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 Gehrunger (tim.gehrunger@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

1. Einfuehrung

Lecture 1

1.1. **Motivation: Fibonacci Folgen.** In diesem Kapitel moechte ich Ihnen eine sehr schoene Anwendung der linearen Algebra zeigen, als Motivation fuer die Stukturen, die wie in diesem Kurs untersuchen 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 $(a_n)_{n\geq 0}$ is definiert durch dir Rekursion*

$$a_0 = 0, \quad a_1 = 1,$$
$$a_n = a_{n-1} + a_{n-2} \quad \textit{for } n \geq 2$$

Beachte 1.1.2. *Die ersten Elemente der Folge sind*

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

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*

$$a_0 = a, \quad a_1 = b,$$
$$a_n = a_{n-1} + a_{n-2} \quad \textit{for } n \geq 2$$

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 $\{\mathcal{F}_{\alpha a,\alpha b} \,:\, \alpha \in \mathbb{R}\} \subset V$ unter der Annahme $(a, b) \neq (0, 0)$ unendlich viele Elemente enthaelt, 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*

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

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 einen Abbildung $T : V \to V$, die die Struktur von V respektiert.

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

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

Wir sagen: T ist eine lineare Abbildung.

Beispiele 1.1.12.

(1) Die Identitäts-Abbildung

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

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

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

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

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

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. For $n \geq 0$ sei $b_n = a_{n+1}$; wir muessen 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-Abbildung[1]

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

Dann ist S eine Symmetrie von V.

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

Beispiel 1.1.16. Es gilt

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

Definition 1.1.17. *Set $T : V \to V$ 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 seinen $\varphi = \frac{1+\sqrt{5}}{2}$ 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*

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

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

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 $\mathcal{A} = \beta \cdot \mathcal{F}_{1,\psi}$.

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

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

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

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

Übung 1.1.20. *Zeigen Sie, dass*

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

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 (3) (bzw. (3)) gegeben durch φ^n (bzw. durch ψ^2). 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. □

Wir sehen also, dass wir die Betrachung des reellen Vektorraums V und seiner Symmetries einige sehr interessante (und konkrete) Saetze ueber Fibonacci-Folgen beweisen koennen. Wir werden diesem Phaenomen in diesem Kurs haeufig begegnen: durch sehr abstrakte Argumente erhalten wir informationen ueber 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

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

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: Lecture 2

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

Beispiel 1.2.6. $M = \{x : x \in \mathbb{Z} \quad \text{und} \quad x^2 < 9\}$ ist die Menge $M = \{0, \pm1, \pm2, \pm3\}$.

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 $P \nsubseteq Q$, 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 $\mathbb{Z} \nsubseteq \mathbb{Q}_+$.
(3) Die leere Menge ist eine Teilmenge von jeder Menge.
(4) Es sei $M = \{1, \{1, 2\}\}$. Die Teilmengen von M sind

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

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

Wenn Q eine Teilmenge von P ist, dann nennen wir P one 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 $\mathcal{P} \cap Q = \emptyset$.

---

[2] Wenn 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*

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

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

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

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

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*

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

1.3. **Funktionen.**

Definition 1.3.1. *Es seien X, Y Mengen. Eine Funktion (oder Abbildung) $f : X \to Y$ ist eine Relation, die* jedem *Element $x \in X$* genau ein *Element $f(x) \in Y$ zuordnet; wir schreiben*[3] $x \mapsto f(x)$. *Wir nennen X die* Definitionsmenge *und Y die* Zielmenge *der Funktion.*

Beispiel 1.3.2.
(1) $f : \mathbb{Z} \to \mathbb{Z}$, $n \mapsto n^2$ ist eine Funktion.
(2) $g : \mathbb{Z} \to \{\pm 1\}$, $x \mapsto \begin{cases} 1 & \text{wenn } x > 0 \\ -1 & \text{wenn } x < 0 \end{cases}$ ist keine Funktion.[4]
(3) $g : \mathbb{Z} \to \{\pm 1\}$, $x \mapsto \begin{cases} 1 & \text{wenn } x > 0 \\ -1 & \text{wenn } x \leq 0 \end{cases}$ ist eine Funktion.
(4) $\mathrm{id}_X : X \to X$, $x \mapsto x$ ist die *Identitaetsfunktion von X*.
(5) Es sei $Y \subseteq X$. Die charakteristische Funktion von Y ist die Funktion

$$\mathrm{char}_Y : X \to \{0, 1\},$$
$$x \mapsto \begin{cases} 1 & \text{wenn } x \in Y \\ 0 & \text{wenn } x \notin Y \end{cases}$$

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

Bemerkung 1.3.3. *Zwei Funktionen $f : X \to Y$ und $g : x' \to Y'$ 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 $f : X \to Y$ eine Funktion und $A \subseteq X$. Die* Beschraenkung von f auf A *ist die Funktion*

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

Definition 1.3.5. *Es sei $f : X \to Y$ 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 $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ ist weder injektiv noch surjektiv.
(2) Die Funktion $g : \mathbb{R}_{\geq 0} \to \mathbb{R}$, $x \mapsto x^2$ ist injektiv aber nicht surjektiv.
(3) Die Funktion $h : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, $x \mapsto x^2$ ist bijektiv.

