der p -Norm fuer verschiedene p .



13.2. Innere Produkte. Wir konzentrieren uns nun auf Vektorräume über $\mathbb{R}$ oder $\mathbb{C}$ mit einer besonderen geometrischen Struktur: einem *inneren Produkt*. Wir erhalten

- ⎧ Euklidische Räume über $\mathbb{R}$
- ⎧ Hermitesche Räume über $\mathbb{C}$

Definition 13.2.1. Es sei V ein $\mathbb{R}$ -Vektorraum. Ein inneres Produkt ist eine Funktion

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$$

mit folgenden Eigenschaften:

(1) Linearität in der ersten Variablen:

$$\begin{aligned} \langle v_1 + v_2, w \rangle &= \langle v_1, w \rangle + \langle v_2, w \rangle \quad \forall v_1, v_2, w \in V, \\ \langle \alpha v, w \rangle &= \alpha \langle v, w \rangle \quad \forall \alpha \in \mathbb{R}, \forall v, w \in V. \end{aligned}$$

(2) Linearität in der zweiten Variablen:⁶³

$$\begin{aligned} \langle v, w_1 + w_2 \rangle &= \langle v, w_1 \rangle + \langle v, w_2 \rangle \quad \forall v, w_1, w_2 \in V, \\ \langle v, \alpha w \rangle &= \alpha \langle v, w \rangle \quad \forall \alpha \in \mathbb{R}, \forall v, w \in V. \end{aligned}$$

(3) Symmetrie: $\langle v, w \rangle = \langle w, v \rangle \quad \forall v, w \in V$;

(4) Positivität: $\langle v, v \rangle > 0 \quad \forall v \in V, v \neq 0_V$.

Das Paar $(V, \langle \cdot, \cdot \rangle)$ ist ein euklidischer Raum.

Beispiel 13.2.2. Es sei $V = \mathbb{R}^n$. Für $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ und $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ definiere

$$\langle u, v \rangle = \sum_{i=1}^n u_i v_i.$$

Dann ist $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Raum; das innere Produkt ist das Standardprodukt auf $\mathbb{R}^n$.

Lecture 39

Beispiel 13.2.3. Es sei $V = \mathbb{R}^2$. Für $u_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$, $u_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ definiere

$$\langle u_1, u_2 \rangle' = 2x_1x_2 - x_1y_2 - y_1x_2 + y_1y_2.$$

Dann ist $\langle \cdot, \cdot \rangle'$ ein inneres Produkt: Eigenschaften (1) - (3) sind klar; für Eigenschaft (4), beachte, dass für $u = \begin{pmatrix} x \\ y \end{pmatrix}$

$$\langle u, u \rangle' = x^2 + (x - y)^2 \geq 0,$$

mit Gleichheit genau dann, wenn $u = 0$.

Bemerkung 13.2.4. Wir können das innere Produkt $\langle u, v \rangle$ aus Beispiel 13.2.2 als Matrixmultiplikation betrachten: es gilt $\langle u, v \rangle = u^t v$, d.h. es ist die Matrixmultiplikation der $(1 \times n)$ -Matrix u^t und der $(n \times 1)$ -Matrix v .

Auch das innere Produkt (13.2.3) lässt sich als Matrixmultiplikation schreiben:

$$\langle u_1, u_2 \rangle' = u_1^t A u_2 \quad \text{mit } A = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

⁶³Eine Abbildung $V \times V \rightarrow \mathbb{R}$, die linear in jeder der Variablen ist, heißt *bilinear*.

Es stellt sich die Frage, ob wir nicht andere innere Produkte konstruieren koennen, indem wir A durch eine andere Matrix in $M_{n \times n}(\mathbb{R})$ ersetzen. Fuer welche anderen Matrizen erhalten wir ebenfalls ein inneres Produkt? Wir kommen auf diese Frage im naechsten Kapitel zurueck.

Beispiel 13.2.5. Es seien $a, b \in \mathbb{R}$, $a < b$, und es sei $C[a, b]$ der Vektorraum der stetigen Funktionen $f : [a, b] \rightarrow \mathbb{R}$. Wir definieren

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx.$$

Die erfuehlt die Axiome des inneren Produkts:

- Linearitaet in jedem der Argumente und Symmetrie sind klar.
- Es sei nun $f \in V$, $f \neq 0$. Dann gibt es $y \in (a, b)$ und $\delta > 0$, so dass $[y - \delta, y + \delta] \subset [a, b]$ und $f(x) \neq 0 \forall x \in [y - \delta, y + \delta]$. Dann gilt

$$\langle f, f \rangle = \int_a^b f(x)^2 dx \geq \int_{y-\delta}^{y+\delta} f(x)^2 dx \geq 2\delta \min_{x \in [y-\delta, y+\delta]} f(x)^2 > 0.$$

Definition 13.2.6. Es sei V ein $\mathbb{C}$ -Vektorraum. Ein Hermetisches Produkt (oder inneres Produkt) ist eine Funktion

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$$

mit folgenden Eigenschaften:

(1) *Linearitaet in der erste Variablen:*

$$\begin{aligned} \langle v_1 + v_2, w \rangle &= \langle v_1, w \rangle + \langle v_2, w \rangle \quad \forall v_1, v_2, w \in V, \\ \langle \alpha v, w \rangle &= \alpha \langle v, w \rangle \quad \forall \alpha \in \mathbb{C}, \forall v, w \in V. \end{aligned}$$

(2) *Sesquilinearitaet in der zweiten Variablen:*

$$\begin{aligned} \langle v, w_1 + w_2 \rangle &= \langle v, w_1 \rangle + \langle v, w_2 \rangle \quad \forall v, w_1, w_2 \in V, \\ \langle v, \alpha w \rangle &= \overline{\alpha} \langle v, w \rangle \quad \forall \alpha \in \mathbb{C}, \forall v, w \in V. \end{aligned}$$

(3) *Hermitesche Eigenschaft:* $\langle v, w \rangle = \overline{\langle w, v \rangle} \quad \forall v, w \in V$;

(4) *Positivitaet:* $\langle v, v \rangle > 0 \quad \forall v \in V, v \neq 0_V$.

Das Paar $(V, \langle \cdot, \cdot \rangle)$ ist ein Hermitescher oder unitaerer Raum.

Beispiel 13.2.7. Es sei $V = \mathbb{C}^n$. Fuer $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ und $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ definiere

$$\langle u, v \rangle = \sum_{i=1}^n u_i \overline{v_i}.$$

Dann ist $(V, \langle \cdot, \cdot \rangle)$ ein Hermitescher Raum, und das innere Produkt ist das Standard IP auf $\mathbb{C}^n$.

Bemerkung 13.2.8. Wir koennen auch das Standard IP auf $\mathbb{C}^n$ als Matrixmultiplikation betrachten: es gilt $\langle u, v \rangle = u^t \bar{v}$, d.h. es ist die Matrixmultiplikation der $(1 \times n)$ -Matrix u^t und der $(n \times 1)$ -Matrix $\bar{v}$.

Wieder koennen wir die Frage stellen: gibt es Matrizen $A \in M_{n \times n}(\mathbb{C})$, so dass $(u, v) \mapsto u^t A \bar{v}$ ein Hermitesches Produkt ist?

Lemma 13.2.9. Es sei V ein Euklidischer oder Hermitescher Raum. Dann gilt

- (1) $\langle 0_V, v \rangle = \langle v, 0_V \rangle = 0 \quad \forall v \in V$;
- (2) Es sei $w \in V$. Wenn $\langle v, w \rangle = 0$ fuer alle $v \in V$, dann gilt $w = 0_V$;
- (3) Es seien $w_1, w_2 \in V$. Wenn $\langle v, w_1 \rangle = \langle v, w_2 \rangle$ fuer alle $v \in V$, dann gilt $w_1 = w_2$.

Proof.

- (1) $\langle 0_V, v \rangle = 0 \cdot \langle 0_V, w \rangle = 0$.
- (2) Wenn $w \neq 0_V$, dann gilt $\langle w, w \rangle > 0$.
- (3) Folgt von (2) fuer $w = w_1 - w_2$.

□

Satz 13.2.10. Es sei $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer oder Hermitescher Raum. F ur alle $v \in V$, definiere

$$\|v\| = \sqrt{\langle v, v \rangle} \in \mathbb{R}_{\geq 0}.$$

Dann ist $\|\cdot\|$ eine Norm.

Um diesen Satz zu beweisen, brauchen wir folgendes Lemma:

Lemma 13.2.11 (Cauchy-Schwartz Ungleichung). Es sei $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer oder Hermitescher Raum. Dann gilt

$$|\langle u, v \rangle| \leq \|u\| \cdot \|v\| \quad \forall u, v \in V$$

mit Gleichheit genau dann, wenn u und v linear abh angig sind.

Proof. Wenn $u = 0_V$, dann h elt die Ungleichung automatisch. Nimm daher an, dass $u \neq 0_V$, und definiere $\lambda = \frac{\langle v, u \rangle}{\|u\|^2}$ und

$$w = v - \lambda \cdot u.$$

Beachte, dass

$$\langle w, u \rangle = \langle v, u \rangle - \lambda \langle u, u \rangle = 0.$$

Nun gilt

$$(63) \quad 0 \leq \|w\|^2$$

$$(64) \quad \begin{aligned} &= \langle v - \lambda u, v - \lambda u \rangle \\ &= \|v\|^2 - \lambda \bar{\lambda} \|u\|^2 - \bar{\lambda} \lambda \|u\|^2 + \lambda \bar{\lambda} \|u\|^2 \\ &= \|v\|^2 - |\lambda|^2 \cdot \|u\|^2 \\ &= \|v\|^2 - \frac{|\langle u, v \rangle|^2}{\|u\|^2} \\ \Rightarrow \quad &|\langle u, v \rangle| \leq \|u\| \cdot \|v\|. \end{aligned}$$

In (64) erhalten wir Gleichheit genau dann, wenn $w = 0_V$, d.h. wenn $v = \frac{\langle v, u \rangle}{\|u\|^2} \cdot u$.

Andererseits, wenn $v = \lambda \cdot u$ f ur $\lambda \in K$ und $u \neq 0_V$, dann gilt $\lambda = \frac{\langle v, u \rangle}{\|u\|^2}$. Quod erat demonstrandum. $\square$

von Satz 13.2.10. Wir beweisen den Satz f ur Hermitesche R aume; der Beweis f ur Euklidische R aume ist sehr  hnlich. Wir  uberpr ufen die Axiome:

(1) Dreiecksungleichung: es seien $u, v \in V$. Dann gilt

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \|u\|^2 + 2\Re\langle u, v \rangle + \|v\|^2. \end{aligned}$$

Beachte nun, dass $\Re(z) \leq |z|$ f ur alle $z \in \mathbb{C}$. Angewandt auf $\Re\langle u, v \rangle$ erhalten wir

$$\Re\langle u, v \rangle \leq |\Re\langle u, v \rangle| \leq |\langle u, v \rangle|$$

und daher

$$\begin{aligned} \|u + v\|^2 &= \|u\|^2 + 2\Re\langle u, v \rangle + \|v\|^2 \\ &\leq \|u\|^2 + 2|\langle u, v \rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\| \cdot \|v\| + \|v\|^2 \\ &= (\|u + v\|)^2 \end{aligned}$$

und daher $\|u + v\| \leq \|u\| + \|v\|$.

(2) Die anderen beiden Axiome der Norm sind eine  bung. $\square$

Lemma 13.2.12. Es sei $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer Raum. Dann gilt

$$\langle u, v \rangle = \frac{1}{2} (\|u + v\|^2 - \|u\|^2 - \|v\|^2)$$

f ur alle $u, v \in V$. Mit anderen Worten, das innere Produkt $\langle \cdot, \cdot \rangle$ ist durch die Norm $\|\cdot\|$ vollst andig bestimmt. $\square$

Proof. Explizite Rechnung. $\square$

Uebung. Formulieren und beweisen Sie das Analog von Lemma 13.2.12 fuer Hermitesche Raeume.

Bemerkung 13.2.13. *Nicht jede Norm kann ueber ein inneres Produkt definiert werden! Beispiele da fuer sind die Normen $\|\sim\|'$ und $\|\sim\|''$ aus Beispiel 13.1.2. Koennen Sie das beweisen?*

13.3. Konstruktion innerer Produkte. In diesem Kapitel wollen wir zeigen, dass sich innere Produkte mit Hilfe bestimmter Matrizen definieren lassen.

Beispiel 13.3.1. Es sei $A \in M_{n \times n}(\mathbb{R})$. Definiere

$$\langle \cdot, \cdot \rangle_A : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad \langle u, v \rangle_A = v^T \cdot A \cdot v.$$

Dann ist $\langle \cdot, \cdot \rangle_A$ linear in jeder der beiden Variablen.

Allerdings ist die Funktion nicht unbedingt symmetrisch: es sei $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ und $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. Dann gilt

$$\langle u, v \rangle_A = u_1(v_1 + v_2) + u_2v_2 \quad \text{aber} \quad \langle v, u \rangle_A = (u_1 + u_2)v_1 + u_2v_2.$$

Definition 13.3.2. *Es sei K ein Koerper. Eine Matrix $A \in M_{n \times n}(K)$ ist symmetrisch, wenn $A = A^t$, d.h. wenn $a_{ij} = a_{ji}$ fuer alle $1 \leq i, j \leq n$.*

Lemma 13.3.3. *Es sei $A \in M_{n \times n}(\mathbb{R})$ symmetrisch. Dann gilt*

$$\langle u, v \rangle_A = \langle v, u \rangle_A \quad \forall u, v \in \mathbb{R}^n.$$

Proof. Es gilt

$$\langle v, u \rangle_A = v^t A u = v^t A^t u = (u^t A v)^t = u^t A v = \langle u, v \rangle_A.$$

□

Allerdings erhalten wir immer noch nicht unbedingt ein Skalarprodukt:

Beispiel 13.3.4. Es sei $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Dann gilt

$$\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle_A = -1;$$

die Bilinearform ist nicht positiv.

Definition 13.3.5. *Eine symmetrische Matrix $A \in M_{n \times n}(\mathbb{R})$ ist positiv definit, wenn $v^t A v > 0$ fuer alle $v \in \mathbb{R}^n$, $v \neq 0$.*

Uebung: Bestimmen Sie, wann eine Matrix $B \in M_{2 \times 2}(\mathbb{R})$ positiv definit ist.

Lecture 40

Beispiele 13.3.6. (1) Die Matrix $B = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ ist genau dann positiv definit, wenn $a > 0$ und $\det(B) > 0$.

(2) Es seien $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, $\alpha_i > 0$ fuer alle i . Dann ist die diagonale Matrix

$$A = \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ 0 & 0 & \dots & \alpha_n \end{pmatrix}$$

positiv definit.

Satz 13.3.7. *Es sei $A \in M_{n \times n}(\mathbb{R})$. Dann ist $\langle \cdot, \cdot \rangle_A$ ist genau dann ein inneres Produkt, wenn A positiv definit ist.*

Bemerkung 13.3.8. *Tatsaechlich gibt es eine 1:1-Korrespondenz zwischen positiv definiten $(n \times n)$ -Matrizen und inneren Produkten auf $\mathbb{R}^n$: wenn $\langle \cdot, \cdot \rangle$ ein inneres Produkt ist, definiere $a_{ij} = \langle e_i, e_j \rangle$, wobei $e_1, \dots, e_n$ die Standardbasis ist. Dann gilt fuer $u, v \in \mathbb{R}^n$*

$$\langle u, v \rangle = u^t A v.$$

Wir werden Satz 13.3.7 spaeter zusammen mit seinem Analog fuer komplexe Matrizen beweisen. Zunaechst muessen wir herausfinden, wie wir ‘positiv definit’ fuer komplexe Matrizen passend verallgemeinern.

Definition 13.3.9. Es sei $B \in M_{n \times n}(\mathbb{C})$. Definiere

$$\langle \cdot, \cdot \rangle_B : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}, \quad \langle u, v \rangle_B = u^t B \bar{v}.$$

Definition 13.3.10. Es sei $B \in M_{n \times n}(\mathbb{C})$.

- (1) Schreibe $\bar{B} = (\bar{b}_{ij})_{ij}$. Die adjungierte Matrix von B ist die Matrix $B^* = (\bar{B})^t$.
- (2) Die Matrix B ist Hermitesch, wenn $B = B^*$, das heisst, wenn $b_{ij} = \bar{b}_{ji} \forall 1 \leq i, j \leq n$.

Beispiel 13.3.11. (1) Die Matrix $A = \begin{pmatrix} -1 & 1+i & 2i \\ 1-i & 0 & 1-2i \\ -2i & 1+2i & 1 \end{pmatrix}$ ist Hermitesch.

(2) Die Matrix $B = \begin{pmatrix} -i & 1+i & 2i \\ 1-i & 0 & 1-2i \\ -2i & 1+2i & 1 \end{pmatrix}$ ist nicht Hermitesch.

(3) Eine Matrix $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2 \times 2}(\mathbb{C})$ ist genau dann Hermitesch, wenn $a, d \in \mathbb{R}$ und $b = \bar{c}$.

Beachte 13.3.12. (1) Wenn B Hermitesch ist, dann erfllt $\langle \cdot, \cdot \rangle_B$ Eigenschaften (1) - (3) des Hermiteschen Produktes.

(2) Es sei $B \in M_{n \times n}(\mathbb{C})$ Hermitesch, und es sei $v \in \mathbb{C}^n$. Dann gilt

$$\overline{\langle v, v \rangle_B} = \bar{v}^t \bar{B} v = (\bar{v}^t \bar{B}^t v)^t = v^t \bar{B}^t \bar{v} = \langle v, v \rangle_B,$$

das heisst $v^t B v \in \mathbb{R}$ fr alle $v \in \mathbb{C}^n$.

Mit Hilfe dieser Beobachtung koennen wir eine positiv definite Hermitesche Matrix definieren.

Definition 13.3.13. Eine Hermitesche Matrix $B \in M_{n \times n}(\mathbb{C})$ ist positiv definit, wenn

$$v^t B \bar{v} > 0 \quad \forall v \in \mathbb{C}^n, v \neq 0.$$

Beispiel 13.3.14. Die Hermitesche Matrix $B = \begin{pmatrix} 1 & b \\ \bar{b} & a \end{pmatrix} \in M_{2 \times 2}(\mathbb{C})$ ist genau dann positiv definite, wenn $a \in \mathbb{R}_{>0}$ und $\det(B) \in \mathbb{R}_{>0}$.

Folgender Satz ist die Verallgemeinerung von Satz 13.3.7 auf den komplexen Fall:

Satz 13.3.15. Es sei $B \in M_{n \times n}(\mathbb{C})$. Dann ist $\langle \cdot, \cdot \rangle_B$ genau dann ein inneres Produkt, wenn B positiv definit ist.⁶⁴

Proof. Wir beweisen den komplexen Fall; der reelle Fall ist aehnlich. Wenn B Hermitesch positiv definit ist, dann kann man einfach ueberpruefen, dass $\langle \cdot, \cdot \rangle_B$ die Axiome eines inneren Produktes erfllt.

Nimm nun an, dass $\langle \cdot, \cdot \rangle_B$ ein inneres Produkt ist. Wir zeigen zunaechst, dass B Hermitesch ist. Da $\langle \cdot, \cdot \rangle_B$ ein inneres Produkt ist, besitzt es die Hermitesche Eigenschaft: $\langle u, v \rangle_B = \overline{\langle v, u \rangle_B}$. Es sei $e_1, \dots, e_n$ die Standardbasis von $\mathbb{C}^n$. Dann gilt fr $1 \leq i, j \leq n$ gilt

$$b_{ij} = e_i^t B e_j = \langle e_i, e_j \rangle_B = \overline{\langle e_j, e_i \rangle_B} = \bar{e}_j^t \bar{B} \bar{e}_i = e_j^t \bar{B} e_i = \bar{b}_{ji},$$

das heisst, B ist Hermitesch.

Weiterhin ist die Bedingung, dass $\langle v, v \rangle_B > 0 \forall v \in \mathbb{C}^n, v \neq 0$, genau die Bedingung, dass $v^t B \bar{v} > 0 \forall v \in \mathbb{C}^n, v \neq 0$, das heisst, dass B positiv definit ist. $\square$

Wir koennen also die Vektorraeume $\mathbb{R}^n$ und $\mathbb{C}^n$ mit vielen verschiedenen inneren Produkten ausstatten; die Geometrie der Raume (z.B. die Frage, ob zwei Vektoren orthogonal sind oder nicht) haengt von der Wahl eines inneren Produktes ab.

13.4. Orthogonalitaet. Mit Hilfe eines inneren Produktes koennen wir definieren, wann zwei Vektoren orthogonal sind:

Definition 13.4.1.