Definition 1.3.7. *Es sei $f : X \to Y$ bijektiv. Die* inverse Funktion $f^{-1} : Y \to X$ *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 Funtion h aus Beispiel 1.3.6 (3) ist gegeben durch

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

Manchmal koennen Funktionen miteinander verknuepft werden:

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

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

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

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

Bemerkung 1.3.11. *Es seien $f : X \to Y$, $g : Y \to Z$ 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 $f, g : \mathbb{R} \to \mathbb{R}$ 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 $f : X \to Y$, $g : Y \to Z$ Funktionen.*
(1) Wenn f und g injektiv sind, dann ist auch $g \circ f$ injektiv.

[3]Auf Vorschlag eines Studenten wird $\mapsto$ gesprochen als *Pfeil mit Fuss*.
[4]Warum nicht?

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

Proof. Uebung. □

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

2. Matrizen und lineare Gleichungssysteme

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

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

$$+ : K \times K \to K \qquad \text{(Addition)}$$
$$\cdot : K \times K \to K \qquad \text{(Multiplikation)}$$

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

(1) Assoziativitaet der Addition: $x + (y + z) = (x + y) + z$ fuer alle $x, y, z \in K$,
(2) Kommutativitaet der Addition: $x + y = y + x$ fuer alle $x, y \in K$,
(3) Neutrales Element der Addition: $0 + x = x$ fuer alle $x \in K$,
(4) Additives inverses Element: $\forall\, x \in K$ gibt es $x' \in K$, so dass $x + x' = 0$;
(5) Assoziativitaet der Multiplikation: $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ fuer alle $x, y, z \in K$,
(6) Kommutativitaet der Multiplikation: $x \cdot y = y \cdot x$ fuer alle $x, y \in K$,
(7) Neutrales Element der Multiplikation: $1 \cdot x = x$ fuer 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) Distributivitaet: $x \cdot (y + z) = x \cdot y + x \cdot z$ fuer alle $x, y, z \in K$.

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

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 Koerper. Ebenso ist die Menge

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

ein Koerper. Diese Arten von Koerpern heissen Zahlkoerper *(engl. number fields); sie sind sehr wichtig in der algebraischen Zahlentheorie.*

Eine weitere sehr interessante Klasse von Koerpern sind die Koerper mit endlich vielen Elementen.

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

Bemerkung 2.1.4. *Die einzige Schwerigkeit zu zeigen, dass $\mathbb{F}_p$ die Koerperaxiome erfuellt, 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 Element von M ist durch p teilbar.
Beweis von Behaptung 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 enthaelt, folgt aus Behauptung 2, dass es ein $1 \leq y < p$ gibt, so dass xy den Rest 1 bei der Division durch p hat. □

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

Proof. Uebung. □

Beispiele 2.1.7.

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

+	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

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

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

$\cdot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Bemerkung 2.1.8.

(1) Der Koerpers $\mathbb{F}_p$ kann mit Hilfe von modularem Rechnen / Kongruenzen *wesentlich eleganter definiert werden. Aber das gehoert in den Bereich der Zahlentheorie.*

(2) Es sei $n \geq 2$. Sie koennen dann natuerlich auch die Menge der Reste betrachten, die bei der Division durch n entstehen. Addition und Mutiplikation koennen 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) Uebung fuer Ehrgeizige: es sei M die Menge der Reste, die bei der Division durch 10 *entstehen. Finden Sie heraus, welche Elemente von M ein mutliplikatives 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) erfuellen, 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 dann ein Koerper ist, wenn n eine Primzahl ist.*

2.2. **Matrizen.** Es sei K ein Körper.

Definition 2.2.1. *Seinem $m, n \geq 1$. Eine $(m \times n)$ Matrix A mit Werten in K ist eine Tabelle mit m Zeilen und n Spalten, deren Eintraege 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 $\mathbf{1}_n$ 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 $\mathbf{1}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ und $\mathbf{1}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Definition 2.2.5.

(1) *Seinen $A = (a_{ij})$ 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 Eintraegen (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) *$A + 0_{m+n} = 0_{m+n} + A = A$;*

(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. $\square$

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 $(m \times p)$-Matrix*[6] *$C = (c_{ik})$ mit*

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

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*

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

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

[6] Wir werden in Kapitel 4 sehen, warum dies eine gute Definition ist: wenn man Matrizen als lineare Abbildungen interpraetiert, 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

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

(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 seinen 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 dass*[7]

$$AB = BA = \mathbf{1}_n. \tag{5}$$

[7]Wir 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* (42) *erfuellt. 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 (42) erfuellen. 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 naeher 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]

[8] Diese 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

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

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 erfuellt ist:*

1. *in jeder Zeile ist das erste von null verschiedene Element (falls es eins gibt) eine 1 (wir nenen 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 erfuellt.
- 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) erfuellt.
- 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 erfuellt.
- 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) erfuellt. (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 erfuellt 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 $A = \begin{pmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{pmatrix} \in M_{3\times 4}(\mathbb{R})$.