- (1) Zwei Vektoren v, w sind orthogonal wenn $\langle v, w \rangle = 0$; wir schreiben $v \perp w$.
- (2) Eine Untermenge $S \subset V$ ist ein orthogonales System wenn $u \perp v$ fr alle $u, v \in S$.
- (3) Ein orthogonales System $S \subset V$ ist orthonormal, wenn $\|v\| = 1$ fr alle $v \in S$.

Beispiele 13.4.2.

⁶⁴Auch im komplexen Fall gibt es wieder eine 1:1 Aequivalenz zwischen IPs und Hermiteschen positiv definiten Matrizen: einem IP $\langle \cdot, \cdot \rangle$ ordnen wir die Matrix $B = (\langle e_i, e_j \rangle)$ zu. Dann gilt $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_B$.

- (1) Die Menge der Standardvektoren im $\mathbb{R}^n$ ist orthonormal bezueglich des Standard IPs.
- (2) In einem Vektorraum mit innerem Produkt ist jeder Vektor orthogonal zum Nullvektor.
- (3) Betrachte $(\mathbb{R}^2, \langle \cdot, \cdot \rangle')$ wie in Beispiel 13.2.3. Dann ist $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ orthogonal zu $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.
- (4) Die Funktionen $f(x), g(x) \in C[-\pi, \pi]$, $f(x) = \sin(nx)$ und $g(x) = \cos(mx)$ sind ein orthogonales System bezueglich des Skalarproduktes von Beispiel 13.2.5.
- (5) Orthogonalitaet haengt von der Wahl des inneren Produktes ab. Es sei $A = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$. Dann gilt

$$\left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle_A = 0,$$

das heisst, die Vektoren $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ sind orthogonal bezueglich dieses inneren Produktes.

Andererseits sind sie selbstverstaendlich nicht orthogonal bezueglich des Standard IPs auf $\mathbb{R}^2$.

Satz 13.4.3 (Satz des Pythagoras). *Es sei $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer oder Hermitescher Raum, und es seien $u, v \in V$, $u \perp v$. Dann gilt*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

Proof. Es gilt

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle \\ &= \|u\|^2 + \|v\|^2. \end{aligned}$$

□

Die Definition der Orthogonalitaet haengt eng mit der *Projektion eines Vektors auf einen anderen Vektor* zusammen:

Definition 13.4.4. *Es sei $(V, \langle \cdot, \cdot \rangle)$ ein Vektorraum mit einem inneren Produkt, und es sei $v \in V$, $v \neq 0_V$. Definiere*

$$\text{proj}_v : V \rightarrow V, \quad u \mapsto \frac{\langle u, v \rangle}{\langle v, v \rangle} \cdot v.$$

Bemerkung 13.4.5. *Ein Vektor $u \in V$ ist genau dann orthogonal zu $v \neq 0_V$, wenn $\text{proj}_v(u) = 0_V$.*

Beispiele 13.4.6.

- (1) Betrachte $\mathbb{R}^n$ mit dem Standard IP, und es sei $v = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$. Dann ist $\text{proj}_{e_i}(v) = a_i \cdot e_i$ fuer alle

$$1 \leq i \leq n.$$

- (2) Es sei $(\mathbb{R}^2, \langle \cdot, \cdot \rangle')$ wie in Beispiel 13.2.3 und $u = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Dann ist die Projektion von u auf den Vektor $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ gegeben durch $\text{proj}_v(u) = v$.

Folgendes Lemma wird wichtig sein, wenn wir orthogonale Basen von Vektorraeumen konstruieren.

Lecture 41

Lemma 13.4.7. *Es sei $(V, \langle \cdot, \cdot \rangle)$ ein Vektorraum mit einem inneren Produkt, und es sei $v \in V$ so dass $v \neq 0_V$. Dann ist fuer alle $u \in V$ der Vektor*

$$w = u - \text{proj}_v(u)$$

orthogonal zu v .

Proof. Uebung.

□

Beispiel 13.4.8. Wir ueberpruefen das in Beispiel 13.4.6 (2):

$$\langle u - \text{proj}_v(u), v \rangle' = \left\langle \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle' = 0.$$

13.5. Gram-Schmidt Orthogonalisierung. Kehren wir zu der Frage von Basen von Vektorraeumen zurueck. Wenn der Vektorraum ein inneres Produkt besitzt, koennen wir besonders schoene Basen konstruieren, naemlich solche, die aus orthogonalen Systemen bestehen. Zur Erinnerung:

- (1) Zwei Vektoren v, w sind *orthogonal* wenn $\langle v, w \rangle = 0$; wir schreiben $v \perp w$.
- (2) Eine Untermenge $S \subset V$ ist ein *orthogonales System* wenn $u \perp v$ fuer alle $u, v \in S$.
- (3) Ein orthogonales System $S \subset V$ ist *orthonormal*, wenn $\|v\| = 1$ fuer alle $v \in S$.

Wir beginnen mit ein paar einfachen Beobachtungen:

Satz 13.5.1. *Es sei $(V, \langle \cdot, \cdot \rangle)$ ein K -Vektorraum mit einem inneren Produkt.*

- (1) *Wenn $S \subseteq V$ ein orthogonales System ist, das nicht den Nullvektor enthaelt, dann ist S linear unabhangig.*
- (2) *Wenn $v_1, \dots, v_n$ ein orthogonales System ist, $v_i \neq 0$ fuer alle $1 \leq i \leq n$, und $v = a_1 v_1 + \dots + a_n v_n$, dann gilt*

$$a_i = \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle} \quad \forall 1 \leq i \leq n.$$

Proof. (1) Nimm an, es gibt $v_1, \dots, v_n \in S$ und $a_1, \dots, a_n \in K$, so dass

$$a_1 v_1 + \dots + a_n v_n = 0_V.$$

Dann gilt

$$\langle a_1 v_1 + \dots + a_n v_n, v_i \rangle = a_i \langle v_i, v_i \rangle = 0.$$

Da $v_i \neq 0_V$ gilt $\langle v_i, v_i \rangle \neq 0$ und daher $a_i = 0$ fuer all $1 \leq i \leq n$.

(2) bung.

□

Bemerkung 13.5.2. *Es folgt von Satz 13.5.1 (1), dass ein orthogonales System in einem n -dimensionalen Vektorraum, das nicht den Nullvektor enthaelt, hoechstens n Elemente haben kann.*

Bemerkung 13.5.3. *In Zukunft schreiben wir $\langle \cdot, \cdot \rangle$ ($\sim$) $\langle \cdot, \cdot \rangle$ fuer die lineare Huelle von Vektoren, damit keine Verwirrung mit dem inneren Produkt entsteht.*

Hat jeder Vektorraum mit innerem Produkt eine orthogonale Basis?

Theorem 13.5.4 (Gram-Schmidt Orthogonalisierungsverfahren). *Es sei $(V, \langle \cdot, \cdot \rangle)$ ein Vektorraum ueber K ($K = \mathbb{R}$ oder $\mathbb{C}$) mit einem inneren Produkt. Es sei $n = \dim_K V$, und es sei $v_1, \dots, v_n$ eine beliebige Basis von V . Es sei $w_1, \dots, w_n$ gegeben durch*

- *es sei $w_1 = v_1$;*
- *fuer $2 \leq j \leq n$ sei*

$$w_j = v_j - \sum_{i=1}^{j-1} \text{proj}_{w_i}(v_j).$$

Dann ist $w_1, \dots, w_n$ eine orthogonale Basis von V .

Proof. Wir zeigen ueber Induktion nach j , dass $w_1, \dots, w_j$ folgende Eigenschaften hat

- (1) $w_i \neq 0_V$ fuer alle $1 \leq i \leq j$;
- (2) $w_1, \dots, w_j$ ist ein orthogonales System;
- (3) $\langle w_1, \dots, w_j \rangle = \langle v_1, \dots, v_j \rangle$.

Da jedes orthogonale System linear unabhangig ist, folgt daraus, dass $w_1, \dots, w_n$ eine orthogonale Basis ist.

Der Fall $j = 1$ ist klar. Nimm an, die Behauptungen (1) - (3) sind wahr fuer $j - 1$. Wenn $w_j = 0_V$, dann gilt

$$v_j = \sum_{i=1}^{j-1} \frac{\langle v_j, w_i \rangle}{\langle w_i, w_i \rangle} \cdot w_i \in \langle w_1, \dots, w_{j-1} \rangle.$$

Doch $\langle w_1, \dots, w_{j-1} \rangle = \langle v_1, \dots, v_{j-1} \rangle$, und daher

$$v_j \in \langle v_1, \dots, v_{j-1} \rangle,$$

was ein Widerspruch zur linearen Unabhangigkeit von $v_1, \dots, v_n$ ergibt. Daher gilt (1).

Um (2) zu ueberpruefen, berechnen wir $\langle w_j, w_i \rangle$ fuer $1 \leq i < j$:

$$\langle w_j, w_k \rangle = \left\langle v_j - \sum_{i=1}^{j-1} \text{proj}_{w_i}(v_j), w_k \right\rangle$$

Doch $\text{proj}_{w_i}(v_j)$ ist ein Vielfaches von w_i und aufgrund der Induktionshypothese $\langle w_i, w_k \rangle = 0$ wenn $i \neq k$. Daher gilt

$$(13.5) = \langle v_j - \text{proj}_{w_k}(v_j), w_k \rangle = 0$$

aufgrund von Lemma 13.4.7.

Nun muessen wir nur noch zeigen, dass $\langle w_1, \dots, w_j \rangle = \langle w_1, \dots, v_j \rangle$.

Aufgrund der Hypothese wissen wir, dass $\langle w_1, \dots, w_{j-1} \rangle = \langle w_1, \dots, v_{j-1} \rangle$.

Doch $w_j = v_j - u$ fuer $u \in \langle w_1, \dots, w_{j-1} \rangle$. Das beendet den Beweis. $\square$

Korollar 13.5.5. Es sei $(V, \langle \cdot, \cdot \rangle)$ ein Vektorraum ueber K ($K = \mathbb{R}$ oder $\mathbb{C}$) mit einem inneren Produkt. Dann besitzt V eine orthonormale Basis. Weiterhin gilt: wenn $v_1, \dots, v_n$ eine orthonormale Basis von V ist und $v = \sum_{i=1}^n a_i v_i$, dann gilt

$$\|v\|^2 = |a_1|^2 + \dots + |a_n|^2.$$

Proof. Es folgt von Theorem 13.5.4, dass V eine orthogonale Basis $w_1, \dots, w_n$ besitzt. Es sei

$$v_i = \frac{1}{\|w_i\|} \cdot w_i \quad \forall 1 \leq i \leq n.$$

Dann ist $\{v_1, \dots, v_n\}$ eine orthonormale Basis.

Wenn $v = \sum_{i=1}^n a_i v_i$, dann gilt

$$\begin{aligned} \|v\|^2 &= \langle v, v \rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i \bar{a}_j \langle v_i, v_j \rangle \\ &= \sum_{i=1}^n |a_i|^2. \end{aligned}$$

$\square$

Beispiel 13.5.6. Es seien $v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ und $v_2 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$, und es sei $V = \langle v_1, v_2 \rangle \leq \mathbb{R}^3$ mit dem Standard IP. Es sei $w_1 = v_1$ und

$$\begin{aligned} w_2 &= v_2 - \text{proj}_{w_1}(v_2) \\ &= \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} - \frac{3}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}. \end{aligned}$$

Dann gilt $\langle w_1, w_2 \rangle = 0$, das heisst, w_1 und w_2 sind eine orthogonale Basis von V .

Beispiel 13.5.7. Betrachte die Basis v_1, v_2, v_3 von $\mathbb{R}^3$ (mit dem Standard IP) mit

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

Dann ist $w_1 = v_1$,

$$\begin{aligned} w_2 &= v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 \\ &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}. \end{aligned}$$

und

$$\begin{aligned} w_3 &= v_3 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 \\ &= \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}. \end{aligned}$$

Dann ist $\frac{w_1}{\|w_1\|}, \frac{w_2}{\|w_2\|}, \frac{w_3}{\|w_3\|}$ eine orthonormale Basis.

Beispiel 13.5.8. Es sei V der Vektorraum aller stetigen Funktionen $f : [-1, 1] \rightarrow \mathbb{R}$ mit dem inneren Produkt aus Beispiel 13.2.5 und $U \leq V$ der Unterraum aller Polynome vom Grad ≤ 3 . Dann ist $1, x, x^2, x^3$ eine Basis von U ; schreibe $f_i = x^i$. Wir benutzen Gram-Schmidt, um eine orthonormale Basis zu konstruieren: es ist $g_0 = f_0$,

$$\begin{aligned} g_1 &= f_1 - \frac{\langle f_1, g_0 \rangle}{\langle g_0, g_0 \rangle} \cdot g_0 \\ &= x - \frac{1}{2} \int_{-1}^1 t \, dt \\ &= x, \\ g_2 &= f_2 - \frac{\langle f_2, g_0 \rangle}{\langle g_0, g_0 \rangle} \cdot g_0 - \frac{\langle f_2, g_1 \rangle}{\langle g_1, g_1 \rangle} \cdot g_1 \\ &= x^2 - \frac{\langle x^2, x \rangle}{2} \cdot 1 - \frac{\langle x^2, x \rangle}{\langle x, x \rangle} \cdot x \\ &= x^2 - \frac{1}{3}, \\ g_3 &= f_3 - \frac{\langle f_3, g_0 \rangle}{\langle g_0, g_0 \rangle} \cdot g_0 - \frac{\langle f_3, g_1 \rangle}{\langle g_1, g_1 \rangle} \cdot g_1 - \frac{\langle f_3, g_2 \rangle}{\langle g_2, g_2 \rangle} \cdot g_2 \\ &= x^3 - \frac{3}{5}x. \end{aligned}$$

Dann ist g_0, g_1, g_2, g_3 eine orthogonale Basis von U . Durch Normalisierung erhalten wir eine orthonormale Basis.

Bemerkung 13.5.9. Es sei V ein $\mathbb{R}$ -Vektorraum und $v_1, v_2, \dots, v_n$ eine Basis. Dann gibt es ein IP $(\langle \cdot, \cdot \rangle)$ on V , bezueglich dessen die gegebene Basis orthonormal ist: wir definieren einfach $\langle v_i, v_j \rangle = \delta_{ij}$. Dieses bestimmt aufgrund von Bilinearitaet (bzw. Sesquilinearitaet) ein IP auf V mit der geforderten Eigenschaft.

Lecture 42

13.6. Das orthogonale Komplement.

Bemerkung 13.6.1. Es sei $K = \mathbb{R}$ oder $\mathbb{C}$ und $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler IP Raum, und es sei $U \leq V$. Es sei $u_1, \dots, u_k$ eine Basis von U , die wir zu einer Basis $u_1, \dots, u_k, u_{k+1}, \dots, u_n$ von V erweitern. Indem wir das GS Verfahren, darauf anwenden, erhalten wir eine orthogonale Basis $w_1, \dots, w_n$ von V , deren Untermenge $w_1, \dots, w_k$ eine Basis von U ist.⁶⁵ Es sei nun $W = \langle w_{k+1}, \dots, w_n \rangle$. Dann hat W folgende Eigenschaften:

- (1) es gilt $V = U \oplus W$, d.h. W ist ein Komplement von U ;
- (2) fuer alle $u \in U, w \in W$ gilt $\langle u, w \rangle = 0$.

⁶⁵Dies folgt von der Konstruktion der orthogonalen Basis - siehe den Beweis von Theorem 13.5.4.

Wie wir sehen werden ist dieses Komplement von U kanonisch: wir nennen es das orthogonale Komplement von U .

Definition 13.6.2. Es sei $V, \langle \cdot, \cdot \rangle$ ein IP Raum und $\emptyset \neq S \subseteq V$ eine Untermenge. Wir definieren das orthogonale Komplement von S als

$$S^\perp = \{v \in V : \langle v, s \rangle = 0 \quad \forall s \in S\}.$$

Wenn $S = \{v\}$, dann schreiben wir $v^\perp$ fuer $\{v\}^\perp$.

Bemerkung 13.6.3. Es gilt $0_V^\perp = V$ und $V^\perp = \{0_V\}$.

Lemma 13.6.4. Es seien $\emptyset \neq S \subseteq V$. Dann gilt

- (1) $S^\perp$ ist ein Unterraum von V ;
- (2) es gilt entweder $S \cap S^\perp = \{0_V\}$ oder $S \cap S^\perp = S$;
- (3) wenn $T \subseteq V$ so dass $S \subseteq T$, dann gilt $T^\perp \subseteq S^\perp$;
- (4) es gilt $\left(\left(\bigcup_{i=1}^n S_i \right)^\perp \right)^\perp = \bigcap_{i=1}^n S_i$;
- (5) es gilt $S \subseteq (S^\perp)^\perp$.

Proof. (1) Ueberpruefe die Axiome.

(2) Nimm an, dass $S \cap S^\perp \neq \{0_V\}$; es sei $v \in S \cap S^\perp$. Dann gilt

$$\langle v, v \rangle = 0 \quad \Rightarrow \quad v = 0_V.$$

(3) Uebung.

(4) Da $S \subseteq \left(\bigcup_{i=1}^n S_i \right)^\perp$, folgt von (3), dass $\left(\bigcup_{i=1}^n S_i \right)^\perp \subseteq S^\perp$. Es sei nun $v \in S^\perp$, das heisst

$$\langle s, v \rangle = 0 \quad \forall s \in S.$$

Wenn $w \in \left(\bigcup_{i=1}^n S_i \right)^\perp$, dann gibt es $s_1, \dots, s_n \in S$ und $\alpha_1, \dots, \alpha_n \in K$ so dass

$$w = \alpha_1 s_1 + \dots + \alpha_n s_n \quad \Rightarrow \quad \langle w, v \rangle = \sum_{i=1}^n \alpha_i \langle s_i, v \rangle = 0,$$

und daher $v \in \left(\bigcup_{i=1}^n S_i \right)^\perp$.

(5) Es sei $s \in S$. Dann gilt (aufgrund der Definition von $S^\perp$), dass $\langle s, v \rangle = 0$ fuer alle $v \in S^\perp$, und daher $s \in (S^\perp)^\perp$. $\square$

Das folgende Theorem zeigt, dass sich Bemerkung 13.6.1 in einem gewissen Sinne auf unendlich dimensionale IP Raeume verallgemeinern laesst.

Theorem 13.6.5. Es sei $(V, \langle \cdot, \cdot \rangle)$ ein IP Raum (moeglicherweise unendlich-dimensional), und es sei U ein endlich-dimensionaler Unterraum. Dann gilt $V = U \oplus U^\perp$.

Proof. Lemma 13.6.4 zeigt, dass $U \cap U^\perp = \{0_V\}$; es ist daher ausreichend zu zeigen, dass $U + U^\perp = V$. Es sei $r = \dim U$, und es sei $\mathcal{B} = \{u_1, \dots, u_r\}$ eine orthonormale Basis von U - diese existiert aufgrund von Theorem 13.5.4. Definiere die Projektion⁶⁶

$$\begin{aligned} \text{pr}_{\mathcal{B}} : V &\rightarrow U, \\ v &\mapsto \langle v, u_1 \rangle u_1 + \dots + \langle v, u_r \rangle u_r, \end{aligned}$$

ds heisst $\text{pr}_{\mathcal{B}} = \text{pr}_{u_1} + \dots + \text{pr}_{u_r}$. Dann gilt

$$v = \text{pr}_{\mathcal{B}}(v) + (v - \text{pr}_{\mathcal{B}}(v)).$$

Behauptung: $(v - \text{pr}_{\mathcal{B}}(v)) \in U^\perp$.

Beweis der Behauptung: es ist ausreichend zu zeigen, dass

$$\langle v - \text{pr}_{\mathcal{B}}(v), u_i \rangle = 0 \quad \forall 1 \leq i \leq r.$$

⁶⁶A priori haengt die Projektion von der Wahl einer ON Basis von U ab. Wir werden spaeter sehen, dass die Projektion kanonisch ist, d.h. unabhaengig von der Wahl einer Basis - wir nennen sie die orthogonale Projektion auf U .

Wir rechnen:

$$\begin{aligned}\langle v - \text{pr}_{\mathcal{B}}(v), u_i \rangle &= \langle v - (\langle v, u_1 \rangle u_1 + \cdots + \langle v, u_r \rangle u_r), u_i \rangle \\ &= \langle v, u_i \rangle - \sum_{j=1}^r \langle v, u_j \rangle \langle u_j, u_i \rangle \\ &= \langle v, u_i \rangle - \langle v, u_i \rangle \\ &= 0,\end{aligned}$$

da $\langle u_j, u_i \rangle = 0$ for all $i \neq j$. □

Beispiel 13.6.6. Wir betrachten $\mathbb{R}^3$ mit dem Standard IP. Es sei $U = \{x_1 + x_2 - x_3 = 0\}$. Dann ist $u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ und $u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ eine Basis von U . Wir berechnen $U^\perp$: wenn $w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \in U^\perp$, dann gilt $\langle w, u_i \rangle = 0$ fuer $i = 1, 2$, das heisst

$$\begin{aligned}w_1 + w_3 &= 0 \quad \text{und} \quad w_2 + w_3 = 0 \\ \Rightarrow w_1 &= w_2 = -w_3 \\ \Rightarrow U^\perp &= \left(\begin{array}{c} 1 \\ 1 \\ -1 \end{array} \right) \end{aligned}$$

Beispiel 13.6.7. Es sei V der $\mathbb{R}$ -Vektorraum aller Polynome vom Grad ≤ 3 mit dem inneren Produkt aus Beispiel 13.2.5, und es sei U der Unterraum aller Polynome vom Grad ≤ 1 . Dann ist $\dim U^\perp = 2$. Um $U^\perp$ zu bestimmen, waehe eine Basis von U , z.B. $f_0(x) = 1$ und $f_1(x) = x$. Wenn $g(x) = a + bx + cx^2 + dx^3 \in U^\perp$, dann gilt

$$\langle g, f_0 \rangle = 0 \quad \text{und} \quad \langle g, f_1 \rangle = 0.$$

Explizit gibt das Relationen zwischen den Koeffizienten, aus denen wir eine Basis g_2, g_3 bestimmen koennen (Uebung). Um eine *orthonormale* Basis von $U^\perp$ zu bestimmen, wende das Gram-Schmidt Verfahren auf g_2, g_3 an (Uebung).

Wir koennen jetzt die Projektion auf einen Unterraum definieren:

Definition 13.6.8. Es sei $(V, \langle \cdot, \cdot \rangle)$ ein IP Raum (moeglicherweise unendlich-dimensional), und es sei U ein endlich-dimensionaler Unterraum. Dank Theorem 13.6.5 kann jeder Vektor $v \in V$ eindeutig geschrieben werden als

$$v = u + w \quad \text{mit } u \in U \text{ und } w \in U^\perp.$$

Definiere die orthogonale Projektion von v als $\text{pr}_U(v) = u$.

Lemma 13.6.9. (1) pr_U ist eine lineare Abbildung.

(2) $\ker(\text{pr}_U) = U^\perp$ und $\text{im}(\text{pr}_U) = U$.

(3) Fuer alle $v \in V$ gilt $v - \text{pr}_U(v) \in U^\perp$.

(4) Wenn $u_1, \dots, u_r$ eine orthonormale Basis von U ist, dann ist $\text{pr}_U = \tilde{\text{pr}}_U$.

Proof. (1) - (3) sind eine Uebung.

Es sei $\mathcal{B} = \{u_1, \dots, u_r\}$ eine orthonormale Basis von U . Dann gilt $v = x + w$ mit

$$x = \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_r \rangle u_r \in U \quad \text{und} \quad w \in U^\perp.$$

Dann gilt

$$\text{pr}_U(v) = x = \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_r \rangle u_r = \text{pr}_{u_1}(v) + \cdots + \text{pr}_{u_r}(v) = \text{pr}_{\mathcal{B}}(v)$$

□

Als eine einfache Konsequenz erhalten wir folgendes schoenes Resultat:

Satz 13.6.10. Es sei $(V, \langle \cdot, \cdot \rangle)$ ein IP Raum und $U \leq V$ endlich-dimensional. Dann gilt $U = (U^\perp)^\perp$.

Proof. Wir wissen bereits von Lemma 13.6.4, dass $U \subseteq (U^\perp)^\perp$; es reicht daher, die umgekehrte Inklusion zu zeigen. Es sei $v \in (U^\perp)^\perp$. Da $V = U \oplus U^\perp$, gibt es eindeutig bestimmte $u \in U$, $w \in U^\perp$, so dass $v = u + w$. Wir zeigen, dass $w = 0_V$. Da

$$u \in U \subseteq (U^\perp)^\perp \quad \text{und} \quad v \in (U^\perp)^\perp,$$

folgt, dass $w = v - u \in (U^\perp)^\perp$. Aber w ist ebenfalls ein Element von $U^\perp$, das heisst, $w \in U^\perp \cap (U^\perp)^\perp = \{0_V\}$. □

Korollar 13.6.11. Wenn $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler IP Raum ist und $U \leq V$ ist ein Unterraum, dann gilt

$$\dim V = \dim U + \dim U^\perp.$$

13.7. QR Zerlegung. In diesem Kapitel werden wir das Gram-Schmidt Orthogonalisierungsverfahren dazu benutzen um zu zeigen, dass sich jede reelle oder komplexe $(n \times n)$ -Matrix sich in ein Produkt zwei besonders einfacher Matrizen zerlegen lässt.

Zunächst definieren wir zwei neue Familien von sehr wichtigen Matrixgruppen.

Definition 13.7.1.

- (1) Eine Matrix $A \in M_{n \times n}(\mathbb{R})$ ist orthogonal, wenn die Spaltenvektoren eine orthonormale Basis von $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ bilden; hier ist $\langle \cdot, \cdot \rangle$ das Standard IP. Wir schreiben $O(n)$ fuer die Menge aller orthogonalen $(n \times n)$ -Matrizen.
- (2) Eine Matrix $B \in M_{n \times n}(\mathbb{C})$ ist unitaer, wenn die Spaltenvektoren eine orthonormale Basis von $\mathbb{C}^n$ bilden bezueglich des Standard IP auf $\mathbb{C}^n$. Wir schreiben $U(n)$ fuer die Menge aller unitaeren $(n \times n)$ -Matrizen.

Beispiele 13.7.2. (1) Die Matrizen $A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ und $B = \frac{1}{3} \begin{pmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & 2 \end{pmatrix}$ sind orthogonal.

(2) Die Matrix $C = \frac{1}{\sqrt{5}} \begin{pmatrix} i & -2i \\ -2i & -1 \end{pmatrix}$ ist unitaer.

Lemma 13.7.3.

- (1) Eine Matrix $A \in M_{n \times n}(\mathbb{R})$ ist genau dann orthogonal, wenn $A^{-1} = A^t$.
- (2) Eine Matrix $B \in M_{n \times n}(\mathbb{C})$ ist genau dann unitaer, wenn $B^{-1} = B^*$.

Proof. Wir beweisen (1); der Beweis fuer (2) ist aehnlich. Wir muessen zeigen, dass $A^t A = \mathbf{1}_n$. Schreibe $A = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ mit $\mathbf{v}_i \in \mathbb{R}^n$. Dann gilt

$$(A^t A)_{ij} = \mathbf{v}_i^t \mathbf{v}_j = \langle \mathbf{v}_i, \mathbf{v}_j \rangle \quad \forall 1 \leq i, j \leq n.$$

Doch da die $\mathbf{v}_i$ eine orthonormale Basis bilden, gilt $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij}$ und daher $A^t A = \mathbf{1}_n$. □

Bemerkung 13.7.4. Insbesondere folgt, dass $A \in O(n)$ genau dann, wenn $A^t \in O(n)$, das heisst genau dann, wenn die Zeilenvektoren von A eine orthonormale Basis von $\mathbb{R}^n$ sind. Analog gilt, dass $B \in U(n)$ genau dann, wenn $B^* \in U(n)$, das heisst genau dann, wenn die Zeilenvektoren von B eine orthonormale Basis von $\mathbb{C}^n$ sind.

Diese Charakterisierungen von Lemma 13.7.3 sind nuetzlich, um zu zeigen, dass $O(n)$ und $U(n)$ eine Gruppenstruktur unter Matrixmultiplikation haben:

Satz 13.7.5. Die Menge $O(n)$ (bzw. $U(n)$) ist eine Untergruppe von $GL_n(\mathbb{R})$ (bzw. von $GL_n(\mathbb{C})$).

Proof. Wir beweisen den Fall $O(n)$; der Beweis fuer $U(n)$ ist aehnlich. Es ist klar, dass $\mathbf{1}_n \in O(n)$. Wenn $A^{-1} = A^t$, dann gilt

$$(A^{-1})^{-1} = (A^t)^{-1} = (A^{-1})^t,$$

und daher $A^{-1} \in O(n)$. Es seien nun $A, B \in O(n)$. Dann gilt

$$(AB)^{-1} = B^{-1} A^{-1} = B^t A^t = (AB)^t,$$

und daher $AB \in O(n)$. □

Bemerkung 13.7.6. Wir werden diese Gruppen spaeter (siehe Kapitel 15) noch genauer untersuchen, wenn wir ueber Isometrien sprechen.

Wir koennen nun folgenden Satz beweisen:

Theorem 13.7.7.

- (1) Es sei $A \in GL_n(\mathbb{R})$. Dann gibt es $Q \in O(n)$ und eine obere Dreiecksmatrix $R \in M_{n \times n}(\mathbb{R})$, so dass $A = QR$.
- (2) Es sei $A \in GL_n(\mathbb{C})$. Dann gibt es $Q \in U(n)$ und eine obere Dreiecksmatrix $R \in M_{n \times n}(\mathbb{C})$, so dass $A = QR$.

Proof. Es sei $K = \mathbb{R}$ oder $\mathbb{C}$ und $A \in \text{GL}_n(K)$. Schreibe $A = (v_1, \dots, v_n)$ mit $v_i \in \mathbb{R}^n$. Da A invertierbar ist, ist $v_1, \dots, v_n$ eine Basis von K^n . Wir wenden das GS Orthogonalisierungsverfahren auf diese Basis an und erhalten eine orthonormale Basis $w_1, \dots, w_n$ mit der Eigenschaft, dass

$$\begin{aligned} v_1 &= \text{pr}_{w_1}(v_1) \\ v_2 &= \text{pr}_{w_1}(v_2) + \text{pr}_{w_2}(v_2) \\ &\vdots \\ v_n &= \text{pr}_{w_1}(v_n) + \dots + \text{pr}_{w_n}(v_n). \end{aligned}$$

Es sei nun $Q = (w_1, \dots, w_n)$ und

$$R = \begin{pmatrix} \langle v_1, w_1 \rangle & \langle v_2, w_1 \rangle & \dots & \langle v_i, w_1 \rangle & \dots & \langle v_n, w_1 \rangle \\ 0 & \langle v_2, w_2 \rangle & \dots & \langle v_i, w_2 \rangle & \dots & \langle v_n, w_2 \rangle \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \langle v_1, w_i \rangle & \dots & \langle v_n, w_i \rangle \\ 0 & 0 & \dots & 0 & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \dots & \langle v_n, w_n \rangle \end{pmatrix}.$$

Dann laesst sich obige Eigenschaft als Matrix-Multiplikation ausdruecken: $A = QR$. $\square$

Beispiel 13.7.8. Es sei $A = \begin{pmatrix} -1 & 3 \\ 1 & 5 \end{pmatrix}$ mit Spaltenvektoren v_1, v_2 . Das GS Verfahren erzeugt die Basis

$$w_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad w_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

und daher

$$Q = \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \quad \text{und} \quad R = \sqrt{2} \begin{pmatrix} 1 & 1 \\ 0 & 4 \end{pmatrix}.$$

Wir ueberpruefen die Orthogonalitaet von Q : es gilt

$$Q^t Q = Q^2 = \frac{1}{\sqrt{2}^2} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Beispiel 13.7.9. Es sei $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ mit Spaltenvektoren v_1, v_2, v_3 . Durch Anwendung des GS Verfahrens erhalten wir eine orthonormale Basis

$$w_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad w_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \quad w_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

Daher gilt

$$Q = (w_1, w_2, w_3) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

und

$$R = \begin{pmatrix} \frac{2}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{3}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{pmatrix}.$$

Tatsaechlich laesst sich die QR Zerlegung auf alle $M_{n \times n}(K)$ -Matrizen verallgemeinern:

Theorem 13.7.10. Es sei $A \in M_{n \times n}(K)$ und $r = \text{rk}(A)$. Dann gibt es Matrizen

$$Q \in \begin{cases} O(n) & \text{wenn } K = \mathbb{R} \\ U(n) & \text{wenn } K = \mathbb{C} \end{cases}$$

und $R = \begin{pmatrix} C & \star \\ 0 & 0 \end{pmatrix} \in M_{n \times n}(K)$ mit $C \in M_{r \times r}(K)$ eine obere Dreiecksmatrix, so dass $A = QR$.

Proof. Übung fuer die Enthusiasten. □

Beweis der Eindeutigkeit der Zerlegung von Philipp Steiner:

Satz 13.7.11. *Es sei $A \in \text{GL}_n(\mathbb{R})$. Dann gibt es eindeutig bestimmte $Q \in O(n)$ und eine obere Dreiecksmatrix $R \in M_{n \times n}(\mathbb{R})$ mit positive Elementen entlang der Diagonalen, so dass $A = QR$.*

Proof. Nimm an, dass es zwei solcher Zerlegungen gibt:

$$A = Q_1 R_1 = Q_2 R_2$$

Dann folgt, dass $Q_2^t Q_1 = R_2 R_1^{-1}$. Das sowohl die orthogonalen Matrizen als auch die oberen Dreiecksmatrizen eine Untergruppe von $\text{GL}_n(\mathbb{R})$ bilden, ist die Matrix $Q = Q_2^t Q_1$ eine orthogonale Dreiecksmatrix mit positiven Elementen entlang der Diagonalen. Aber die einzige Matrix mit einer solchen Eigenschaft ist 1_n . Die Eindeutigkeit der Zerlegung folgt. □

13.8. Dualraeume von Inneren Produkt Raeumen. Es sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler IP Raum ueber K . In diesem Kapitel werden wir zeigen, dass der Dualraum

$$V^* = \text{Hom}_K(V, K)$$

sich mit Hilfe des inneren Produkts beschreiben laesst. Insbesondere gibt es eine *kanonische* bijektive Abbildung $V \cong V^*$.

Definition 13.8.1. *Es sei $u \in V$. Definiere die Abbildung*

$$\varphi_u : V \rightarrow K, \quad v \mapsto \langle v, u \rangle.$$

Lemma 13.8.2. *Fuer alle $u \in V$ gilt $\varphi_u \in V^*$; mit anderen Worten, die Abbildung $\varphi_u : V \rightarrow K$ ist linear.*

Proof. Klar, da das innere Produkt linear in the ersten Variable ist. □

Theorem 13.8.3 (Darstellungssatz von Riesz). *Es sei $(V, \langle \cdot, \cdot \rangle)$ ein IP Raum ueber K , und es sei $\varphi \in V^*$. Dann gibt es ein eindeutig bestimmtes Element $u \in V$, so dass $\varphi = \varphi_u$; mit anderen Worten,*

$$\varphi(v) = \langle v, u \rangle \quad \forall v \in V.$$

Proof. Nimm an, dass $K = \mathbb{C}$; der Beweis fuer $K = \mathbb{R}$ ist aehnlich.

Es sei $e_1, \dots, e_n$ eine orthonormale Basis von V . Dann gilt fuer $v \in V$, dass

$$v = \sum_{i=1}^n \langle v, e_i \rangle e_i.$$

Es sei nun $\varphi \in V^*$. Dann gilt

$$\begin{aligned} \varphi(v) &= \varphi \left(\sum_{i=1}^n \langle v, e_i \rangle e_i \right) \\ &= \sum_{i=1}^n \langle v, e_i \rangle \varphi(e_i) \\ &= \sum_{i=1}^n \langle v, \overline{\varphi(e_i)} \cdot e_i \rangle \\ &= \left\langle v, \sum_{i=1}^n \overline{\varphi(e_i)} \cdot e_i \right\rangle. \end{aligned}$$

Es sei $u = \sum_{i=1}^n \overline{\varphi(e_i)} \cdot e_i \in V$. Dann gilt $\varphi(v) = \langle v, u \rangle$ fuer alle $v \in V$, und daher $\varphi = \varphi_u$.

Nimm an, dass $\varphi_u = \varphi_w$ fuer $u, w \in V$. Dann gilt

$$\langle v, u \rangle = \langle v, w \rangle \quad \forall v \in V \quad \Rightarrow \quad u = w$$

aufgrund von Lemma 15.2.9 (3). □

Bemerkung 13.8.4. Theorem 13.8.3 zeigt, dass die Abbildung

$$\Phi : V \rightarrow V^*, \quad u \mapsto \varphi_u$$

bijektiv ist. Wenn $K = \mathbb{R}$, dann ist Φ eine lineare Abbildung, da

$$\langle \sim, \alpha u + \beta w \rangle = \alpha \cdot \langle \sim, u \rangle + \beta \cdot \langle \sim, w \rangle \quad \forall \alpha, \beta \in \mathbb{R}, u, w \in V.$$

Wenn aber $K = \mathbb{C}$, dann ist die Abbildung $\Phi : V \rightarrow V^*$ bijektiv und anti-linear:

$$\langle \sim, \alpha u + \beta w \rangle = \bar{\alpha} \cdot \langle \sim, u \rangle + \bar{\beta} \cdot \langle \sim, w \rangle \quad \forall \alpha, \beta \in \mathbb{C}, u, w \in V.$$

Beispiel 13.8.5. Es sei V der Vektorraum aller stetigen Funktionen $f : [-1, 1] \rightarrow \mathbb{R}$ mit dem inneren Produkt aus Beispiel 13.2.5 und $U \leq V$ der Unterraum aller Polynome vom Grad ≤ 2 . Definiere

$$\varphi : U \rightarrow \mathbb{R}, \quad f(x) \mapsto \int_{-1}^1 f(t) \cos(\pi t) dt.$$

Dann ist $\varphi \in V^*$; laut Theorem 13.8.3 gibt es also ein Element $u(x) \in U$, so dass $\varphi(f(x)) = \langle f, u \rangle \quad \forall f(x) \in U$. Der Beweis von dem Theorem sagt uns, wie wir dieses Element $u(x)$ konstruieren koennen. Wir bestimmen zunaechst eine orthonormals Basis von U mit Hilfe des GS Verfahrens; ein Beispiel ist $g_0 = \frac{1}{\sqrt{2}}, g_1(x) = \sqrt{\frac{3}{2}}x, g_2 = \sqrt{\frac{45}{8}}(x^2 - \frac{1}{3})$. Dann ist

$$\begin{aligned} u &= \sum_{i=0}^2 \varphi(g_i) \cdot g_i \\ &= \sum_{i=0}^2 \left(\int_{-1}^1 g_i(t) \cos(\pi t) dt \right) \cdot g_i \\ &= -\frac{45}{2\pi^2} \left(x^2 - \frac{1}{3} \right). \end{aligned}$$

Die Abbildung Φ gibt uns auch die Moeglichkeit, eine Beziehung zwischen Strukturen des Dualraums V^* und Strukturen von V Herzustellen. Ein Beispiel ist die Beziehung zwischen dem Annihilator und dem orthogonalen Komplement eines Unterraums:

Satz 13.8.6. Es sei $(V, \langle \cdot, \cdot \rangle)$ ein IP Raum ueber K , und es sei $U \leq V$. Dann ist $\Phi(U^\perp) \leq V^*$ der Annihilator von U .⁶⁷

Proof. Schreibe $\text{Ann}(U)$ fuer den Annihilator von U . Laut Definition gilt

$$\begin{aligned} w \in U^\perp &\Leftrightarrow \langle u, w \rangle = 0 \quad \forall u \in U \\ &\Leftrightarrow \phi_w(u) = 0 \quad \forall u \in U \\ &\Leftrightarrow U \leq \ker(\varphi_w) \\ &\Leftrightarrow \varphi_w \in \text{Ann}(U). \end{aligned}$$

Da dank Theorem 13.8.3 jedes Element in V^* von der Form φ_w fuer ein $w \in V$ ist, folgt das Resultat. $\square$

13.9. Die adjungierte Abbildung. Es seien $(V, \langle \cdot, \cdot \rangle_V)$ und $(W, \langle \cdot, \cdot \rangle_W)$ endlichdimensionale IP Raume ueber K , und es sei $T : V \rightarrow W$ eine lineare Abbildung.

Definition 13.9.1. Die adjungierte Abbildung T^* von T ist die Abbildung $T^* : W \rightarrow V$, fuer die gilt

$$(65) \quad \langle Tv, w \rangle_W = \langle v, T^*w \rangle_V \quad \forall v \in V, w \in W.$$

Bemerkung 13.9.2. Es ist nicht klar, dass T^* ueberhaupt existiert!