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

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*

$$2x - y = 3$$
$$-x + 3y = -1$$

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 seinen $m, n \geq 1$ 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:

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

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.*

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

- *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 Matix $A|0$; man nennt ein lineares Gleichungssystem der Form $Ax = 0$* homogen.

Theorem 2.4.5. *Es seien*

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

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:

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

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

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

d.h. x ist eine Loesung von (S'). □

[9]Durch 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 (EZU)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

$$\begin{pmatrix} 0 & -9 & 3 & 4 & | & 9 \\ 1 & 4 & 0 & -1 & | & 5 \\ 2 & 6 & -1 & 5 & | & -5 \end{pmatrix}.$$

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

$$\begin{pmatrix} 1 & 0 & 0 & \frac{17}{3} & | & -\frac{22}{5} \\ 0 & 1 & 0 & -\frac{5}{3} & | & \frac{12}{5} \\ 0 & 0 & 1 & -\frac{11}{3} & | & \frac{51}{5} \end{pmatrix}$$

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

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

Lecture 6

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

$$A = \begin{pmatrix} 1 & -2 & 1 & | & b_1 \\ 2 & 1 & 1 & | & b_2 \\ 0 & 5 & -1 & | & b_3 \end{pmatrix}.$$

Wir wandeln sie in reduzierte Zeilenstufenform um:

$$\begin{aligned}
S(1,2,-2) \quad &\to \quad \begin{pmatrix} 1 & -2 & 1 & | & b_1 \\ 0 & 5 & -1 & | & b_2-2b_1 \\ 0 & 5 & -1 & | & b_3 \end{pmatrix} \\
S(2,3,-1) \quad &\to \quad \begin{pmatrix} 1 & -2 & 1 & | & b_1 \\ 0 & 5 & -1 & | & b_2-2b_1 \\ 0 & 0 & 0 & | & b_3-b_2+2b_1 \end{pmatrix} \\
M(2,\frac{1}{5}) \quad &\to \quad \begin{pmatrix} 1 & -2 & 1 & | & b_1 \\ 0 & 1 & -\frac{1}{5} & | & \frac{1}{5}(b_2-2b_1) \\ 0 & 0 & 0 & | & b_3-b_2+2b_1 \end{pmatrix} \\
S(2,1,2) \quad &\to \quad \begin{pmatrix} 1 & 0 & \frac{3}{5} & | & \frac{1}{5}(b_1+2b_2) \\ 0 & 1 & -\frac{1}{5} & | & \frac{1}{5}(b_2-2b_1) \\ 0 & 0 & 0 & | & b_3-b_2+2b_1 \end{pmatrix}
\end{aligned}$$

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

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

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

3.1. **Definition und Beispiele.**

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

$$+ : V \times V \to V, \quad (v_1, v_2) \mapsto v_1 + v_2, \tag{6}$$

$$\times : K \times V \to V \quad (\lambda, v) \mapsto \lambda v \tag{7}$$

so dass die folgenden Bedingungen erfuellt 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 schreiben*[10] $w = -v$;
(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
$$V = \{f : (0,1) \to \mathbb{R} : f \text{ ist stetig}\}.$$
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) Inverse[11] *von v.*

Proof. Nimm an, es gibt $u, u' \in V$ so dass
$$v + u = v + u' = 0_V.$$
Dann gilt
$$u = u + 0_V = u + (v + u') = (u + v) + u' = 0_V + u' = u'.$$
□

[10] Diese 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.

[11] Falls 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) Uebung.
(4) Wir zeigen folgende Aussage[12]: 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

$$\lambda v = 0_V \quad \Leftrightarrow \quad \lambda^{-1}(\lambda v) = 0_V \qquad \text{wegen (1)} \tag{8}$$

$$\Leftrightarrow \quad (\lambda^{-1}\lambda)v = 0_V \qquad \text{wegen (VR7)} \tag{9}$$

$$\Leftrightarrow \quad 1.v = v = 0_V \qquad \text{wegen (VR8)} \tag{10}$$

□

[12]Vergewissern Sie sich, dass diese Aussage aequivalent zu (4) ist.