Satz 13.9.3. Die adjungierte Abbildung T^* ist wohl-definiert und linear.⁶⁸

Proof. Es sei $w \in W$, und betrachte die Linearform $\lambda_w \in V^*$ gegeben durch

$$\lambda_w(v) = \langle Tv, w \rangle_W.$$

⁶⁷Das erklart, warum sowohl der Annihilator als auch das orthogonale Komplement mit $U^\perp$ bezeichnet werden.

⁶⁸Die adjungierte Abbildung ist durch Gleichung (65) vollstaendig bestimmt! Mit anderen Worten, wenn wir etwas ueber die adjungierte Abbildung beweisen wollen, benutzen wir diese Gleichung.

Dann existiert laut Theorem 13.8.3 ein eindeutig bestimmtes Element $v' \in V$, so dass

$$\begin{aligned}\lambda_w(v) &= \langle v, v' \rangle_V \quad \forall v \in V \\ \Leftrightarrow \quad \langle Tv, w \rangle_W &= \langle v, v' \rangle_V \quad \forall v \in V.\end{aligned}$$

Definiere $T^*(w) = v'$. Mir muessen nun zeigen, dass T^* eine lineare Abbildung ist.

Es seien $w_1, w_2 \in W$; wir zeigen zunaechst, dass

$$T^*(w_1 + w_2) = T^*w_1 + T^*w_2.$$

Da die adjungierte Abbildung durch (65) vollstaendig bestimmt ist, reicht es zu zeigen, dass

$$\langle v, T^*(w_1 + w_2) \rangle_V = \langle v, T^*w_1 + T^*w_2 \rangle_V \quad \forall v \in V.$$

Doch

$$\begin{aligned}\langle v, T^*(w_1 + w_2) \rangle_V &= \langle Tv, w_1 + w_2 \rangle_W \\ &= \langle Tv, w_1 \rangle_W + \langle Tv, w_2 \rangle_W \\ &= \langle v, T^*w_1 \rangle_V + \langle v, T^*w_2 \rangle_V \\ &= \langle v, T^*w_1 + T^*w_2 \rangle_V,\end{aligned}$$

was zu beweisen war. Die Gleichung

$$T^*(\alpha w) = \alpha T^*(w)$$

kann aehnlich bewiesen werden. □

Bemerkung 13.9.4.

(1) Es gilt $(\text{id}_V)^* = \text{id}_V$.

(2) Es gilt

$$\langle w, Tv \rangle_W = \overline{\langle Tv, w \rangle_W} = \overline{\langle v, T^*w \rangle_V} = \langle T^*w, v \rangle_V$$

und daher $(T^*)^* = T$.

Beispiel 13.9.5. Wir betrachten $\mathbb{C}^m$ als einen IP Raum mit dem Standard IP. Es sei $A = \begin{pmatrix} 2i & 1+i & 7 \\ 1 & 2 & 3 \end{pmatrix}$ und $T_A : \mathbb{C}^3 \rightarrow \mathbb{C}^2$ die induzierte Abbildung. Was ist die Darstellungsmatrix von T_A^* bezueglich der Standardbasen?

Es gilt

$$\langle T_A x, y \rangle_{\mathbb{C}^2} = (Ax)^t \bar{y} = x^t A^t \bar{y} = x^t \overline{A^t y} = x^t \overline{A^* y} = \langle x, T_A^* y \rangle_{\mathbb{C}^3}.$$

Wir werden spaeter sehen, dass sich dieses Beispiel verallgemeinern laesst.

Die adjungierte Abbildung haengt eng mit der dualen Abbildung zusammen. Wir erinnern uns: die *duale Abbildung* $W^* \rightarrow V^*$ ist definiert durch $\ell \mapsto \ell \circ T$ fuer $\ell \in W^*$.

Satz 13.9.6. Schreibe $T_{\text{dual}}^* : W^* \rightarrow V^*$ fuer die *duale Abbildung*. Dann gilt

$$T^* = \Phi_V^{-1} \circ T_{\text{dual}}^* \circ \Phi_W,$$

wobei $\Phi_V : V \rightarrow V^*$ und $\Phi_W : W \rightarrow W^*$ die bijektiven Abbildungen aus Theorem 13.8.3 sind.

Proof. Es reicht aus zu zeigen, dass folgendes Diagramm kommutiert:

$$\begin{array}{ccc} W & \xrightarrow{T^*} & V \\ \Phi_W \downarrow & & \downarrow \Phi_V \\ W^* & \xrightarrow{T_{\text{dual}}^*} & V^* \end{array}$$

Aber das folgt direkt von der Definition der adjungierten Abbildung. □

Bemerkung 13.9.7. Satz 13.9.6 erlaubt uns, T^* und T_{dual}^* kanonisch zu identifizieren. Das erklart, warum beide mit T^* bezeichnet werden. Zusammenfassend gilt fuer IP Raeume: wir identifizieren

- den Dualraum V^* mit V ;
- den Annihilator eines Unterraums $U \leq V$ mit $U^\perp$;
- die Duale Abbildung $T^* : W^* \rightarrow V^*$ mit der adjungierten Abbildung $T^* : W \rightarrow V$.

Lemma 13.9.8. Es seien V, W, U IP R ume ueber $K \in \{\mathbb{R}, \mathbb{C}\}$, und es seien

$$S, T : V \longrightarrow W, \quad R : W \longrightarrow U$$

lineare Abbildungen. Dann gilt

- (1) $(S + T)^* = S^* + T^*$;
- (2) $(\lambda T)^* = \bar{\lambda} \cdot T^*$ fuer alle $\lambda \in K$;
- (3) $(T^*)^* = T$;
- (4) $(R \circ T)^* = T^* \circ R^*$.

Proof. (1) Es sei $w \in W$. Dann gilt $\forall v \in V$

$$\begin{aligned} \langle v, (S + T)^* w \rangle_V &= \langle (S + T)v, w \rangle_W \\ &= \langle Sv, w \rangle_W + \langle Tv, w \rangle_W \\ &= \langle v, S^* w \rangle_V + \langle v, T^* w \rangle_V \\ &= \langle v, (S^* + T^*)w \rangle_V \end{aligned}$$

und daher $(S + T)^* w = (S^* + T^*)w \forall w \in W$.

(2) Uebung.

(3) Es sei $w \in W$. Dann gilt fuer alle $v \in V$

$$\langle (T^*)^* v, w \rangle_W = \langle v, T^* w \rangle_V = \langle Tv, w \rangle_W$$

und daher $(T^*)^* = T$.

(4) Uebung. □

Lemma 13.9.9. Es seien $(V, \langle \cdot, \cdot \rangle_V)$ und $(W, \langle \cdot, \cdot \rangle_W)$ IP R ume ueber K , und es sei $T : V \rightarrow W$ linear.

- $\ker(T^*) = \text{im}(T)^\perp$ und $\ker(T) = \text{im}(T^*)^\perp$;
- $\text{im}(T^*) = (\ker T)^\perp$ und $\text{im}(T) = \ker(T^*)^\perp$.

(Vergl. Satz 6.4.6)

Proof. (1) Es sei $w \in W$. Dann gilt

$$\begin{aligned} w \in \ker(T^*) &\Leftrightarrow T^* w = 0 \\ &\Leftrightarrow \langle v, T^* w \rangle_V = 0 \quad \forall v \in V \\ &\Leftrightarrow \langle Tv, w \rangle_W = 0 \quad \forall v \in V \\ &\Leftrightarrow w \in \text{im}(T)^\perp. \end{aligned}$$

Wir erhalten die zweite Aussage, indem wir die erste auf T^* anwenden.

(2) Wenden Sie die Operation $(\sim)^\perp$ auf die Gleichungen in (1) an. □

13.10. Die Abbildungsmatrix der adjungierten Abbildung. Wir wollen nun Beispiel 13.9.5 verallgemeinern. Genauer wollen wir zeigen, dass die Darstellungsmatrix einer linearen Abbildung zwischen zwei IP R umen durch die adjunkte Matrix gegeben ist, wenn die gewaehlten Basen orthonormal sind.

Wir beginnen mit folgender Beobachtung:

Lemma 13.10.1. Es seien $(V, \langle \cdot, \cdot \rangle_V)$ und $(W, \langle \cdot, \cdot \rangle_W)$ IP R ume ueber K , und es sei $T : V \rightarrow W$ linear. Es seien $\mathcal{B} = (v_1, \dots, v_n)$, bzw. $\mathcal{C} = (w_1, \dots, w_m)$ orthonormale Basen von V , bzw. W , und es sei $A = (a_{ij})$ die Abbildungsmatrix von T . Dann gilt

$$a_{ji} = \langle Tv_i, w_j \rangle_W.$$

Proof. Per Definition ist die i te Spalte von A gegeben durch die Koeffizienten von Tv_i als Linearkombination der Basis $\mathcal{C}$, mit anderen Worten,

$$Tv_i = a_{1i}w_1 + \dots + a_{mi}w_m.$$

Doch da $\mathcal{C}$ orthonormal ist, erhalten wir $a_{ji} = \langle Tv_i, w_j \rangle_W$. □

Satz 13.10.2. Es seien $(V, \langle \cdot, \cdot \rangle_V)$ und $(W, \langle \cdot, \cdot \rangle_W)$ IP R ume ueber K , und es sei $T : V \rightarrow W$ linear. Es seien $\mathcal{B} = (v_1, \dots, v_n)$, bzw. $\mathcal{C} = (w_1, \dots, w_m)$ orthonormale Basen von V , bzw. W . Dann gilt⁶⁹

$$[T^*]_{\mathcal{B}}^{\mathcal{C}} = \left(\overline{[T]_{\mathcal{C}}^{\mathcal{B}}} \right)^t = ([T]_{\mathcal{C}}^{\mathcal{B}})^*.$$

Proof. Es sei $A = (a_{ij}) = [T]_{\mathcal{C}}^{\mathcal{B}}$ und $B = (b_{ij}) = [T^*]_{\mathcal{B}}^{\mathcal{C}}$. Wir wollen zeigen, dass

$$b_{ij} = \overline{a_{ji}} \quad \forall i, j.$$

Nun gilt

$$\begin{aligned} b_{ij} &= \langle T^* w_j, v_i \rangle \\ &= \overline{\langle v_i, T^* w_j \rangle} \\ &= \overline{\langle T v_i, w_j \rangle} \\ &= \overline{a_{ji}}. \end{aligned}$$

□

Korollar 13.10.3. Es sei $A \in M_{n \times m}(K)$ fuer $K = \mathbb{R}$ oder $\mathbb{C}$. Wir betrachten K^n und K^m als IP R ume mit dem Standard IP. Dann gilt

$$(T_A)^* = T_{A^*}.$$

14. SPEKTRALTHEORIE

14.1. Normale Endomorphismen. Wir wollen uns nun mit der Frage besch ftigen, wie orthonormale Basen in der Theorie von Eigenvektoren zusammenhaengen.

Definition 14.1.1. Es sei $(V, \langle \cdot, \cdot \rangle)$ ein IP Raum ueber K mit $K \in \{\mathbb{R}, \mathbb{C}\}$ und $T : V \rightarrow V$ eine lineare Abbildung. Dann ist T orthogonal diagonalisierbar, wenn V eine orthonormale Basis besitzt, die aus Eigenvektoren besteht.

Beispiele 14.1.2. (1) Es sei $V = \mathbb{R}^2$ mit dem Standard IP, $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ und $T_A : V \rightarrow V$ die dazugehoerige lineare Abbildung. Dann sind die Eigenwerte die Wurzeln des Polynoms

$$\lambda^2 - 6\lambda + 8 = 0 \quad \Rightarrow \quad \lambda_{1,2} = 3 \pm 1,$$

das heisst $\lambda_1 = 2$ und $\lambda_2 = 4$. Wir berechnen die Eigenvektoren: $v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ und $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Die Vektoren v_1, v_2 sind orthogonal, d.h. nach der Normalisierung erhalten wir eine orthonormale Basis von Eigenvektoren von V .

(2) Wir betrachten wieder $V = \mathbb{R}^2$ mit dem Standard IP und $T_B : V \rightarrow V$ mit $B = \begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix}$. Dann hat das charakteristische Polynom keine reellen Loesungen, d.h. T_B ist ueber $\mathbb{R}$ nicht diagonalisierbar. Wie ist es ueber $\mathbb{C}$? Die Eigenwerte sind gegeben durch $\lambda_{1,2} = 1 \pm 3i$ mit dazugehoerigen Eigenvektoren $w_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$ und $w_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}$. Diese Vektoren sind orthogonal bezueglich des Standard IP auf $\mathbb{C}^n$, das heisst, T_B ist orthogonal diagonalisierbar.

(3) Es sei $C = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}$. Wir haben in Beispiel 12.1.3 (2) gesehen, dass T_C diagonalisierbar ist, aber nicht ON diagonalisierbar.

Lemma 14.1.3. Wenn T orthogonal diagonalisierbar ist, dann gilt $T \circ T^* = T^* \circ T$.

Proof. Es sei $\mathcal{B}$ eine orthonormale Basis von Eigenvektoren von V . Dann ist $[T]_{\mathcal{B}}^{\mathcal{B}}$ diagonal. Aufgrund von Satz 13.10.2 gilt $[T^*]_{\mathcal{B}}^{\mathcal{B}} = ([T]_{\mathcal{B}}^{\mathcal{B}})^*$, und diese Matrix ist ebenfalls diagonal. Doch alle diagonalen Matrizen kommutieren, das heisst

$$[T \circ T^*]_{\mathcal{B}}^{\mathcal{B}} = [T]_{\mathcal{B}}^{\mathcal{B}} \cdot [T^*]_{\mathcal{B}}^{\mathcal{B}} = [T^*]_{\mathcal{B}}^{\mathcal{B}} \cdot [T]_{\mathcal{B}}^{\mathcal{B}} = [T^* \circ T]_{\mathcal{B}}^{\mathcal{B}}$$

und daher (aufgrund von Bemerkung 4.4.4) gilt $T \circ T^* = T^* \circ T$. □

⁶⁹Wenn $K = \mathbb{R}$, dann ist die komplexe Konjugation die Identitaet und daher die adjunkte Matrix gleich der transponierten Matrix.

Endomorphismen mit der Eigenschaft von Lemma 14.1.3 sind sehr wichtig; wir machen deshalb folgende Definition:

Definition 14.1.4. Es sei $(V, \langle \cdot, \cdot \rangle)$ ein IP Raum ueber K mit $K \in \{\mathbb{R}, \mathbb{C}\}$. Eine lineare Abbildung $T : V \rightarrow V$ ist normal, wenn

$$T \circ T^* = T^* \circ T.$$

Eine Matrix $A \in M_{n \times n}(K)$ ist normal, wenn $A \cdot A^* = A^* \cdot A$.

Bemerkung 14.1.5. Wenn $A \in M_{n \times n}(K)$ normal ist, dann ist die Abbildung $T_A : K^n \rightarrow K^n$ normal bez. des Standard IPs.

Folgendes Resultat ist ein Spezialfall der Korrespondenz zwischen Endomorphismen und Matrizen (Bemerkung 4.4.4):

Satz 14.1.6. Wenn $T : V \rightarrow V$ normal ist, dann ist die Abbildungsmatrix $[T]_{\mathcal{B}}^{\mathcal{B}}$ normal fuer jede orthonormale Basis $\mathcal{B}$ von V .

Proof. Einfache Konsequenz der Beobachtung, dass $[T^*]_{\mathcal{B}}^{\mathcal{B}} = ([T]_{\mathcal{B}}^{\mathcal{B}})^*$ (Satz 13.10.2). $\square$

Bemerkung 14.1.7.

- (1) Jede orthogonale bez. unitaere Matrix ist normal.
- (2) Jede symmetrische Matrix in $M_{n \times n}(\mathbb{R})$ und jede Hermitesche Matrix in $M_{n \times n}(\mathbb{C})$ ist normal.
- (3) Die Matrix $B = \begin{pmatrix} 2i & i-1 \\ i+1 & i \end{pmatrix}$ ist normal.
- (4) Die Matrix $C = \begin{pmatrix} i & i \\ i & 1 \end{pmatrix}$ ist nicht normal.
- (5) Es seien $A, B \in M_{n \times n}(K)$ normal. Dann ist es i.A. nicht wahr, dass AB normal ist! Allerdings gilt: wenn A, B normal sind und $AB = BA$, dann ist AB normal.

Wir haben in Lemma 14.1.3 gezeigt, dass jeder orthogonal diagonalisierbare Endomorphismus normal ist. Gilt die Umkehrung?

Beachte 14.1.8. Die Antwort ist nein: Wir haben in Beispiel 14.1.2 gesehen, dass T_B ueber $\mathbb{R}$ nicht orthogonal diagonalisierbar ist, obwohl eine einfache Rechnung zeigt, dass $BB^t = B^t B$. Allerdings haben wir gesehen, dass T_B ueber $\mathbb{C}$ orthogonal diagonalisierbar ist – ist der Grundkoerper das einzige Hindernis?

Es sei T ein normaler Operator, dessen Eigenwerte in K enthalten sind. Ist T orthogonal diagonalisierbar? Wir betrachten zunaechst ein paar Beispiele.

Beispiel 14.1.9. (1) Es sei $A = \begin{pmatrix} 2i & i-1 \\ i+1 & i \end{pmatrix}$. Dann ist $A^* = \begin{pmatrix} -2i & 1-i \\ -1-i & -i \end{pmatrix}$, und eine einfache Rechnung zeigt, dass A normal ist. Was koennen wir ueber die Eigenwerte und Eigenvektoren sagen?

- A hat Eigenwerte $\lambda_1 = 0$ und $\lambda_2 = 3i$ mit Eigenvektoren $v_1 = \begin{pmatrix} 1-i \\ 2i \end{pmatrix}$ und $v_2 = \begin{pmatrix} i+1 \\ -1 \end{pmatrix}$;
 - A^* hat Eigenwerte $\mu_1 = 0$ und $\mu_2 = -3i$ mit Eigenvektoren $w_1 = \begin{pmatrix} 1-i \\ 2i \end{pmatrix}$ und $w_2 = \begin{pmatrix} i+1 \\ -1 \end{pmatrix}$.
- (2) Se $B = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$. Dann ist B Hermitesch und damit normal. Eine einfache Rechnung zeigt, dass B die Eigenwerte $\lambda_1 = 0$ und $\lambda_2 = 2$ hat, mit jeweiligen Eigenvektoren $v_1 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$ und $v_2 = \begin{pmatrix} i \\ 1 \end{pmatrix}$.

Was faellt Ihnen auf?

Lemma 14.1.10. Es sei $(V, \langle \cdot, \cdot \rangle)$ ein IP Raum ueber K mit $K \in \{\mathbb{R}, \mathbb{C}\}$, und es sei $T : V \rightarrow V$ normal. Dann gilt

- (1) $\|Tv\| = \|T^*v\|$ fuer alle $v \in V$;
- (2) der Operator $T - \lambda \mathbf{1}_V$ fuer $\lambda \in K$ ist ebenfalls normal;

- (3) wenn v ein Eigenvektor von T ist mit Eigenwert λ , dann ist v ein Eigenvektor von T^* mit Eigenwert $\bar{\lambda}$;
 (4) Es seien v_1, v_2 Eigenvektoren von T mit verschiedenen Eigenwerten. Dann gilt $v_1 \perp v_2$.

Proof. (1) Es gilt

$$\|Tv\|^2 = \langle Tv, Tv \rangle = \langle v, T^*Tv \rangle = \langle v, TT^*v \rangle = \overline{\langle T^*v, v \rangle} = \overline{\langle T^*v, T^*v \rangle} = \langle T^*v, T^*v \rangle = \|T^*v\|^2.$$

(2) Explizite Rechnung.

(3) Wenn v ein Eigenvektor von T ist mit Eigenwert λ , dann gilt $(T - \lambda \mathbf{1}_V)v = 0$. Doch $T - \lambda \mathbf{1}_V$ ist normal aufgrund von (2), daher gilt

$$\|(T - \lambda \mathbf{1}_V)v\| = \|(T - \lambda \mathbf{1}_V)^*v\| = 0.$$

Doch $(T - \lambda \mathbf{1}_V)^* = T^* - \bar{\lambda} \mathbf{1}_V$, daher gilt

$$\|(T^* - \bar{\lambda} \mathbf{1}_V)v\| = 0 \Rightarrow (T^* - \bar{\lambda} \mathbf{1}_V)v = 0$$

und daher $T^*v = \bar{\lambda}v$.

(4) Es seien $\lambda_{1,2}$ die Eigenwerte von v_1 und v_2 . Dann gilt

$$\lambda_1 \langle v_1, v_2 \rangle = \langle \lambda_1 v_1, v_2 \rangle = \langle Tv_1, v_2 \rangle = \langle v_1, T^*v_2 \rangle = \langle v_1, \bar{\lambda}_2 v_2 \rangle = \bar{\lambda}_2 \langle v_1, v_2 \rangle,$$

das heisst $(\lambda_1 - \bar{\lambda}_2) \langle v_1, v_2 \rangle = 0$. Da $\lambda_1 \neq \bar{\lambda}_2$, gilt $\langle v_1, v_2 \rangle = 0$. □

Wir koennen jetzt obige Frage beantworten:

Lecture 46

Theorem 14.1.11 (Spektralsatz ueber $\mathbb{C}$). *Es sei $(V, \langle \cdot, \cdot \rangle)$ ein IP Raum ueber $\mathbb{C}$ und $T : V \rightarrow V$ linear. Dann ist T genau dann orthogonal diagonalisierbar, wenn T normal ist.*

Proof. Dank Lemma 14.1.3 ist es ausreichend zu zeigen, dass ein normaler Operator ueber $\mathbb{C}$ orthogonal diagonalisierbar ist. Es folgt von Theorem 12.1.2, dass es eine Basis $\mathcal{C} = (u_1, \dots, u_n)$ von V gibt, so dass $[T]_{\mathcal{C}}^{\mathcal{C}}$ in Jordanscher Normalenform ist. Indem wir das Gram-Schmidt Verfahren auf $\mathcal{C}$ anwenden, erhalten wir eine orthonormale Basis $\mathcal{B} = (v_1, \dots, v_n)$. Da

$$\langle u_1, \dots, u_i, \quad u_j \rangle = \langle v_1, \dots, v_i, \quad v_j \rangle \quad \forall 1 \leq i \leq n,$$

ist $[T]_{\mathcal{B}}^{\mathcal{B}}$ ebenfalls eine obere Dreiecksmatrix. Schreibe $[T]_{\mathcal{B}}^{\mathcal{B}} = (a_{ij})_{1 \leq i, j \leq n}$; wir werden zeigen, dass diese Matrix diagonal ist, d.h. $a_{ij} = 0$ fuer alle $i < j$.

Da $\mathcal{B}$ orthonormal ist, gilt

$$\|Tv_1\|^2 = |a_{11}|^2 \quad \text{und} \quad \|T^*v_1\|^2 = \sum_{k=1}^n |\overline{a_{1k}}|^2 = |a_{11}|^2 + |a_{12}|^2 + \dots + |a_{1n}|^2.$$

Doch T ist normal, und daher $\|Tv_1\| = \|T^*v_1\|$

$$\Rightarrow |a_{12}|^2 = \dots = |a_{1n}|^2 = 0.$$

Wir wiederholen nun das Argument fuer $\|Tv_2\|^2$ usw. □

Bemerkung 14.1.12. *Die Annahme, dass wir ueber $\mathbb{C}$ arbeiten, ist in dem Beweis von Theorem 14.1.11 nur deshalb wichtig, weil wir dadurch sicher sein koennen, dass eine Jordanbasis existiert.*

Korollar 14.1.13. *Es sei V ein endlich-dimensionaler IP Raum ueber $\mathbb{R}$ und $T : V \rightarrow V$ ein normaler Operator. Wenn V eine Jordanbasis besitzt, dann ist T orthogonal diagonalisierbar.*

Korollar 14.1.14. *Es sei $A \in M_{n \times n}(\mathbb{C})$ normal. Dann gibt es eine unitaere Matrix $U \in \text{GL}_N(\mathbb{C})$, so dass U^*AU diagonal ist.*

Proof. Dank Satz 14.1.6 ist die Abbildung $T_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ orthogonal diagonalisierbar, das heisst, es gibt eine ON Basis von $\mathbb{C}^n$ (bezuglich des Standard IPs), die aus Eigenvektoren besteht. Insbesondere ist $[T_A]_{\mathcal{B}}^{\mathcal{B}}$ diagonal. Um das Resultat zu beweisen, muessen wir daher zeigen, dass die Basiswechselmatrix $U = [\text{id}_{\mathbb{C}^n}]_{\mathcal{B}}^{\mathcal{E}}$ unitaer ist⁷⁰; hier ist $\mathcal{E}$ die Standardbasis von $\mathbb{C}^n$, die ebenfalls orthonormal ist. Da $\text{id}_{\mathbb{C}^n}^* = \text{id}_{\mathbb{C}^n}$, folgt von Satz 13.10.2, dass

$$U^* = [\text{id}_V]_{\mathcal{E}}^{\mathcal{B}}.$$

Daher gilt

$$UU^* = [\text{id}_{\mathbb{C}^n}]_{\mathcal{B}}^{\mathcal{E}} [\text{id}_V]_{\mathcal{E}}^{\mathcal{B}} = [\text{id}_{\mathbb{C}^n}]_{\mathcal{E}}^{\mathcal{E}} = \mathbf{1}_n,$$

⁷⁰Dieses Resultat gilt ganz allgemein fuer die Basiswechselmatrix zwischen zwei orthonormalen Basen.

und aehnlich, dass $U^*U = \mathbf{1}_n$. □

Dieses Resultat hat folgende schoene Konsequenz:

Theorem 14.1.15. *Sei $A \in M_{n \times n}(\mathbb{C})$. Dann laesst sich A als das kommutierende Produkt einer positiv definiten Hermiteschen Matrix und einer unitaeren Matrix schreiben. Wenn $A \in GL_N(\mathbb{C})$, dann ist diese Zerlegung eindeutig.*

Proof. Es sei U wie in dem Beweis von Korollar 14.1.14, so dass $D = UAU^*$ diagonal ist. Schreibe $D = RJ$ als Produkt diagonaler Matrizen, wobei $R \in M_{n \times n}(\mathbb{R})$ positiv definit ist und di Eigenwerte von $J \in M_{n \times n}(\mathbb{C})$ alle Betrag 1 haben; insbesondere ist J unitaer. Dann gilt

$$A = U^*RJU = (U^*RU)(U^*JU),$$

die Matrix U^*JU ist natuerlich unitaer, da $U(N)$ eine Gruppe ist, und U^*RU ist Hermitesch. Da eine Hermitesche Matrix genau dann positiv definit ist, wenn alle Eigenwerte > 0 sind (Uebung), folgt die Existenz der Zerlegung. Die Kommutativitaet der beiden Matrizen ist klar.

Die Eindeutigkeit der Zerlegung fuer invertierbare Matrizen ist eine Uebung. □

14.2. Spektraltheorie ueber $\mathbb{R}$. Wir haben in Beispiel 14.1.2 (2) gesehen, dass nicht jeder normale Operator ueber einem reellen IP Raum orthogonal diagonalisierbar ist. Koennen wir ein Kriterium finden, fuer welche normalen Operatoren die Implikation wahr ist?

Lemma 14.2.1. *Es sei V ein IP Raum ueber $\mathbb{R}$ und $T : V \rightarrow V$ orthogonal diagonalisierbar. Dann gilt $T^* = T$.*

Proof. Es sei $\mathcal{B}$ eine orthonormale Basis von Eigenvektoren, d.h. $[T]_{\mathcal{B}}^{\mathcal{B}}$ ist diagonal mit reellen Eintraegen und daher $[T]_{\mathcal{B}}^{\mathcal{B}} = ([T]_{\mathcal{B}}^{\mathcal{B}})^t$. Doch dann gilt

$$[T^*]_{\mathcal{B}}^{\mathcal{B}} = \left(\overline{[T]_{\mathcal{B}}^{\mathcal{B}}} \right)^t = ([T]_{\mathcal{B}}^{\mathcal{B}})^t = [T]_{\mathcal{B}}^{\mathcal{B}}$$

und daher $T^* = T$. □

Definition 14.2.2. *Es sei $(V, \langle \cdot, \cdot \rangle)$ ein IP Raum ueber K mit $K \in \{\mathbb{R}, \mathbb{C}\}$. Eine lineare Abbildung $T : V \rightarrow V$ ist selbst-adjungiert, wenn*

$$T = T^*.$$

Eine Matrix $A \in M_{n \times n}(K)$ ist selbstadjungiert, wenn $A^ = A$.*

Beachte 14.2.3. *Wenn $K = \mathbb{C}$, dann ist selbstadjungiert das Gleiche wie hermitesch; wenn $K = \mathbb{R}$, dann ist selbstadjungiert das Gleiche wie symmetrisch.*

Folgendes Resultat ist das Analogon von Satz 14.1.6 fuer selbstadjungierte Matrizen:

Satz 14.2.4.

- (1) *Wenn $T : V \rightarrow V$ selbst-adjungiert ist, dann ist die Abbildungsmatrix $[T]_{\mathcal{B}}^{\mathcal{B}}$ selbstadjungiert fuer jede orthonormale Basis $\mathcal{B}$ von V .*
- (2) *Wenn $A \in M_{n \times n}(K)$ selbst-adjungiert ist, dann ist die Abbildung $T_A : K^n \rightarrow K^n$ selbst-adjungiert bez. des Standard IPs.*

Proof. Uebung. □

Bemerkung 14.2.5. *Jede selbstadjungierte lineare Abbildung $T : V \rightarrow V$ ist normal.*

Beispiele 14.2.6. (1) Es sei $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Dann sind die Eigenwerte 3 und 1.

(2) Es sei $B = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$. Dann sind die Eigenwerte -2 und 0 .

(3) Es sei $C = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}$. Dann sind die Eigenwerte 5 und 1 .

Was faellt auf?

Satz 14.2.7. *Es sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler IP Raum ueber K , und es sei $T : V \rightarrow V$ ein selbstadjungierter Endomorphismus. Dann gilt*

- (1) *Die Eigenwerte von T sind reell.*
- (2) *Das charakteristische Polynom $\chi_T(x)$ zerfaellt ueber K in Linearfaktoren.*

Proof. (1) Es sei $\lambda \in K$ ein Eigenwert mit Eigenvektor v . Da jeder selbstadjungierte Operator normal ist, ist laut Lemma 14.1.10 v ein Eigenvektor von T^* mit Eigenwert $\bar{\lambda}$. Daher gilt

$$\lambda v = Tv = T^*v = \bar{\lambda}v$$

und daher $\lambda = \bar{\lambda}$, d.h. $\lambda \in \mathbb{R}$.

(2) Nimm an, dass $K = \mathbb{R}$. Es sei $\mathcal{B}$ eine orthonormale Basis von V und $A = [T]_{\mathcal{B}}^{\mathcal{B}} \in M_{n \times n}(\mathbb{R})$, so dass $\chi_T(x) = \chi_A(x)$. Es folgt von Satz 14.2.4 (1), dass A selbstadjungiert ist.

Da wir Tensorprodukte noch nicht kennen⁷¹, behelfen wir uns mit einem Trick. Betrachte den Vektorraum $\mathbb{C}^n$ mit dem Standard IP und die lineare Abbildung $T_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$; per Definition gilt $[T_A]_{\mathcal{E}}^{\mathcal{E}} = A$, wobei $\mathcal{E}$ die Standard Basis von $\mathbb{C}^n$ ist. Dann gilt $\chi_A(x) = \chi_{T_A}(x)$, und es folgt von dem fundamentalen Satz der Algebra, dass das Polynom $\chi_{T_A}(x) \in \mathbb{C}[x]$ in Linearfaktoren zerfaellt:

$$(66) \quad \chi_{T_A}(x) = \prod_{i=1}^n (\lambda_i - x),$$

wobei $\lambda_1, \dots, \lambda_n$ die Eigenwerte von T_A sind.

Nun folgt von Satz 14.2.4 (2), dass T_A selbstadjungiert ist, und daher gilt $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, das heisst, Gleichung (66) gilt ueber $\mathbb{R}$. $\square$

Als einfache Konsequenz erhalten wir folgendes Theorem:

Theorem 14.2.8. *Es sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler IP Raum ueber $\mathbb{R}$, und es sei $T : V \rightarrow V$ linear. Dann ist T genau dann orthogonal diagonalisierbar, wenn T selbst-adjungiert ist.*

Proof. Da das charakteristische Polynom in Linearfaktoren zerfaellt, folgt das Resultat mit dem gleichen Argument wie in dem Beweis von Theorem 14.1.11. $\square$

Dadurch, dass wir dieses Theorem mit Hilfe von Matrizen ausdruecken, erhalten wir (vergl. Korollar 14.1.14)

Korollar 14.2.9. *Wenn $A \in M_{n \times n}(\mathbb{R})$ selbstadjungiert (d.h. symmetrisch) ist, dann gibt es eine orthogonale Matrix $O \in GL_n(\mathbb{R})$, so dass $O^t A O$ diagonal ist.*

Proof. Uebung. $\square$

15. ISOMETRIEN

Lecture 47

15.1. Definition und erste Eigenschaften. Wir wollen uns nun mit einer anderen Klasse von linearen Abbildungen auf IP Raeumen beschaeftigen, naemlich mit solchen, unter denen sich der Abstand zwischen zwei Vektoren nicht aendert.

Definition 15.1.1. *Es seien $(V, \langle \cdot, \cdot \rangle_V)$, $(W, \langle \cdot, \cdot \rangle_W)$ IP Raeume ueber K . Eine Abbildung $T : V \rightarrow W$ ist eine Isometrie, wenn*

$$(67) \quad \|Tv\|_W = \|v\|_V$$

fuer alle $v \in V$. Eine Isometrie ist eine lineare Isometrie, wenn sie eine lineare Abbildung ist.

Beispiel 15.1.2. Es sei $V = \mathbb{R}^2$ mit dem Standard IP. Dann sind Rotationen um den Nullvektor und Relektionen in einer Graden durch 0_V lineare Isometrien. Ein schoenes Beispiel fuer eine Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, die die Gleichung (67) erfuehlt, aber keine Isometrie ist, ist

$$f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{cases} \begin{pmatrix} x \\ y \end{pmatrix} & \text{wenn } x \geq 0 \\ \begin{pmatrix} x \\ -y \end{pmatrix} & \text{wenn } x < 0 \end{cases}$$

Bemerkung 15.1.3. *Es sei T eine lineare Isometrie. Dann ist T injektiv: Wenn $Tv = 0$, dann gilt*

$$\|v\| = \|Tv\| = 0 \quad \Rightarrow \quad v = 0_V.$$

⁷¹Was wir gerne machen wuerden, ist, den reellen Vektorraum V zu einem Vektorraum V' ueber $\mathbb{C}$ und T zu einem Endomorphismus $T' : V' \rightarrow V'$ zu 'erweitern'. Wenn wir dieses in einer Art und Weise tun koennten, ohne dass sich das charakteristische Polynom von T aendert, dann koennten wir Theorem 14.1.11 und Teil (1) dieses Satz anwenden. Eine solche 'Erweiterung' von V und T gibt es tatsaechlich: V' ist das *Tensorprodukt* von V ueber $\mathbb{R}$ mit $\mathbb{C}$, geschrieben $V' = V \otimes_{\mathbb{R}} \mathbb{C}$. Koennen Sie erraten, wie V' und T' definiert sind?

Wir koennen eine lineare Isometrie auch mit Hilfe der IPs charakterisieren:

Lemma 15.1.4. *Eine lineare Abbildung $T : V \rightarrow W$ ist genau dann eine Isometrie, wenn*

$$(68) \quad \langle Tv_1, Tv_2 \rangle_W = \langle v_1, v_2 \rangle \quad \forall v_1, v_2 \in V.$$

Proof. Wir beweisen das Resultat in dem Fall $K = \mathbb{R}$; der komplexe Fall ist aehnlich. Wenn T eine Isometrie ist, dann gilt (Lemma 15.2.12)

$$\begin{aligned} \langle Tv_1, Tv_2 \rangle_W &= \frac{1}{2} (\|Tv_1 + Tv_2\|^2 - \|Tv_1\|^2 - \|Tv_2\|^2) \\ &= \frac{1}{2} (\|T(v_1 + v_2)\|^2 - \|Tv_1\|^2 - \|Tv_2\|^2) \\ &= \frac{1}{2} (\|v_1 + v_2\|^2 - \|v_1\|^2 - \|v_2\|^2) \\ &= \langle v_1, v_2 \rangle_W. \end{aligned}$$

□

Wir betrachten nun den Fall, dass $V = W$. Beispiele von linearen Isometrien sind die linearen Abbildungen, die durch orthogonale bzw. unitaere Matrizen gegeben sind:

Satz 15.1.5. *Es sei $A \in O(n)$ (fuer $K = \mathbb{R}$) bzw. $A \in U(n)$ (fuer $K = \mathbb{C}$), und es sei $T_A : K^n \rightarrow K^n$ die dazugehoerige lineare Abbildung. Betrachte K^n als IP Raum mit dem Standard IP. Dann ist T_A eine Isometrie.*

Proof. Es gilt

$$\langle T_A(u), T_A(v) \rangle = \langle u, T_A^* \circ T_A(v) \rangle = \langle u, T_{A^*} \circ T_A(v) \rangle = \langle u, T_{A^*A}(v) \rangle = \langle u, v \rangle.$$

□

In Satz 15.1.7 werden wir zeigen, dass man umgekehrt jeder linearen Isometrie $V \rightarrow V$ eine orthogonale bzw. unitaere Matrix zuordnen kann.

Lemma 15.1.6. *Es sei $(V, \langle \cdot, \cdot \rangle)$ ein IP Raum und $T : V \rightarrow V$ linear. Dann sind folgende Aussagen aequivalent:*

- (1) T ist eine Isometrie;
- (2) es gibt eine ON Basis $v_1, \dots, v_n$ von V , so dass $Tv_1, \dots, Tv_n$ ebenfalls eine ON Basis ist;
- (3) fuer jede ON Basis $v_1, \dots, v_n$ ist $Tv_1, \dots, Tv_n$ ebenfalls eine ON Basis von V ;
- (4) $T \circ T^* = \text{id}_V$;
- (5) $T^* \circ T = \text{id}_V$;
- (6) T^* ist eine Isometrie.

Proof. (1) $\Rightarrow$ (2), (3): Es sei $v_1, \dots, v_n$ eine ON Basis. Dann gilt aufgrund von Lemma 15.1.4, dass

$$\langle Tv_i, Tv_j \rangle_V = \langle v_i, v_j \rangle = \delta_{ij},$$

das heisst, die $Tv_1, \dots, Tv_n$ sind ebenfalls orthonormal. Da sie die richtige Anzahl von Vektoren haben und linear unabhaengig sind, sind sie eine Basis.

(3) $\Rightarrow$ (4): fuer alle $1 \leq i, j \leq n$ gilt

$$\langle v_i, v_j \rangle = \langle v_i, T^*Tv_j \rangle,$$

und daher $\langle v, v_j \rangle = \langle v, T^*Tv_j \rangle$ fuer alle $v \in V$. Da das IP nicht degeneriert ist, folgt daraus, dass $v_j = T^*Tv_j$ fuer alle $1 \leq j \leq n$ und daher $T^*T = \mathbf{1}_V$.

Die restlichen Implikationen sind eine Uebung.

□

Satz 15.1.6 legt nahe, dass lineare Isometrien eng mit orthogonalen und unitaeren Matrizen zusammenhaengen:

Satz 15.1.7. *Es sei $T : V \rightarrow V$ eine lineare Isometrie und $\mathcal{B}$ eine ON Basis von V . Dann ist die Abbildungsmatrix $[T]_{\mathcal{B}}^{\mathcal{B}}$*

$$\begin{cases} \text{orthogonal} & \text{wenn } K = \mathbb{R} \\ \text{unitaer} & \text{wenn } K = \mathbb{C} \end{cases}.$$

Proof. Wenn T eine lineare Isometrie ist, dann gilt laut Satz 15.1.6 $TT^* = T^*T = \text{id}_V$, das heisst

$$[T]_{\mathcal{B}}^{\mathcal{B}} [T^*]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{1}_n.$$

Doch da $\mathcal{B}$ orthonormal ist, gilt $[T^*]_{\mathcal{B}}^{\mathcal{B}} = ([T]_{\mathcal{B}}^{\mathcal{B}})^*$ (Satz 13.10.2), und daher ist $[T]_{\mathcal{B}}^{\mathcal{B}}$ orthogonal fuer $K = \mathbb{R}$ und unitaer fuer $K = \mathbb{C}$. $\square$

Beachte 15.1.8.

- (1) $\forall A \in O(n)$ gilt $\det(A) = \pm 1$.
- (2) $\forall B \in U(n)$ gilt $|\det(B)| = 1$.