3.2. **Unterräume.**

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

(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 Bedinungen erfuellt 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 bezueglich der Salar-Multiplikation abgeschlossen ist, gilt

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

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 erfuellt sind. Dann ist U nicht die leere Menge da $0_V \in U$. Es seien nun $u, v \in U$. Fuer $\lambda = 1$ erhalten wirt $u + v \in U$. 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. $\square$

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*
$$U = \{\lambda_1 v_1 + \cdots + \lambda_n v_n : \lambda_i \in K \,\forall\, 1 \leq i \leq n\}.$$
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 $\lambda_1 = \cdots = \lambda_n = 0$ 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
$$u = \mu_1 v_1 + \cdots + \mu_n v_n \quad \text{und} \quad w = \nu_1 v_1 + \cdots + \nu_n v_n.$$
Daher
$$\begin{aligned} u + \lambda w &= \mu_1 v_1 + \cdots + \mu_n v_n + \lambda(\nu_1 v_1 + \cdots + \nu_n v_n) \\ &= (\mu_1 + \lambda\nu_1) v_1 + \cdots + (\mu_n + \lambda\nu_n) v_n \end{aligned}$$
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 als*[13]
$$\langle S \rangle = \{v \in V : \exists\, n \geq 0,\ v_1 \dots, v_n \in S \text{ und } \lambda_1 \dots, \lambda_n \in K \text{ so dass } v = \lambda_1 v_1 + \cdots + \lambda_n v_n\}.$$
Insbesondere ist der Null-Vektorraum die lineare Hülle 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 seinen 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 $i \geq i$ 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 as 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 Unterraeumen lassen sich auf mehrere Arten neue Unterraeume 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.

[13] gesprochen *Krokodilklammer von S*

[14] Fuer unendliche Summen bracht 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 Unterraeume 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 Unterraeume sind, ist das Element $u + \lambda w$ in beiden Unterraeumen enthalten, und daher auch in $U \cap W$.

(2) kann mit aehnlichen Argumenten bewiesen werden (Uebung). □

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*

$$\lambda_1 v_1 + \cdots + \lambda_n v_n$$

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:

$$v = xe_1 + ye_2 + ze_3.$$

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

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

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

$$\begin{pmatrix} 1 & 0 & 1 & | & b_1 \\ 0 & 1 & 1 & | & b_1 + b_2 \\ 0 & 0 & 0 & | & b_3 + b_1 - 2b_2 \end{pmatrix}.$$

Mit anderen Worten, (11) hat genau dann eine Loesung, wenn $b_3 = 2b_2 + b_1$; ansonsten hat es keine Loesung. 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*

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

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

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

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: Lecture 8

Satz 3.3.8. *Es seien $v_1, \dots, v_n \in V$ linear unabhaengig. Dann kann jeder Vektor $v \in \langle v_1, \dots, v_m \rangle$* 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

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

fuer $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in K$. Dann gilt

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

da $v_1, \ldots, v_n$ linear unabhaengig sind. Daher gilt

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

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 abhaengig: $2v_1 - v_2 - v_3 = 0$. Hingegen sind die Vektoren v_1, v_2 linear unabhaengig: die einzige Loesung des linearen Gleichungssystems

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

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 unabhaengig? Mit anderen Worten, hat die Gleichung

$$x_1 v_1 + x_2 v_2 + x_3 v_3 + x_4 v_4 = 0 \tag{12}$$

eine Loesung 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 Loesung von (12). Daher sind die Vektoren linear abhaengig.

(3) Generell kann man zeigen, dass je drei Vektoren in $\mathbb{R}^2$ (allgemeiner: je $n+1$ Vektoren in $\mathbb{R}^n$) linear abhaengig sind – wir werden spaeter verstehen, warum dies so ist.

(4) Per Konvention ist die leere Menge $\emptyset$ linear unabhaengig, und es gilt $\langle \emptyset \rangle = \{0\}$.

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

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

Proof. Uebung. □

Lemma 3.3.11.

(1) Es sei V ein Vektorraum und $v \in V$, $v \neq \{0\}$. Dann ist v linear unabhaengig.

(2) Der Nullvektor 0_V ist linear abhaengig.

Proof. Klar. □

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

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

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

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

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

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

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 groesstmoegliche 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. $\langle T \rangle = V$.

□

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

Es liegt nahe, die Dimension eines Vektorraums as die Anzahl der Element einer Basis zu definieren. Aber ist das wohldefiniert? Koennte 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*

$$w = \alpha_1 v_1 + \cdots + \alpha_n v_n \tag{13}$$

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

$$\alpha_j v_j = w - \sum_{i \neq j} \alpha_i v_i$$

$$\Rightarrow \quad v_j = \frac{1}{\alpha_j} w - \sum_{i \neq j} \frac{\alpha_i}{\alpha_j} v_i. \tag{14}$$

[15] Es sei $V = \{0\}$. Dann ist $\emptyset$ is 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

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

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

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

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

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

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 den Raumes aller Fibonacci Folgen sind: von Satz 1.1.9 wissen wir, dass

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

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 einen 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:

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

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

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

□

Lecture 9

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

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

(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 $v_1, \ldots, v_n \in V$. Dann sind folgende Aussagen aequivalent:*

(i) $v_1, \ldots, v_n$ *sind linear unabhaengig;*

(ii) $v_1, \ldots, v_n$ *sind ein Erzeugendensystem von V;*

(iii) $v_1, \ldots, v_n$ *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 $\{v_1, \ldots, v_n\}$ zu einer Basis von V erweitern. Aber jede Basis von V hat n Elemente, daher ist $\{v_1, \ldots, v_n\}$ selber eine Basis.

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

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

(i) Wenn $k < n$, dann sind $v_1, \ldots, v_k$ kein Erzeugendensystem fuer V.

(ii) Wenn $k > n$, dann sind $v_1, \ldots, v_k$ linear abhaengig.

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

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.

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

3.4. **Basen von Unterraeumen.**

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 unabhaengig sind als Elemente von U, dann sind sie ebenfalls linear unabhaengig 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*

$$\dim_K U \le \dim_K V,$$

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 Moeglichkeiten: 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 waehle $u_2 \in U - \langle u_1 \rangle$. Es folgt von Satz 3.3.12, dass u_1, u_2 linear unabhaengig sind.

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

Aufgrund von Satz 3.3.24 muss dieser Prozess nach hoechstens n Schritten enden. Wir erhalten also eine Liste $u_1, \dots, u_k$ fuer $k \le n$ von linear unabhaengigen 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

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

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 unabhaengig; sie sind ebenfalls linear unabhaengig 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 Unterraumen U, W von V die Unterraeume $U \cap W$ 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 $U = W \lneq V$ 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 unabhaengig, und daher gilt
$$\dim(U + W) = 3 = 2 + 1 = \dim U + \dim W.$$
- Es seien nun
$$U' = \langle e_2, e_3 \rangle \qquad \text{und} \qquad W' = \langle e_2 + e_3, e_4 \rangle.$$
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 unabhaengig: $e_2 + e_3 \in \langle e_2, e_3, e_4 \rangle$, 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 unabhaengig 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 einen 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 Unterraeume 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 $\dim_K(U \cap W) = k$. Waehle eine Basis $v_1, \ldots, v_k$ von $U \cap W$. Aufgrund von Theorem 3.3.19 koennen wir diese Basis zu Basen

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

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

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

eine Basis ist von $U + W$.

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

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

Wir zeigen die lineare Unabhaengigkeit. Nimm an, dass

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

Dann folgt, dass Lecture 10

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

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

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

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

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

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

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

Daher gilt

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

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

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

□

Warnung 3.4.5. *Es ist* ***nicht wahr****, dass fuer drei Unterraeume U, W, X von V folgende Formel gilt:*[17]

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

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

$$U = \langle e_1 \rangle, \qquad W = \langle e_2 \rangle, \qquad X = \langle e_1 + e_2 \rangle. \tag{15}$$

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

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

d.h. in (15) *erhalten wir $2 = 3$!*[18]

Philipp Steiner und Marco Vaccaro haben eine schoene Formel fuer den Schnitt von beliebeig vielen Unterraeumen gefunden; sie ist allerdings nicht symmetrisch in den A_i's.

Theorem 1. *Seien $A_1, A_2, \ldots, A_n$ Unterraeume eines Vektorraums V . Dann gilt für die Dimension der Summe dieser Vektorräume:*

$$\dim\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \dim(A_i) - \sum_{i=1}^{n-1} \dim\left(\left(\sum_{t=1}^{i} A_t\right) \cap A_{i+1}\right)$$

[17] Selbst viele professionelle Mathematiker wissen das nicht!

[18] Das Problem ist, dass im allgemeinen

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

Proof. Für $n = 1$ ist der Satz trivial, da die Dimension der Summe eines einzelnen Vektorraums einfach seine eigene Dimension ist:

$$\dim(A_1) = \dim(A_1)$$

Für $n = 2$ gilt die bekannte Dimensionsformel für die Summe zweier Vektorräume:

$$\dim(A_1 + A_2) = \dim(A_1) + \dim(A_2) - \dim(A_1 \cap A_2)$$

Dies ist der Induktionsanfang für $n = 1$ und $n = 2$. **Induktionsannahme** Nehmen wir an, dass der Satz für $n = k$ gilt. Das heißt, wir nehmen an:

$$\dim\left(\sum_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} \dim(A_i) - \sum_{i=1}^{k-1} \dim\left(\sum_{t=1}^{i} A_t \cap A_{i+1}\right)$$

Induktionsschritt Nun wollen wir zeigen, dass der Satz auch für $n = k + 1$ gilt. Es gilt:

$$\dim\left(\sum_{i=1}^{k+1} A_i\right) = \dim\left(\sum_{i=1}^{k} A_i + A_{k+1}\right)$$