Lemma 15.1.9. Die Untermenge

$$SO(n) = \{A \in O(n) \mid \det(A) = 1\} \subseteq O(n)$$

ist eine Untergruppe von $O(n)$; wir nennen sie die speziellen orthogonalen Matrizen. Ebenso ist die Untermenge

$$SU(n) = \{B \in U(n) \mid \det(B) = 1\}$$

eine Untergruppe von $U(n)$.

15.2. Klassifikation der Elemente in $O(2)$ und in $SO(3)$. Wir beginnen mit folgenden elementaren Beobachtungen:

Lemma 15.2.1. Es sei $A \in O(n)$. Wenn λ ein Eigenwert von A ist, dann gilt $\lambda \in \{\pm 1\}$.

Proof. Nimm an, dass $Av = \lambda v$ fuer $v \in \mathbb{R}^n$, $v \neq 0_V$. Wenn $\langle \cdot, \cdot \rangle$ das Standard IP auf $\mathbb{R}^n$ ist, dann gilt

$$\langle v, v \rangle = \langle Av, Av \rangle = \langle \lambda v, \lambda v \rangle = \lambda^2 \langle v, v \rangle.$$

Da $v \neq_V$ folgt, dass $\lambda^2 = 1$. $\square$

Lemma 15.2.2. Es sei $v \in \mathbb{R}^2$ mit $\|v\| = 1$ (bezuglich der Norm des Standard IPs). Dann gibt es $\theta \in [0, 2\pi)$, so dass

$$v = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}.$$

Satz 15.2.3. Es sei $A \in O(2)$.

- (1) Wenn $\det(A) = 1$, dann gibt es $\theta \in [0, 2\pi)$ so dass

$$A = R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Die Abbildung $T_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ist die Rotation gegen den Uhrzeigersinn um den Winkel θ .

- (2) Wenn $\det(A) = -1$, dann gibt es eine Basis $\mathcal{B} = (v_1, v_2)$ von $\mathbb{R}^2$, so dass

$$[T_A]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} :$$

T_A beschreibt die Reflektion in der Geraden $\ell = \begin{pmatrix} \text{---} \end{pmatrix} v_1 \begin{pmatrix} \text{---} \end{pmatrix}$. Weiterhin gibt es $\theta \in [0, 2\pi)$, so dass

$$A = R_\theta \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Explizit ist die Basis $\mathcal{B}$ gegeben durch

$$v_1 = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} \cos\left(\frac{\theta+\pi}{2}\right) \\ \sin\left(\frac{\theta+\pi}{2}\right) \end{pmatrix}.$$

Proof. Betrachte $\mathbb{R}^2$ mit dem standard IP $\langle \cdot, \cdot \rangle$. Schreibe $A = (u_1, u_2)$ mit $u_i \in \mathbb{R}^2$. Dann gilt $\|u_1\| = 1$, und es gilt $u_1 = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$ aufgrund von Lemma 15.2.2. Weiterhin gilt $\|u_2\| = 1$ und $u_1 \perp u_2$. Aber eine einfache Rechnung zeigt, dass

$$\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}^\perp = \begin{pmatrix} \cos\left(\theta + \frac{\pi}{2}\right) \\ \sin\left(\theta + \frac{\pi}{2}\right) \end{pmatrix} \quad \text{---} \quad \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix} \quad \text{---}$$

Aber  $\begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}$  enthaelt genau zwei Elemente mit Norm 1, naemlich

$$u'_2 = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix} \quad \text{und} \quad u''_2 = \begin{pmatrix} \sin(\theta) \\ -\cos(\theta) \end{pmatrix}.$$

Wenn $u_2 = u'_2$, dann ist

$$A = R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

und $\det(A) = 1$.

Wenn $u_2 = u''_2$, dann ist

$$A = R_\theta \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

und $\det(A) = -1$. In diesem Fall zeigt eine einfache Rechnung, dass

$$v_1 = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} \cos\left(\frac{\theta+\pi}{2}\right) \\ \sin\left(\frac{\theta+\pi}{2}\right) \end{pmatrix}$$

Eigenvektoren von A sind mit jeweiligen Eigenwerten $+1$ und -1 . □

Korollar 15.2.4. *Es sei V ein 2-dimensionaler Euklidischer Raum und $T : V \rightarrow V$ eine lineare Isometrie. Wenn $\det(T) = -1$, dann existiert eine ON Basis $\mathcal{B}$ von V , so dass*

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}.$$

Proof. Es sei $\mathcal{C}$ eine ON Basis von V . Dann ist $[T]_{\mathcal{C}}^{\mathcal{C}} \in O(2)$. Laut Satz 15.2.3 (2) hat $[T]_{\mathcal{C}}^{\mathcal{C}}$ (und daher T) Eigenwerte ± 1 , und die normalisierten Eigenvektoren bilden eine ON Basis von V . □

Betrachten wir nun $SO(3)$:

Satz 15.2.5. *Es sei $A \in SO(3)$. Dann ist 1 ein Eigenwert von A . Weiterhin gibt es eine orthonormale Basis $\mathcal{B} = (v_1, v_2, v_3)$ und $\theta \in [0, 2\pi)$, so dass⁷²*

$$[T_A]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & & 0_{1 \times 2} \\ & 0_{2 \times 1} & R_\theta \end{pmatrix}.$$

Weiterhin gilt:

(1) Wenn -1 ein Eigenwert von A ist, dann ist $\theta = \pi$ und daher

$$[T_A]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

(2) Wenn -1 kein Eigenwert von A ist, dann ist $\theta \neq \pi$.

Proof. Betrachte $\mathbb{R}^3$ mit dem standard IP $\langle \cdot, \cdot \rangle$. Das charakteristische Polynom $\chi_A(x) \in \mathbb{R}[x]$ hat Grad 3 und daher mindestens eine reelle Wurzel $\lambda = \pm 1$. Es sei $v_1 \in \mathbb{R}^3$ ein dazugehoeriger Eigenvektor mit $\|v_1\| = 1$. Betrachte das orthogonale Komplement $U = v_1^\perp$, so dass

$$\mathbb{R}^3 = \text{wavy line diagram} \ v_1 \text{ wavy line diagram} \oplus U.$$

Behauptung. Es gilt $T_A(U) \subseteq U$.

Beweis der Behauptung: Es sei $u \in U$. Laut Satz 15.1.5 ist T_A eine Isometrie bez. des Standard IPs, d.h.

$$\langle v_1, T_A(u) \rangle = \lambda \cdot \langle T_A(v_1), T_A(u) \rangle = \lambda \cdot \langle v_1, u \rangle = 0,$$

und daher $T_A(u) \in U$ wie erwartet.

Nun ist $(T_A)|_U : U \rightarrow U$ eine lineare Isometrie. Es sei $\mathcal{C} = \{w_1, w_2\}$ eine ON Basis von U ; dann ist laut Satz 15.1.7 die Abbildungsmatrix $[(T_A)|_U]_{\mathcal{C}}^{\mathcal{C}}$ orthogonal, d.h ein Element von $O(2)$. Es sei $\mathcal{B} = (v_1, w_1, w_2)$. Wir unterscheiden nun zwei Faelle:

⁷²Geomtrisch gesehen ist T_A Rotation mit Winkel θ um die Axe  v_1 .

- (1) Falls $\lambda = +1$, dann ist die Abbildungsmatrix von T_A bezüglich der Basis $\mathcal{B}$ gegeben durch

$$[T_A]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0_{1 \times 2} \\ 0_{2 \times 1} & B \end{pmatrix},$$

mit $\det(B) = 1$. Es folgt daher von Satz ??, dass $\exists \theta \in [0, 2\pi)$, so dass $B = R_\theta$.

- (2) Falls $\lambda = -1$, dann gilt

$$[T_A]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} -1 & 0_{1 \times 2} \\ 0_{2 \times 1} & B \end{pmatrix}$$

mit $\det(B) = -1$. Laut Korollary 15.2.4 hat $(T_A)|_U$ Eigenwerte ± 1 mit Eigenvektoren v_2, v_3 . Wenn wir v_2, v_3 normieren, dann ist $\mathcal{B}' = (v_1, v_2, v_3)$ eine ON Basis mit

$$[T_A]_{\mathcal{B}'}^{\mathcal{B}'} = \begin{pmatrix} -1 & & \\ & 1 & \\ & & -1 \end{pmatrix}.$$

□

Herausforderung. Klassifizieren Sie die Elemente von $U(2)$.

15.3. Singulaerwertzerlegung. Wir erinnern uns an Definition 4.5.10: zwei Matrizen $A, B \in M_{m \times n}(K)$ sind äquivalent, wenn es Matrizen $P \in \text{GL}_m(K)$ und $Q \in \text{GL}_n(K)$ gibt, so dass

$$B = P \cdot A \cdot Q.$$

Weiterhin haben wir gesehen (Satz 4.5.12), dass ‘Äquivalenz’ eine Äquivalenzrelation auf $M_{m \times n}(K)$ definiert, mit $\min\{m, n\} + 1$ Äquivalenzklassen (Korollar 4.7.4). Was passiert nun, wenn wir ausserdem fordern, dass P und Q orthogonal bzw. unitär sind?

Definition 15.3.1.

- (1) Zwei Matrizen $A, B \in M_{m,n}(\mathbb{R})$ sind orthogonal äquivalent, wenn es $P \in O(m)$, $Q \in O(n)$ gibt, so dass $B = PAQ$.
- (2) Zwei Matrizen $A, B \in M_{m,n}(\mathbb{C})$ sind unitär äquivalent, wenn es $P \in U(m)$, $Q \in U(n)$ gibt, so dass $B = PAQ$.

Bemerkung 15.3.2. Dies sind Äquivalenzrelationen. Was können wir über die Äquivalenzklassen sagen? Die Frage wurde zuerst von Eugenio Beltrami und Camille Jordan beantwortet.

Theorem 15.3.3. Es seien $(V, \langle \cdot, \cdot \rangle_V)$ und $(W, \langle \cdot, \cdot \rangle_W)$ IP Räume über K und $T : V \rightarrow W$ linear mit Rang r . Dann existieren ON Basen $\mathcal{B} = (v_1, \dots, v_n)$ von V und $\mathcal{C} = (w_1, \dots, w_m)$ von W , sowie eindeutig bestimmte positive reelle Zahlen

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$$

so dass

$$T(v_i) = \begin{cases} \sigma_i w_i & \text{für } i = 1, \dots, r \\ 0_W & \text{für } i > r \end{cases},$$

das heisst

$$[T]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} \sigma_1 & 0 & \dots & \dots & & \\ 0 & \sigma_2 & \dots & \dots & & \\ & & \ddots & \ddots & & \\ 0 & 0 & \dots & \sigma_r & & \\ & & 0_{m-r,r} & & 0_{m-r,n-r} & \end{pmatrix}$$

Hierbei sind $\sigma_1, \dots, \sigma_r$ die Wurzeln der positiven Eigenwerte von $TT^* \in \text{End}(W)$ geordnet nach ihrer Grösse; sie heissen die Singulaerwerte von T .

Bemerkung 15.3.4. $\sigma_1, \dots, \sigma_r$ sind ebenfalls die Wurzeln der positiven Eigenwerte von $T^*T \in \text{End}(V)$.

Korollar 15.3.5. Zwei Matrizen $A, B \in M_{m,n}(\mathbb{R})$ (bzw. $A, B \in M_{m,n}(\mathbb{C})$) sind genau dann orthogonal (bzw. unitär) äquivalent, wenn sie den gleichen Rang und die gleichen Singulaerwerte haben.

Um Theorem 15.3.3 zu beweisen, brauchen wir folgendes Resultat:

Lemma 15.3.6. Es sei T wie in Theorem 15.3.3. Dann gilt:

- (1) $\text{rk}(T) = \text{rk}(T^*)$;
- (2) $\ker(TT^*) = \ker(T^*)$;

- (3) $\text{im}(TT^*) = \text{im}(T)$;
 (4) fuer alle $w \in W$ gilt $\langle w, TT^*w \rangle_W \geq 0$;
 (5) TT^* ist selbst-adjungiert;
 (6) die Eigenwerte von TT^* sind nicht-negatieve reelle Zahlen.

Proof. (1) Der Rang einer linearen Abbildung ist definiert als der Rang der Abbildungsmatrix in beliebigen Basen von V, W . Es seien $\mathcal{B}, \mathcal{C}$ ON Basen. Dann gilt

$$\text{rk}(T) = \text{rk}[T]_{\mathcal{C}}^{\mathcal{B}} = \text{rk}([T]_{\mathcal{C}}^{\mathcal{B}})^* = \text{rk}[T^*]_{\mathcal{B}}^{\mathcal{C}} = \text{rk}(T^*).$$

(2) Klarerweise gilt $\ker(T^*) \subseteq \ker(TT^*)$, wir muessen also die umgekehrte Inklusion zeigen. Es sei $w \in \ker(TT^*)$. Dann gilt

$$\langle T^*w, T^*w \rangle_V = \langle w, TT^*w \rangle_W = \langle w, 0_W \rangle_W = 0$$

und daher $T^*w = 0$, d.h. $w \in \ker(T^*)$.

(3) Klarerweise gilt $\text{im}(TT^*) \subseteq \text{im}(T)$. Da beide Unterraeume die gleiche Dimension haben (Uebung) muessen sie gleich sein.

(4) Fuer alle $w \in W$ gilt

$$\langle w, TT^*w \rangle_W = \langle T^*w, T^*w \rangle_V \geq 0.$$

(5) Klar.

(6) Da TT^* selbst-adjungiert ist, sind laut Satz 14.2.7 die Eigenwerte reell. Es sei λ ein Eigenwert mit Eigenvektor w . Dann gilt

$$0 \leq \langle w, TT^*w \rangle = \langle w, \lambda w \rangle = \lambda \langle w, w \rangle_W$$

und daher $\lambda \geq 0$. □

Wir koennen jetzt Theorem 15.3.3 beweisen:

Proof. Da TT^* selbstadjungiert und insbesondere normal ist, koennen wir den Spektralsatz (Theorem 14.2.8) anwenden und erhalten eine ON Basis $w_1, \dots, w_m$ von W , die aus Eigenvektoren von TT^* besteht, mit jeweiligem Eigenwert $\lambda_i \geq 0$. Durch eine Permutation der Basiselemente koennen wir annehmen, dass

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_m = 0.$$

Fuer $1 \leq i \leq r$ sei $u_i = T^*w_i$. Dann gilt

$$\text{im}(T^*) = \langle T^*w_1, \dots, T^*w_m \rangle = \langle u_1, \dots, u_r \rangle.$$

Weiterhin sind die $\{u_i\}_{i=1}^r$ paarweise orthogonal: es gilt

$$\langle u_i, u_j \rangle_V = \langle T^*w_i, T^*w_j \rangle_V = \langle w_i, TT^*w_j \rangle_W = \lambda_j \langle w_i, w_j \rangle_W = \lambda_j \cdot \delta_{ij}.$$

Mit anderen Worten, $u_1, \dots, u_r$ sind eine orthogonale Basis von $\text{im}(T^*)$, und

$$\|u_i\|_V = \sqrt{\langle u_i, u_i \rangle_V} = \sqrt{\lambda_i} = \sigma_i.$$

Weiterhin folgt von Lemma 13.9.9, dass

$$\langle u_1, \dots, u_r \rangle^\perp = \text{im}(T^*)^\perp = \ker(T)$$

und daher

$$V = \langle u_1, \dots, u_r \rangle \oplus \ker(T).$$

Beachte, dass

$$\dim \ker(T) = \dim V - \text{rk}(T) = n - r.$$

Es sei $u_{r+1}, \dots, u_n$ eine orthogonale Basis von $\ker(T)$. Dann ist $u_1, \dots, u_n$ eine orthogonale Basis von V , und durch Normalisierung erhalten wir eine ON Basis $v_1, \dots, v_n$ mit der Eigenschaft, dass

$$T(v_i) = T\left(\frac{u_i}{\|u_i\|}\right) = \begin{cases} \sigma_i^{-1} T(u_i) = \sigma_i^{-1} \cdot TT^*(w_i) = \frac{\lambda_i}{\sigma_i} \cdot w_i = \sigma_i w_i & \text{fuer } 1 \leq i \leq r \\ \|u_i\|^{-1} \cdot T(u_i) = 0_W & \text{fuer } i > r. \end{cases}$$

□

Bemerkung 15.3.7. Zwei Matrizen $A, B \in M_{n \times n}(\mathbb{R})$ sind orthogonal aehnlich, wenn es eine Matrix $P \in O(n)$ gibt, so dass $B = P^{-1}AP$. Was sind die Aequivalenzklassen?

16. DIREKTE SUMMEN UND TENSORPRODUKTE VON VEKTORRAEUMEN

16.1. The aeussere direkte Summe zweier Vektorraeume. Es sei V ein Vektorraum ueber K , uns es seien U, W Unterraume mit $U \cap W = \{0_V\}$. Wir haben bereits gesehen, dass wir einen neuen Vektorraum $U \oplus W$ von V bilden koennen, dessen Dimension gleich $\dim U + \dim W$ ist. Die definierende Eigenschaft von $U \oplus W$ ist, dass sich jeder Vektor $v \in U \oplus W$ eindeutig als Summe $v = u + w$ mit $u \in U$ und $w \in W$ schreiben laesst. Mit anderen Worten, wir koennen v mit dem geordneten Paar (u, w) identifizieren. Diese Idee wollen wir abstrahieren.

Definition 16.1.1. Es seien V, W Vektorraeume ueber K . Wir definieren die aeussere direkte Summe⁷³ $V \oplus W$ von V und W wie folgt: die Vektoren sind die geordneten Paare (v, w) mit $v \in V$ und $w \in W$.⁷⁴ Mit den Operationen

$$\begin{aligned}(v_1, w_1) + (v_2, w_2) &= (v_1 + v_2, w_1 + w_2) \\ \alpha \cdot (v, w) &= (\alpha v, \alpha w)\end{aligned}$$

erhaelt $V \oplus W$ die Struktur eines K -Vektorraums.

Wir schreiben oft $v \oplus w$ anstatt (v, w) , um die Vektorraumstruktur zu betonen.

Beispiele 16.1.2.

- (1) Wir betrachten $\mathbb{R}$ als einen 1-dimensionalen Vektorraum ueber sich selber. Dann ist $\mathbb{R} \oplus \mathbb{R}$ die Menge aller geordneten Paare $x \oplus y$ mit $x, y \in \mathbb{R}$. Die Abbildung $x \oplus y \mapsto \begin{pmatrix} x \\ y \end{pmatrix}$ ist ein kanonischer Isomorphismus $\mathbb{R} \oplus \mathbb{R} \rightarrow \mathbb{R}^2$.
- (2) Es sei V ein Vektorraum und $U, W \leq V$ mit $U \cap W = \{0_V\}$. Dann ist die aeussere direkte Summe von U und W isomorph zu der direkten Summe von U und W als ein Unterraum von V ; der Isomorphismus ist gegeben durch $v \oplus w \mapsto u + w$.

Bemerkung 16.1.3. Wenn $v_1, \dots, v_n$ eine Basis von V ist und $w_1, \dots, w_m$ eine Basis von W , dann ist

$$v_1 \oplus 0_W, \dots, v_n \oplus 0_W, 0_V \oplus w_1, \dots, 0_V \oplus w_m$$

eine Basis von $V \oplus W$: die Dimension von $V \oplus W$ ist $n + m$.

Lemma 16.1.4. Die Abbildungen

$$\begin{aligned}\iota_V : V &\rightarrow V \oplus W & v &\mapsto v \oplus 0_W \\ \iota_W : W &\rightarrow V \oplus W & w &\mapsto 0_V \oplus w\end{aligned}$$

identifizieren V (bzw. W) mit Unterraumen von $V \oplus W$, das heisst, die Abbildungen ι_V und ι_W sind Injektionen. Weiterhin gilt

$$\text{im}(\iota_V) \cap \text{im}(\iota_W) = \{0_V \oplus 0_W\},$$

und die natuerliche Abbildung

$$\iota_V \oplus \iota_W : V \oplus W \rightarrow \iota_V(V) \oplus \iota_W(W)$$

ist ein kanonischer Isomorphismus, d.h. sie identifiziert die aeussere direkte Summe von V und W mit der direkten Summe der Unterraume $\iota_V(V)$ und $\iota_W(W)$.

16.2. Vektorraeume ueber einer freien Menge.

Definition 16.2.1. Es sei S eine (endliche oder unendliche) Menge und K ein Koerper. Der von S erzeugte K -Vektorraum $K(S)$ ist definiert wie folgt:

- Die Elemente von $K(S)$ sind formale Summen

$$v = \sum_{s \in S} \alpha_s \cdot s,$$

wobei nur fuer endlich viele Element $\alpha_s \neq 0$;

- Vektoraddition ist definiert als

$$\sum_{s \in S} \alpha_s \cdot s + \sum_{s \in S} \beta_s \cdot s = \sum_{s \in S} (\alpha_s + \beta_s) \cdot s;$$

⁷³Das Adjektiv *aeussere* bezieht sich darauf, dass a priori V und W keine Unterraume eines gemeinsamen Vektorraums sind.