Verwenden wir die Dimensionsformel für die Summe zweier Vektorräume:

$$\dim(A + B) = \dim(A) + \dim(B) - \dim(A \cap B)$$

Dann ergibt sich:

$$\dim\left(\sum_{i=1}^{k+1} A_i\right) = \dim\left(\sum_{i=1}^{k} A_i\right) + \dim(A_{k+1}) - \dim\left(\left(\sum_{i=1}^{k} A_i\right) \cap A_{k+1}\right) \tag{16}$$

Nun setzen wir die Induktionsannahme für $\dim\left(\sum_{i=1}^{k} A_i\right)$ ein:

$$(16) = \left(\sum_{i=1}^{k} \dim(A_i) - \sum_{i=1}^{k-1} \dim\left(\left(\sum_{t=1}^{i} A_t\right) \cap A_{i+1}\right)\right) + \dim(A_{k+1}) - \dim\left(\left(\sum_{i=1}^{k} A_i\right) \cap A_{k+1}\right)$$

Dies lässt sich umformen zu:

$$(16) = \sum_{i=1}^{k+1} \dim(A_i) - \sum_{i=1}^{k} \dim\left(\left(\sum_{t=1}^{i} A_t\right) \cap A_{i+1}\right)$$

Somit ist die Aussage für $n = k + 1$ ebenfalls bewiesen. □

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

(i) $\dim(U + W) = \dim U + \dim W$*;*
(ii) $\dim(U \cap W) = 0$*;*
(iii) $U \cap W = \{0_V\}$*.*
(iv) fuer 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 Loesung.

Proof. (i) ⇔ (ii): folgt direkt von Theorem 3.4.4.
(ii) ⇔ (iii): klar, da es genau einen Vektorraum der Dimension null gibt.
(iii) ⇒ (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) ⇒ (v): die Gleichung $u + w = 0_V$ hat die Loesung $u = w = 0_V$, und da U, W Unterraeume sind, elthalten beide das Element 0_V. Indemn wir nun (iv) auf $v = 0_V$ andwenden, sehen wir, dass dieses die einzige Loesung ist.
(v) ⇒ (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. □

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

$$\begin{aligned} & v = \sum_{i=1}^{\ell} \alpha_i u_1 = \sum_{j=1}^{m} \lambda_j w_j \\ \Rightarrow \quad & \alpha_1 v_1 + \cdots + \alpha_\ell v_\ell - (\lambda_1 w_1 + \cdots + \lambda_m w_m) = 0_V \\ \Rightarrow \quad & \alpha_1 = \cdots = \alpha_\ell = \lambda_1 = \lambda_m = 0, \end{aligned}$$

da die Vektoren $u_1, \ldots, e_\ell, w_1, \ldots, w_m$ liear unabhaengig 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 unabhaengig 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 it in Beispiel 3.4.9 z.B. $\langle e_2 \rangle$ ein Komplement von U. Das Gleiche gilt fuer $\left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle$ oder fuer $\left\langle \begin{pmatrix} 3 \\ -1 \end{pmatrix} \right\rangle$.*

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

$$w = \lambda_1 v_1 + \dots \lambda_n v_n. \tag{17}$$

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 (17) 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 Spaltenraeume.**

Bemerkung 3.6.1. *Wir koennen 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*

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

Was koennen wir ueber die Dimensionen dieser Unterraeume sagen?

Definition 3.6.3. *Wir definieren*

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

Beispiele 3.6.4.

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

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

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

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

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

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

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

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

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

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

Was faellt ihnen auf?

Hier ist das erste wirklich ueberraschende 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*

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

Fuer Matrizen in reduzierter Zeilenspaltenform ist dieses Resultat offensichtlich.

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

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

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

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

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

[19] Ich 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 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 Spalten(A), d.h.

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

$\square$

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*

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

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.

Lecture 11

4.1. **Definition und Beispiele.**

Definition 4.1.1. *Es seien V, W Vektorraeume ueber K. Eine Funktion $T : V \to W$ ist eine* lineare Abbildung, *wenn sie folgende Bedingungen erfuellt:*

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

Mit anderen Worten, T respektiert die Vektoraddition und Skalarmultiplikation.

Eine lineare Abbildung $T : V \to V$ 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 Identitaetsabbildung $\mathrm{id}_V : V \to V$, $v \mapsto v$ ist linear.
(2) Es seien V, W Vektorraeume. Die Null-Abbildung $V \to W$, $v \mapsto 0_W$ ist linear.
(3) Es sei $K[x]^{\leq n}$ der Vektorraum aller Polynome vom Grad $\leq n$. The Ableitungsabbildung
$$D : K[x]^{\leq n} \to K[x]^{\leq n}, \quad a_0 + a_1x + \cdots + a_nx^n \mapsto a_1 + 2a_2x + \cdots + na_nx^{n-1}$$
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
$$\mathrm{Tr} : V \to K, \qquad A = (a_{ij})_{1\leq i,j\leq n} \mapsto a_{11} + \cdots + a_{nn}.$$
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*
$$T_A : K^n \to K^m, \qquad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Lemma 4.1.5. *Die Abbildung T_A ist linear.*

Proof. Uebung. □

Bemerkung 4.1.6. *Eine Abbildung $T : V \to W$ ist genau dann linear, wenn*
$$T(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 T(v_1) + \alpha_2 T(v_2)$$
fuer alle $v_1, v_2 \in V$ und $\alpha_1, \alpha_2 \in K$.