⁷⁴Mit anderen Worten, die Vektoren von $V \oplus W$ sind die Elemente des direkten Produktes $V \times W$. Allerdings hat $V \times W$ keine Vektorraumstruktur.

- Skalar Multiplikation ist definiert als

$$\beta \left(\sum_{s \in S} \alpha_s \cdot s \right) = \sum_{s \in S} (\beta \alpha_s) \cdot s.$$

Bemerkung 16.2.2.

- (1) Wenn S endlich ist, dann ist die Dimension von $K(S)$ gleich $|S|$, und $\{s \in S\}$ ist eine Basis von $K(S)$.
- (2) S kann hier jede beliebige Menge sein (auch Krokodile oder Studenten).

Beispiel 16.2.3. Es sei $S = \{1, x, x^2, \dots, x^k\}$. Dann ist $K(S) = K[x]^{\leq k}$.

Bemerkung 16.2.4. Alternativ kann man $K(S)$ auch als Vektorraum von Funktionen definieren:

$$K(S) = \{f : S \rightarrow K : f(s) \neq 0 \text{ fuer nur endlich viele } s\}.$$

In jeder der beiden (aequivalenten) Definitionen spielt S die Rolle einer Indexmenge.

16.3. Konstruktion des Tensorprodukts. Wir koennen nun das Tensorprodukt von zwei endlich-dimensionalen Vektorraeumen definieren.

Definition 16.3.1. Es seien V und W endlich-dimensionale Vektorraeume ueber K , mit jeweiligen Basen $v_1, \dots, v_n$ und $w_1, \dots, w_m$. Es sei S die Menge aller Symbole

$$S = \{v_i \otimes w_j : 1 \leq i \leq n, 1 \leq j \leq m\}.$$

Wir definieren das Tensoprodukt $V \otimes_K W$ als den Vektorraum $K(S)$. Mit anderen Worten, die Elemente von $V \otimes_K W$ sind Linearkombinationen der Form

$$\sum_{i,j} a_{ij} \cdot v_i \otimes w_j \quad \text{mit } a_{ij} \in K.$$

Bemerkung 16.3.2. Es ist offensichtlich von der Definition, dass $\dim_K V \otimes_K W = mn$.

Beispiele 16.3.3.

- (1) Es gilt $K \otimes_K K \cong K$: eine Basis fuer K ist gegeben durch $1 \otimes 1$ (zum Beispiel), und die lineare Abbildung

$$K \rightarrow K \otimes_K K, \quad x \mapsto x \cdot (1 \otimes 1)$$

ist ein Isomorphismus von K -Vektorraeumen.

- (2) Es sei V ein endlich-diemsnionaler K -Vektorraum. Dann gilt

$$V \otimes_K K \cong V.$$

- (3) Fuer jeden K -Vektorraum gilt

$$V \otimes_K \{0\} \cong \{0\}.$$

- (4) Es sei V ein n -dimensionaler $\mathbb{R}$ -Vektorraum mit Basis $v_1, \dots, v_n$. Betrachte $\mathbb{C}$ als einen $\mathbb{R}$ -Vektorraum mit Basis $1, i$. Dann ist $V \otimes_{\mathbb{R}} \mathbb{C}$ ein $2n$ -dimensionaler $\mathbb{R}$ -Vektorraum mit Basis

$$v_1 \otimes 1, \dots, v_n \otimes 1, v_1 \otimes i, \dots, v_n \otimes i.$$

Beachte 16.3.4. Wir koennen eine Abbildung

$$\otimes : V \times W \rightarrow V \otimes_K W$$

folgendermassen definieren: fuer $v = \sum_i a_i v_i$ und $w = \sum_j b_j w_j$ definieren wir

$$v \otimes w = \sum_{i,j} a_i b_j \cdot (v_i \otimes w_j).$$

Die Abbidlugn $\otimes$ ist bilinear:

- $\forall v_1, v_2 \in V$ und $\forall w_1, w_2 \in W$ gilt

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w \quad \text{und} \quad v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2;$$

- fuer alle $\alpha \in K$, $v \in V$ und $w \in W$ gilt⁷⁵

$$(\alpha v) \otimes w = v \otimes (\alpha w) = \alpha \cdot v \otimes w.$$

⁷⁵Ich stelle mir das Tensorsymbol $\otimes$ als eine Art Sperrre vor, durch die sich nur Elemente von K von der einen auf die andere Seite schieben lassen, aber keine Vektoren.

Bemerkung 16.3.5. Elemente der Form $v \otimes w$ heissen reine Tensoren. Vorsicht: nicht jedes Element von $V \otimes_K W$ ist ein reiner Tensor! Es ist eine endliche Summe endlicher Tensoren, aber diese Darstellung ist im Allgemeinen nicht eindeutig. Folgendes Beispiel eines Tensors, der kein reiner Tensor ist, hat Lukas Achermann gefunden:

Wir betrachten $V \otimes V$ mit $V \cong \mathbb{R}^2$ mit Basis $\{e_1, e_2\}$. Sei des Weiteren $T \in V \otimes_{\mathbb{R}} V$ gegeben durch:

$$T = e_1 \otimes e_1 + e_2 \otimes e_2$$

Behauptung: T ist ein kein reiner Tensor.

Um diese Behauptung zu beweisen argumentieren wir per Widerspruch.

Angenommen T wäre rein, dann müsste $T = v \otimes w$, für $v = ae_1 + be_2$ und $w = ce_1 + de_2$.

Dann gilt:

$$\begin{aligned} T &= v \otimes w \\ &= (ae_1 + be_2) \otimes (ce_1 + de_2) \\ &= ac \cdot e_1 \otimes e_1 + ad \cdot e_1 \otimes e_2 + bc \cdot e_2 \otimes e_1 + bd \cdot e_2 \otimes e_2 \end{aligned}$$

Verglichen wir dies mit $T = e_1 \otimes e_1 + e_2 \otimes e_2$ ergibt das folgende Koeffizienten:

$$ac = 1, \quad bd = 1, \quad ad = 0, \quad bc = 0$$

- Aus $ad = 0$ folgt $a = 0$ oder $d = 0$
- Falls $a = 0$: Dann ist $ac = 0$, jedoch ist $ac = 1 \rightarrow$ Widerspruch!
- Falls $d = 0$: Dann ist $bd = 0$, jedoch ist $bd = 1 \rightarrow$ Widerspruch!

Da dies zu einem Widerspruch führt, folgt das T kein reiner Tensor ist, womit wir gezeigt haben, dass nicht jedes Element von $V \otimes V$ ein reiner Tensor ist.

Die Definition 16.3.1 ist ()zunaechst) etwas unbefriedigend, weil sie von der Wahl von Basen der beiden Vektorraeume V und W abhaengt. Wir werden mit Hilfe des universelle Eigenschaft des Tensorproduktes zeigen, dass die Definition unabhaengig von der Wahl der Basen ist. Wir erinnern uns zunaechst an die Definition einer bilinearen Abbildung:

Definition 16.3.6. Es seien V, W, X K -Vektorraeume. Eine Abbildung $\Phi : V \times W \rightarrow X$ ist bilinear, wenn folgende Bedingungen erfuehlt sind:

- (1) $\Phi(\alpha v_1 + \beta v_2, w) = \alpha \Phi(v_1, w) + \beta \Phi(v_2, w)$;
- (2) $\Phi(v, \alpha w_1 + \beta w_2) = \alpha \Phi(v, w_1) + \beta \Phi(v, w_2)$.

Bemerkung 16.3.7. Es seien $v_1, \dots, v_n$ und $w_1, \dots, w_m$ Basen von V und W . Dann ist die bilineare Abbildung Φ durch die Werte $\{\Phi(v_i, w_j) : 1 \leq i \leq n, 1 \leq j \leq m\}$ vollstaendig bestimmt.

Satz 16.3.8. Es seien U, V, W K -Vektorraeume, und $v_1, \dots, v_n$ und $w_1, \dots, w_m$ jeweilige Basen von V und W . Es sei $\Phi : V \times W \rightarrow U$ eine bilineare Abbildung. Dann gibt es genau eine lineare Abbildung

$$\phi : V \otimes W \rightarrow U,$$

so dass das folgende Diagramm kommutiert:

$$\begin{array}{ccc} V \times W & \xrightarrow{\Phi} & U \\ \downarrow \otimes & \searrow \phi & \\ V \otimes W & & \end{array}$$

Mit anderen Worten, es gilt $\Phi = \phi \circ \otimes$.

Proof. Wir definieren die lineare Abbildung ϕ durch

$$\phi \left(\sum_{i,j} c_{ij} \cdot (v_i \otimes w_j) \right) = \sum_{i,j} c_{ij} \Phi(v_i \otimes w_j).$$

Wir pruefen, dass das Diagramm kommutiert: da sowohl Φ und $\phi \circ \otimes$ bilinear sind, reicht es aufgrund von Bemerkung 16.3.7 zu zeigen, dass

$$\Phi(v_i, w_j) = \phi(v_i \otimes w_j) \quad \forall i, j.$$

Doch das ist offensichtlich.

Da eine lineare Abbildung durch die Bilder einer Basis eindeutig festgelegt ist, ist ϕ eindeutig definiert. $\square$

Wir koennen nun zeigen, dass Definition 16.3.1 unabhaengig von der Wahl der Basen von V und W ist.

Lemma 16.3.9. *Es sei X ein K -Vektorraum mit einer bilinearen Abbildung $\tau : V \times W \rightarrow X$, der (statt $V \otimes W$) auch die universelle Eigenschaft in Satz 16.3.8 erfuehlt. Dann ist X kanonisch isomorph zu $V \otimes W$.*

Proof. Wir benutzen zunaechst die universelle Eigenschaft von $V \otimes W$ und wenden sie auf $U = W$ an: dann erhalten wir von Satz 16.3.8 das kommutative Diagramm

$$\begin{array}{ccc} V \times W & \xrightarrow{\tau} & X \\ \downarrow \otimes & \nearrow \phi & \\ V \otimes W & & \end{array}$$

Wenn wir hingegen die universelle Eigenschaft von X betrachten und auf $U = V \otimes W$ anwenden, erhalten wir ein kommutatives Diagramm

$$\begin{array}{ccc} V \times W & \xrightarrow{\otimes} & V \otimes W \\ \downarrow \tau & \nearrow \phi & \\ X & & \end{array}$$

Insbesondere erhalten wir also Abbildungen

$$\phi : V \otimes W \rightarrow X \quad \text{und} \quad \tilde{\phi} : X \rightarrow V \otimes W.$$

Ich behaupte, dass ϕ ein Isomorphismus ist, mit $\phi^{-1} = \tilde{\phi}$. Wir betrachten zunaechst die Abbildung $\tilde{\phi} \circ \phi$, durch die das Diagramm

$$\begin{array}{ccc} V \times W & \xrightarrow{\otimes} & V \otimes W \\ \downarrow \otimes & \nearrow \tilde{\phi} \circ \phi & \\ V \otimes W & & \end{array}$$

kommutiert. Aber die Identitaetsabbildung $\mathbf{1}_{V \otimes W}$ (statt $\tilde{\phi} \circ \phi$) laesst dieses Diagramm ebenfalls kommutieren. Aufgrund der Eindeutigkeit in der universellen Eigenschaft folgt daher $\tilde{\phi} \circ \phi = \mathbf{1}_{V \otimes W}$. Ebenso zeigen wir, dass $\phi \circ \tilde{\phi} = \mathbf{1}_X$. Das Resultat folgt. $\square$

Als direkte Konsequenz erhalten wir folgenden Satz:

Satz 16.3.10. *Die Definition 16.3.1 ist unabhaengig von der Wahl der Basen.*

Bemerkung 16.3.11. *Alternativ kann man $V \otimes W$ auch folgendermassen definieren: betrachte den (riesigen) Vektorraum $K(V \times W)$. Es sei*

$$(69) \quad R = \left\{ \begin{array}{l} \{(v_1 + v_2, w) - (v_1, w) - (v_2, w) : v_1, v_2 \in V, w \in W\} \\ \cup \{(v, w_1 + w_2) - (v, w_1) - (v, w_2) : v \in V, w_1, w_2 \in W\} \\ \cup \{(\alpha v, w) - \alpha(v, w) : \alpha \in K, v \in V, w \in W\} \\ \cup \{(v, \alpha w) - \alpha(v, w) : \alpha \in K, v \in V, w \in W\} \end{array} \right\}$$

und $X = \left(\begin{array}{c} \text{---} \end{array} \right) R \left(\begin{array}{c} \text{---} \end{array} \right) \leq K(V \otimes W)$. Definiere $V \otimes_K W$ als den Quotientenraum

$$V \otimes_K W = K(V \otimes W)/X.$$

Diese Definition hat den Vorteil, dass man zu keinem Zeitpunkt Basen waehlen muss; allerdings ist sie zum Rechnen sehr unpraktisch.

16.4. Erste Eigenschaften.

Satz 16.4.1. *Es seien $v \in V$, $w \in W$, keiner von beiden gleich dem Nullvektor. Dann ist $v \otimes w \neq 0_{V \otimes_K W}$.*

Proof. Die Idee ist es, eine bilineare Abbildung

$$\Phi : V \times W \rightarrow K$$

zu konstruieren, fuer die gilt $\Phi(v, w) \neq 0$.⁷⁶ Dann gibt es eine lineare Abbildung $\phi : V \otimes_K W \rightarrow K$, so dass das Diagramm aus Satz 16.3.8 kommutiert; insbesondere folgt daraus, dass $\phi(v \otimes w) \neq 0$ und daher $v \otimes w$ nicht der Nullvektor sein kann.

Da v (bzw. w) nicht der Nullvektor ist, gibt es ein $f \in V^*$ (bzw. $g \in W^*$), so dass $f(v) \neq 0$ (bzw. $g(w) \neq 0$.) Definiere

$$\Phi : V \times W \rightarrow K, \quad (x, y) \mapsto f(x)g(y).$$

□

Satz 16.4.2. *Die Vektorraeume $V \otimes_K W$ und $W \otimes_K V$ sind kanonisch isomorph.*

Proof. Wir benutzen wieder die universelle Eigenschaft. Definiere

$$\Phi : V \times W \rightarrow W \otimes_K V, \quad (v, w) \mapsto w \otimes v.$$

Dann ist Φ bilinear, und wir erhalten eine lineare Abbildung

$$\phi : V \otimes_K W \rightarrow W \otimes_K V,$$

so dass das Diagramm

$$\begin{array}{ccc} V \times W & \xrightarrow{\Phi} & W \otimes V \\ \downarrow \otimes & \nearrow \phi & \\ V \otimes W & & \end{array}$$

kommutiert; insbesondere gilt $\phi(v \otimes w) = w \otimes v$. Wir erhalten analog eine lineare Abbildung

$$\tilde{\phi} : W \otimes_K V \rightarrow V \otimes_K W,$$

so dass

$$\begin{array}{ccc} W \times V & \xrightarrow{\Phi} & V \otimes_K W \\ \downarrow \otimes & \nearrow \tilde{\phi} & \\ W \otimes_K V & & \end{array}$$

kommutiert, d.h. $\tilde{\phi}(w \otimes v) = v \otimes w$. Es folgt, dass

$$\phi \otimes \tilde{\phi} = \text{id}_{W \otimes V} \quad \text{und} \quad \tilde{\phi} \circ \phi = \text{id}_{V \otimes W}.$$

□

Folgende Eigenschaften des Tensorprodukts lassen sich ebenfalls mit Hilfe der universellen Eigenschaft ueberpruefen:

Lemma 16.4.3. *Es seien U, V, W endlich-dimensionale K -Vektorraeume. Dann gibt es kanonische Isomorphismen*

- (1) $U \otimes (V \otimes W) \cong (U \otimes V) \otimes W$;
- (2) $U \otimes (V \oplus W) \cong (U \otimes V) \oplus (U \otimes W)$.

Proof. Uebung. □

⁷⁶Natuerlich koennen Sie es alternativ auch direkt beweisen, indem Sie Basen waehlen.

16.5. Beispiel: Komplexifizierung. Es sei V ein endlich-dimensionaler $\mathbb{R}$ -Vektorraum. Wir betrachten das Tensorprodukt $V \otimes_{\mathbb{R}} \mathbb{C}$ als $\mathbb{R}$ -Vektorraum. Es sei $v_1, \dots, v_n$ eine Basis von V . Dann ist eine $\mathbb{R}$ -Basis von $V \otimes_{\mathbb{R}} \mathbb{C}$ gegeben durch $\{v_j \otimes 1, v_k \otimes i : 1 \leq j, k \leq n\}$. Weiterhin koennen wir eine Einbettung von V in $V \otimes \mathbb{C}$ definieren durch $i_V : v \mapsto v \otimes 1$.

Definition 16.5.1. Wir definieren eine $\mathbb{C}$ -Skalarmultiplikation auf $V \otimes_{\mathbb{R}} \mathbb{C}$ wie folgt: es sei $v \otimes \alpha \in V \otimes_{\mathbb{R}} \mathbb{C}$ und $\beta \in \mathbb{C}$. Definiere⁷⁷

$$\beta \cdot (v \otimes \alpha) = v \otimes (\beta \alpha).$$

Dadurch erhaelt $V \otimes_{\mathbb{R}} \mathbb{C}$ die Struktur eines $\mathbb{C}$ -Vektorraums⁷⁸.

Bemerkung 16.5.2. Es sei $u \otimes 1 + v \otimes i \in V \otimes_{\mathbb{R}} \mathbb{C}$ und $a + bi \in \mathbb{C}$. Dann gilt⁷⁹

$$\begin{aligned} (a + bi)(u \otimes 1 + v \otimes i) &= a(u \otimes 1) + a(v \otimes i) + u \otimes (bi) + v \otimes (bi^2) \\ &= a(u \otimes 1) + a(v \otimes i) + b(u \otimes i) - b(v \otimes 1) \\ &= (au - bv) \otimes 1 + (bu + av) \otimes i. \end{aligned}$$

Lemma 16.5.3. Es gilt $\dim_{\mathbb{C}} V \otimes_{\mathbb{R}} \mathbb{C} = n$. Weiterhin sind die Elemente $v_1 \otimes 1, \dots, v_n \otimes 1$ eine $\mathbb{C}$ -Basis von $V \otimes_{\mathbb{R}} \mathbb{C}$.⁸⁰

Proof. Uebung. □

Bemerkung 16.5.4. Definition 16.5.1 ist ein Beispiel einer viel allgemeineren Konstruktion. Es sei V ein K -Vektorraum und L ein Koerper, der K enthaelt und ein endlich-dimensionaler K -Vektorraum ist.⁸¹ Dann besitzt $V \otimes_K L$ die Struktur eines L -Vektorraums: fuer $v \otimes \alpha \in V \otimes_K L$ und $\beta \in L$ definieren wir $\beta \cdot (v \otimes \alpha) = v \otimes (\alpha \beta)$. Wir nennen diesen L -Vektorraum auch die Erweiterung der Skalare von V von K nach L .

16.6. Tensorprodukte linearer Abbildungen. Das Tensorprodukt ist ebenfalls mit linearen Abbildungen kompatibel:

Satz 16.6.1. Es seien V, V', W und W' endlich-dimensionale K -Vektorraeume, und es seien $T : V \rightarrow V'$ und $S : W \rightarrow W'$ lineare Abbildungen. Dann existiert eine eindeutige lineare Abbildung

$$T \otimes S : V \otimes W \rightarrow V' \otimes W',$$

so dass

$$(T \otimes S)(v \otimes w) = T(v) \otimes S(w) \quad \forall v \in V, w \in W.$$

Proof. Es seien $v_1, \dots, v_n$ und $w_1, \dots, w_m$ jeweilige Basen von V und W . Wir definieren $T \otimes S$ durch

$$(T \otimes S)(v_i \otimes w_j) = T(v_i) \otimes S(w_j).$$

Alternativ koennen wir auch die universelle Eigenschaft des Tensorprodukts benutzen: definiere die Bilinearform

$$\Phi : V \times W \rightarrow V' \otimes W' \quad (v, w) \mapsto T(v) \otimes S(w).$$

Dann erhalten wir laut Satz 16.3.8 eine Abbildung $\phi : V \otimes W \rightarrow V' \otimes W'$, so dass folgendes Diagramm kommutiert:

$$\begin{array}{ccc} V \times W & \xrightarrow{\Phi} & V' \otimes W' \\ \downarrow \otimes & \nearrow \phi & \\ V \otimes W & & \end{array}$$

Wir setzen $T \otimes S = \phi$. □

⁷⁷Jedes Element von $V \otimes_{\mathbb{R}} \mathbb{C}$ laesst sich als eine endlich Summe reiner Tensoren ausdruecken; daher erhalten wir eine $\mathbb{C}$ -Skalarmultiplikation auf $V \otimes_{\mathbb{R}} \mathbb{C}$.