Satz 4.1.7. *Es seien V, W Vektorraeume ueber K und $T : V \to W$ eine lineare Abbildung.*

(i) Es seien $v_1, \dots, v_n \in V$ und $\alpha_1, \dots, \alpha_n \in K$. Dann gilt
$$T(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n).$$
(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 Vektorraeume ueber K, V endlich-dimensional, und es sei $v_1, \dots, v_n$ eine Basis von V. Es sei $T : V \to W$ eine lineare Abbildung. Dann ist T durch $T(v_1), \dots, T(v_n)$ eindeutig bestimmt.*[20]

[20] Mit anderen Worten, wenn wir $T(v_1), \dots, T(v_n)$ kennen, dann kennen wir auch $T(v)$ fuer 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. □

Tatsaechlich 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 $T : V \to W$ so dass*

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

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

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

  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 liienar unabhaengig ist. Betrachte z.B. das Erzeugendensystem $e_1, e_2, e_1 + e_2 \in \mathbb{R}^2$. Gibt es eine lineare Abbildung $T : \mathbb{R}^2 \to \mathbb{R}$, 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 $T : V \to U$ und $S : U \to W$ lineare Abbildungen. Dann ist die Abbildung*

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

ebenfalls linear.

Proof. Uebung. □

4.2. **Kernel and Image.**

Definition 4.2.1. *Es seien V, W Vektorraeume ueber K, und es sei $T : V \to W$ 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 is
$$\operatorname{im}(T) = \{T(v) : v \in V\} \subseteq W.$$

Beispiele 4.2.2.

(1) Es sei $A = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$, und es sei $T_A : \mathbb{R}^3 \to \mathbb{R}^3$ die dazugehoerige 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 Loesung eines linearen Gleichungssystems. Wir haben die Loesung dieses Gleichungssystems bereits in Beispiel 3.3.5 bestimmt: die Loesung ist
$$\ker(T_A) = \left\langle \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right\rangle.$$
Was ist $\operatorname{im}(T_A)$? Laut Definition ist
$$\operatorname{im}(T_A) = \left\{ b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} : \exists\, x \in \mathbb{R}^3 \text{ so dass } A.x = b \right\}.$$
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 $T_B : \mathbb{R}^2 \to \mathbb{R}^3$ die dazugehoerige lineare Abbildung. Eine einfache Rechung 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 $T(v_i) = w_i$ fuer $i = 1, 2$. 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 betrachen 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 $T : V \to W$ eine lineare Abbildung. Wenn V endlich-dimensional ist, dann sind auch* $\ker(T)$ *und* $\mathrm{im}(T)$ *endlich-dimensional.*[21]

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

Um die Aussage fuer $\mathrm{im}(T)$ 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 $\mathrm{im}(T)$ ist. Daher ist $\mathrm{im}(T)$ endlich-dimensional. □

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

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

Proof. $\Rightarrow$: klar von der Definition. Lecture 12

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

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

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

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

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

Der folgende Satz bringt $\mathrm{rk}(T)$ 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 $T : V \to W$ eine lineare Abbildung. Dann gilt*

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

Proof. Es sei $u_1, \dots, u_n$ einen Basis fuer $s\ker(T)$. 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 $\mathrm{im}(T)$.

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

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

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

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

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

Wir erhalten die Gleichung

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

Da die Vektoren linear unabhaengig sind, folgt, dass

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

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

Korollar 4.2.10. *Es sei $T : V \to W$ 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 bijectiv $\Leftrightarrow$ T ist injectiv $\Leftrightarrow$ T ist surjectiv.

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

Proof. (1) Da $\mathrm{im}(T) \leq W$, gilt $\mathrm{rk}(T) \leq \dim W$. Daraus folgt, dass
$$\dim \ker(T) = \dim V - \mathrm{rk}(T) \geq \dim V - \dim W > 0$$
und T ist nicht injektiv (Satz 4.2.6).

(2) Aus Theorem 4.2.9 folgt, dass
$$\mathrm{rk}\,\mathrm{im}(T) = \dim V - \dim \ker(T) \leq \dim V < \dim W,$$
so dass T nicht surjectiv sein kann.