⁷⁸was sich anhand der Axiome leicht ueberpruefen laesst.

⁷⁹Vergl. mit Definition 19.2.2.

⁸⁰Vergl. mit Satz 19.2.6.

⁸¹Koerper dieser Art werden Sie in der Galois-Theorie und in der algebraischen Zahlentheorie kennenlernen.

Beispiel 16.6.2. Es sei $V = \mathbb{R}^2$, und $T = T_A : V \rightarrow V$ und $S = S_B : V \rightarrow V$ gegeben durch

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}.$$

Was ist die Abbildung $T \otimes S : V \otimes V \rightarrow V \otimes V$? Es sei e_1, e_2 die Standardbasis von V . Dann ist

$$\mathcal{B} = \{e_1 \otimes e_1, e_1 \otimes e_2, e_2 \otimes e_1, e_2 \otimes e_2\}$$

eine Basis von $V \otimes V$. Was ist die Abbildungsmatrix von $T \otimes S$ bezüglich dieser Basis? Wir berechnen

$$\begin{aligned} (T \otimes S)(e_1 \otimes e_1) &= T(e_1) \otimes S(e_1) \\ &= (ae_1 + ce_2) \otimes (a'e_1 + b'e_2) \\ &= aa'(e_1 \otimes e_1) + ab'(e_1 \otimes e_2) + a'b(e_2 \otimes e_1) + bb'(e_2 \otimes e_2) \\ (T \otimes S)(e_1 \otimes e_2) &= \dots \end{aligned}$$

Die Abbildungsmatrix von $T \otimes S$ ist gegeben durch

$$[T \otimes S]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} a \cdot B & b \cdot B \\ c \cdot B & d \cdot B \end{pmatrix}.$$

Lemma 16.6.3. Es seien $T : V \rightarrow V$ und $S : W \rightarrow W$ lineare Abbildungen. Es gilt $\text{Tr}(T \otimes S) = \text{Tr}(T) \cdot \text{Tr}(S)$.

Proof. Übung. □

Bemerkung 16.6.4. Können Sie $\det(T \otimes S)$ mit Hilfe von $\det(T)$ und $\det(S)$ ausdrücken?

Lemma 16.6.5. Es seien $U_1, U_2, V_1, V_2, W_1, W_2$ endlich-dimensionale K -Vektorräume und

$$T_i : U_i \rightarrow V_i \quad \text{und} \quad S_i : V_i \rightarrow W_i$$

lineare Abbildungen. Dann gilt

$$(S_1 \otimes S_2) \circ (T_1 \otimes T_2) = (S_1 \circ T_1) \otimes (S_2 \circ T_2).$$

Proof. Es sei $u_i \in U_i$ fuer $i = 1, 2$. Dann gilt

$$\begin{aligned} (S_1 \otimes S_2) \circ (T_1 \otimes T_2)(u_1 \otimes u_2) &= (S_1 \otimes S_2)(T_1(u_1) \otimes T_2(u_2)) \\ &= (S_1 T_1(u_1)) \otimes (S_2 T_2(u_2)) \\ &= ((S_1 \circ T_1) \otimes (S_2 \circ T_2))(u_1 \otimes u_2). \end{aligned}$$

Da die reinen Tensoren ein Erzeugendensystem von $U \otimes U'$ sind, gilt

$$(S_1 \otimes S_2) \circ (T_1 \otimes T_2) = (S_1 \circ T_1) \otimes (S_2 \circ T_2).$$

□

Als einfache Konsequenz erhalten wir folgenden Satz:

Satz 16.6.6. Es seien V, V', W und W' endlich-dimensionale K -Vektorräume, und es seien $T : V \rightarrow V'$ und $S : W \rightarrow W'$ Isomorphismen. Dann ist $T \otimes S$ ein Isomorphismus, und es gilt

$$(T \otimes S)^{-1} = T^{-1} \otimes S^{-1}.$$

16.7. Tensorprodukte und duale Abbildungen.

Theorem 16.7.1. Es seien U, V endlich-dimensionale K -Vektorräume. Dann existiert genau ein Isomorphismus

$$\chi : U^* \otimes V^* \rightarrow (U \otimes V)^*$$

so dass

$$\chi(f \otimes g)(u \otimes v) = f(u)g(v)$$

fuer alle $f \in U^*, g \in V^*$ und $u \otimes v \in U \otimes V$.

Proof. Wir benutzen die universelle Eigenschaft des Tensorprodukts: es seien $f \in U^*, g \in V^*$. Definiere

$$\Phi_{f,g} : U \times V \rightarrow K, \quad (u, v) \mapsto f(u)g(v).$$

Eine einfache Rechnung zeigt, dass Φ bilinear ist. Daher existiert eine lineare Abbildung⁸²

$$\phi_{f,g} : U \otimes V \rightarrow K$$

so dass

$$\phi_{f,g}(u \otimes v) = \Phi_{f,g}(u, v) = f(u)g(v).$$

Mit anderen Worten, wir haben eine Abbildung

$$(70) \quad \begin{aligned} \Psi : U^* \times V^* &\rightarrow (U \otimes V)^*, \\ (f, g) &\mapsto \phi_{f,g}. \end{aligned}$$

Dann ist Ψ bilinear: eine einfache Rechnung (Uebung) zeigt, dass

$$\Psi(\alpha_1 f_1 + \alpha_2 f_2, g) = \alpha_1 \Psi(f_1, g) + \alpha_2 \Psi(f_2, g),$$

und aehnlich fuer die zweite Variable. Wir erhalten daher eine lineare Abbildung

$$\chi : U^* \otimes V^* \rightarrow (U \otimes V)^*,$$

so dass

$$\chi(f \otimes g) = \Psi(f, g) = \phi_{f,g}.$$

Behauptung: χ ist ein Isomorphismus.

Beweis der Behauptung: es seien $\mathcal{B} = (u_1, \dots, u_m)$ und $\mathcal{C} = (v_1, \dots, v_n)$ Basen von U und V . Dann ist

$$\mathcal{D} = \{u_i \otimes v_j : 1 \leq i \leq m, 1 \leq j \leq n\}$$

eine Basis von $U \otimes V$ und

$$\{\chi(u_i^* \otimes v_j^*) : 1 \leq i \leq m, 1 \leq j \leq n\}$$

hat die Eigenschaft, dass

$$\chi(u_i^* \otimes v_j^*)(u_k \otimes v_\ell) = u_i^*(u_k)v_j^*(v_\ell) = \delta_{i,k} \cdot \delta_{j,\ell},$$

das heisst, es ist die duale Basis zu $\mathcal{D}$. Mit anderen Worten, χ bildet eine Basis von $U^* \otimes V^*$ auf eine Basis von $(U \otimes V)^*$ ab und ist daher ein Isomorphismus. $\square$

Bemerkung 16.7.2. Alternativ kann man zeigen, dass $(U \otimes V)^*$ zusammen mit der natuerlichen Abbildung

$$\Psi : U^* \times V^* \rightarrow (U \otimes V)^*, \quad (f, g) \mapsto \phi_{f,g}$$

(siehe (70)) die universelle Eigenschaft von $U^* \otimes V^*$ erfuehlt. Dieses Argument hat den Vorteil, dass wir keine Basis waehlen muessen.

Mit einem aehnlichen Argument kann man folgenden Satz zeigen:

Satz 16.7.3. Es seien U, V endlich-dimensionale K -Vektorraeume. Dann existiert genau eine lineare Abbildung

$$\Theta : U^* \otimes V \rightarrow \text{Hom}(U, V)$$

mit folgender Eigenschaft: fuer all $f \in U^*$ und $v \in V$ gilt

$$\Theta(f \otimes v)(u) = f(u) \cdot v.$$

Diese Abbildung ist ein kanonischer Isomorphismus, das heisst

$$\text{Hom}(U, V) \cong U^* \otimes V.$$

Proof. Uebung. $\square$

Korollar 16.7.4. Es sei V ein endlich-dimensionaler K -Vektorraum. Dann gibt es einen kanonischen Isomorphismus

$$\text{End}(V) \cong V \otimes V^*.$$

⁸²d.h. $\phi_{f,g} \in (U \otimes V)^*$

16.8. Darstellung von Tensoren (fuer Physiker).

Bemerkung 16.8.1. Es seien V und W endlich-dimensionale Vektorraume mit jeweiligen Basen $v_1, \dots, v_n$ und $w_1, \dots, w_m$. Dann kann ein beliebiges Element $x \in V \otimes W$ dargestellt werden als

$$x = \sum_{i,j} \alpha_{ij} v_i \otimes w_j.$$

Mit anderen Worten, nachdem Sie Basen von V und W festgelegt haben, koennen Sie x als eine Koeffizientenmatrix darstellen:

$$M_x = \begin{pmatrix} \alpha_{1,1} & \dots & \alpha_{1,n} \\ \vdots & \ddots & \vdots \\ \alpha_{m,1} & \dots & \alpha_{m,n} \end{pmatrix}$$

Beispiel 16.8.2. Es sei $V = \mathbb{R}^2$ mit der Standardbasis e_1, e_2 , und es sei $x \in V^{\otimes 2}$ gegeben durch

$$x = -e_1 \otimes e_1 + 2e_1 \otimes e_2 - 2e_2 \otimes e_1 + e_2 \otimes e_2.$$

Dann koennen wir x durch die Koeffizientenmatrix

$$M_x = \begin{pmatrix} -1 & 2 \\ -2 & -1 \end{pmatrix}$$

darstellen.

In der Physik werden Sie das Konzept eines (p, q) -Tensors kennengelernt haben:

Definition 16.8.3. Es sei V ein Vektorraum ueber K , und es seien $p, q \geq 0$. Ein (p, q) -Tensor ist ein Element des Vektorraums⁸³ $V^{\otimes p} \otimes (V^*)^{\otimes q}$.

Bemerkung 16.8.4. Es sei $v_1, \dots, v_n$ eine Basis von V und $v_1^*, \dots, v_n^*$ die duale Basis von V^* . Dann ist

$$\left\{ v_{i_1} \otimes v_{i_2} \otimes \dots \otimes v_{i_p} \otimes v_{j_1}^* \otimes v_{j_2}^* \otimes \dots \otimes v_{j_q}^* : 1 \leq i_k, j_\ell \leq n \right\}$$

eine Basis von $V^{\otimes p} \otimes (V^*)^{\otimes q}$. Mit anderen Worten, jedes Element $x \in V^{\otimes p} \otimes (V^*)^{\otimes q}$ kann geschrieben werden als eine Linearkombination

$$x = \sum_{i_k, j_\ell} \alpha_{j_1, \dots, j_q}^{i_1, \dots, i_p} \left(v_{i_1} \otimes v_{i_2} \otimes \dots \otimes v_{i_p} \otimes v_{j_1}^* \otimes v_{j_2}^* \otimes \dots \otimes v_{j_q}^* \right).$$

Die Menge der Elemente

$$\left(\alpha_{j_1, \dots, j_q}^{i_1, \dots, i_p} \right)_{1 \leq i_k, j_\ell \leq n}$$

ist das, was die Physiker einen Tensor nennen. Dieser Tensor (d.h. diese Darstellung des Elements x) haengt aber von der Wahl einer Basis von V ab – **solche Definitionen sehen die Mathematiker sehr ungerne!!**

Bemerkung 16.8.5. Wir unterscheiden hier zwischen upper and lower Indices, die anzeigen, ob ein Element aus V oder V^* beschrieben wird. Auf diese Weise behalten wir den Ueberblick, wie sich der Tensor aendert, wenn wir einen Basiswechsel von V vornehmen, der ja auch einen Basiswechsel von V^* induziert, allerdings mit verschiedenen Basiswechselmatrizen.

Lecture 49

16.9. Das symmetrische und alternierende Produkt. Wir wollen zwei sehr wichtige Unterraeume von $V^{\otimes r}$ definieren, naemlich das symmetrische und das alternierende Produkt.

Bemerkung 16.9.1. Es sei V ein endlich-dimensionaler K -Vektorraum und $r \geq 0$. Es sei $\sigma \in S_r$. Fuer einen reinen Tensor $v_1 \otimes v_2 \otimes \dots \otimes v_r \in V^{\otimes r}$ definieren wir⁸⁴

$$\sigma(v_1 \otimes v_2 \otimes \dots \otimes v_r) = v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \dots \otimes v_{\sigma(r)}.$$

Dieses induziert eine lineare (bijektive) Abbildung

$$T_\sigma : V^{\otimes r} \rightarrow V^{\otimes r}.$$

Definition 16.9.2. Definiere das r -fache symmetrische Produkt von V als

$$\text{Sym}^r V = \{v \in V^{\otimes r} : \sigma(v) = v \quad \forall \sigma \in S_r\} \leq V^{\otimes r}.$$

⁸³Wir definieren $V^{\otimes 0} = (V^*)^{\otimes 0} = K$.

⁸⁴Vorsicht: die Permutationen vertauschen nicht die Basisektoren von V , sondern Komponenten der reinen Tensoren! Daher erhalten wir eine Operation von S_r auf $V^{\otimes r}$, unabhangig von der Dimension von V !

Beispiel 16.9.3. Es sei $V = \mathbb{R}^2$, und wir betrachten $V^{\otimes 3}$. Dann sind

$$x = e_2 \otimes e_2 \otimes e_2 \quad \text{und} \quad y = e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_2$$

ein Element von $\text{Sym}^3 V$, aber

$$z = e_1 \otimes e_2 \otimes e_3 + e_3 \otimes e_1 \otimes e_2 + e_2 \otimes e_3 \otimes e_1 \notin \text{Sym}^3(V).$$

Beispiel 16.9.4. Es sei $V = \mathbb{R}^2$ mit Standardbasis e_1, e_2 . Dann sind

$$(71) \quad e_1 \otimes e_1, e_1 \otimes e_2 + e_2 \otimes e_1, e_2 \otimes e_2$$

linear unabhängige Elemente in $\text{Sym}^2 V$. Ist es eine Basis? Es sei

$$v = ae_1 \otimes e_1 + be_1 \otimes e_2 + ce_2 \otimes e_1 + de_2 \otimes e_2 \in V^{\otimes 2}.$$

Dann ist $v \in \text{Sym}^2 V$ genau dann, wenn

$$v = \sigma(v) \quad \text{für das nicht-triviale Element } \sigma \in S_2,$$

das heisst genau dann, wenn $b = c$. Daher ist (71) eine Basis von $\text{Sym}^2 V$.

Das symmetrische Produkt $\text{Sym}^r V$ lässt sich statt als Unterraum auch als Quotient von $V^{\otimes r}$ definieren.

Definition 16.9.5. Betrachte den Unterraum $U \leq V^{\otimes r}$ gegeben durch

$$U = \left\langle \sum_{\sigma \in S_r} v - \sigma(v) : v \in V^{\otimes r} \right\rangle \leq V^{\otimes r},$$

und definiere

$$S^r V = V^{\otimes r} / U.$$

Satz 16.9.6. Nimm an⁸⁵, dass $\text{char}(K) = 0$. Es sei $\pi : V^{\otimes r} \rightarrow S^r V$ die Quotientenabbildung. Dann gibt die Beschränkung von π auf $\text{Sym}^r V$ einen kanonischen Isomorphismus

$$\pi : \text{Sym}^r V \cong S^r V.$$

Proof. Es ist ausreichend, ein Inverses von π zu konstruieren. Es sei $[w] \in S^r V$ repräsentiert durch ein Element $w \in V^{\otimes r}$. Definiere⁸⁶

$$\Sigma_r([w]) = \frac{1}{r!} \sum_{\sigma \in S_r} \sigma(w).$$

Dann ist Σ_r wohl-definiert (warum?), und es gilt

$$\Sigma_r \circ \pi = \text{id}_{\text{Sym}^r V} \quad \text{und} \quad \pi \circ \Sigma_r = \text{id}_{S^r V}.$$

□

Bemerkung 16.9.7. Der Vektorraum $S^r V$ besitzt eine universelle Eigenschaft. Welche?

Theorem 16.9.8. Es sei V ein n -dimensionaler K -Vektorraum. Dann gilt

$$\dim_K \text{Sym}^r V = \binom{n+r-1}{r}.$$

Proof. Es sei $v_1, \dots, v_n$ eine Basis von V . Dann ist ein Element

$$w = \sum_{i_1, \dots, i_r} \alpha_{i_1, \dots, i_r} \cdot v_{i_1} \otimes \dots \otimes v_{i_r}$$

genau dann in $\text{Sym}^r V$, wenn

$$\alpha_{i_1, \dots, i_r} = \alpha_{\sigma(i_1), \dots, \sigma(i_r)} \quad \forall \sigma \in S_r.$$

Wir definieren nun eine Äquivalenzrelation auf

$$X = \{v_{i_1} \otimes \dots \otimes v_{i_r} : 1 \leq i_1, \dots, i_r \leq n\} :$$

es gilt

$$v_{i_1} \otimes \dots \otimes v_{i_r} \sim v_{j_1} \otimes \dots \otimes v_{j_r}$$

genau dann, wenn es ein $\sigma \in S_r$ gibt, so dass

$$v_{i_1} \otimes \dots \otimes v_{i_r} = v_{\sigma(j_1)} \otimes \dots \otimes v_{\sigma(j_r)}$$

⁸⁵Es ist ausreichend anzunehmen, dass $r!$ in K invertierbar ist.

⁸⁶Die Funktion Σ_r heisst *Symmetrisierung*.

Die Dimension ist dann die Anzahl der Äquivalenzklassen. Nun hat jede Äquivalenzklasse genau ein Element $v_{i_1} \otimes \cdots \otimes v_{i_r}$ mit $i_1 \leq i_2 \leq \cdots \leq i_r$. Mit anderen Worten, die Dimension von $\text{Sym}^r V$ ist gleich der Anzahl k aller r -Tupel

$$\{(i_1, \dots, i_r) : 1 \leq i_1 \leq \cdots \leq i_r \leq n\}.$$

Wieviele solcher Tupel gibt es? Um diese Frage zu beantworten, benutzen wir die sogenannte *stars and bars* Methode aus der Kombinatorik. Wir representieren ein solches r -Tupel durch eine Reihe von r Sternen

$$\star \star \star \star \star \star$$

und fügen $(n-1)$ Striche ein, um die Zahlen festzulegen: der erste Strich teilt die Menge der Einsen und zweien, der zweite die Zweien und Dreien etc. Zum Beispiel, wenn $r = 6$ und $n = 7$, dann representiert

$$| \star || \star \star | \star \star || \star$$

das 6-Tupel und

$$\star ||| \star ||| \star \star \star \star$$

das 6-Tupel

$$(1, 4, 7, 7, 7, 7).$$

Eine solche Reihe von Sternen und Strichen enthaelt als $r + n - 1$ Elemente, und wir waehlen r (bzw. $n - 1$) als Sterne (bzw. Striche) aus. Es gibt daher insgesamt

$$k = \binom{r+n-1}{r}$$

solcher Reihen. □

Beispiel 16.9.9. Es sei V ein n -dimensionaler K -Vektorraum, und betrachte $V \otimes V$ mit der natuerlichen Operation von S_2 (Vertauschung der Komponenten des Tensorprodukts). Es sei $\sigma \in S_2$ das nicht-triviale Element, und

$$\sigma : V \otimes V \rightarrow V \otimes V$$

die induzierte lineare Abbildung. Dann gilt

$$x \in \text{Sym}^2 V \quad \Leftrightarrow \quad \sigma(x) = x,$$

das heisst, wenn x ein Element des $(+1)$ -Eigenraums von σ ist. Es folgt von Theorem 16.9.8, dass dieser Eigenraum Dimension

$$\binom{n+1}{2} = \frac{1}{2}n(n+1)$$

hat. Nun gilt $\sigma^2 = 1$, das heisst, ± 1 sind die einzigen Eigenwerte von σ . Was ist der (-1) -Eigenraum?

Definition 16.9.10. Es sei V ein endlich-dimensionaler K -Vektorraum. Definiere das r -fache alternierende Produkt von V als

$$\text{Alt}^r V = \{v \in V^{\otimes r} : \sigma(v) = \text{sgn}(\sigma)v \quad \forall \sigma \in S^r\} \leq V^{\otimes r}.$$