(3)
$$\begin{aligned} T \text{ ist injektiv} \quad &\Leftrightarrow \quad \ker(T) = \{0_V\} \\ &\Leftrightarrow \quad \mathrm{rk}(T) = \dim V \\ &\Leftrightarrow \quad \mathrm{rk}(T) = \dim W \\ &\Leftrightarrow \quad \mathrm{im}(T) = W \\ &\Leftrightarrow \quad T \text{ ist surjektiv} \end{aligned}$$

□

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-aequivalent 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 Loesung des homogenen linearen Gleichungssystems $B.x = 0_{\mathbb{R}^3}$, 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 $T : V \to W$ 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 $T : V \to W$ ist ein* Isomorphismus*, wenn es eine lineare Abbildung $S : W \to V$ gibt, so dass*
$$S \circ T = \mathrm{id}_V \qquad \text{und} \qquad T \circ S = \mathrm{id}_W;$$
in diesem Fall schreiben wir $S = T^{-1}$.

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

Bemerkung 4.2.14. *Es sei X die Menge aller Vektorraeume ueber K. Dann ist "$\cong$" eine Aequivalenzrelation 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 $V \to \mathbb{R}^3$ konstruieren: wir wissen, dass $1, x, x^2$ eine Basis von V ist. Dann gibt es genau eine lineare Abbildung $T : V \to \mathbb{R}^3$ fuer die gilt
$$T(1) = e_1, \qquad T(x) = e_2, \qquad T(x^2) = e_3.$$

Um zu zeigen, dass T ein Isomorphismus ist, konstruieren wir eine inverse lineare Abbildung: Definiere $S : \mathbb{R}^3 \to V$ durch
$$S : \begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto a + bx + cx^2.$$
Dann ist klar, dass $S \circ T = \mathrm{id}_V$ und $T \circ S = \mathrm{id}_{\mathbb{R}^3}$.

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 $T : V \to V$ 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 $T : V \to W$ einen Isomorphismus $T : X \cong \operatorname{im}(T)$.*

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

Lemma 4.2.19. *Es sei $T : V \to W$ eine bijektive lineare Abbildung. Dann ist die inverse Abbildung $T^{-1} : W \to V$ 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 $v_1, \ldots, v_n$ und $w_1, \ldots, w_n$ von V und W. Aus Theorem 4.1.9 folgt, dass es eine lineare Abbildung $T : V \to W$ gibt, so dass $T(v_i) = w_i \ \forall\, 1 \leq i \leq n$. Dann ist T surjectiv: wenn $\alpha_1 w_1 + \cdots + \alpha_n w_n \in W$, dann gilt

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

Es folgt von Korollar 4.2.10, dass T bijektiv ist und daher dank Lemma 4.2.19 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 $T : V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen K-Vektorraeumen, 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 bezueglich der Basen $\mathcal{B}$ und $\mathcal{C}$ ist die Matrix $[T]_{\mathcal{C}}^{\mathcal{B}} = (a_{ij}) \in M_{m\times n}(K)$, deren Eintraege definiert sind durch*

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

Mit anderen Worten, die Eintraege 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 bezueglich jeder Basis ist die Nullmatrix.
(2) Die Abbildungsmatrix der Identitaetsabbildung bezueglich der Basis $\mathcal{B}$ ist die Einheitsmatrix: $[\mathrm{id}]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{1}_n$.

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 $D : V \to W$. 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 $S : V \to V$ 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 schoene Form: sie ist diagonal.

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

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

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

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

Beispiel 4.3.6. Zurueck 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 dieses 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-Vektorraeume V und W mit jeweiligen Basen $\mathcal{B} = (v_1, v_2)$ und $\mathcal{C} = (w_1, w_2, w_3)$. Es sei $T : V \to W$ 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 $T : V \to W$ und $S : W \to U$ 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

$$T(v_j) = a_{1j}w_1 + \dots + a_{mj}w_m, \tag{18}$$

$$S(w_i) = b_{1i}u_1 + \dots + b_{pi}u_p, \tag{19}$$

$$(S \circ T)(v_j) = c_{1j}u_1 + \dots + c_{pj}u_p \tag{20}$$

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

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

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 Vektorraeume mit jeweiligen Basen $\mathcal{B} = (v_1, \ldots, v_n)$ und $\mathcal{C} = (w_1, \ldots, w_m)$. Wir haben bereits gesehen, dass jede lineare Abbildung $T : V \to W$ eine Abbildungsmatrix $[T]_{\mathcal{C}}^{\mathcal{B}}$ hat.*

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

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

***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 Vektorraeume mit jeweiligen Basen $\mathcal{B} = (v_1, \ldots, v_n)$ und $\mathcal{C} = (w_1, \ldots, w_m)$.*

(1) Es sei $T : V \to W$ eine lineare Abbildung. Dann gilt $L_{[T]_{\mathcal{C}}^{\mathcal{B}}} = T$.
(2) Es sei $A \in M_{n\times n}(K)$. 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] $\square$

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

$$\{m \times n\text{-Matrizen mit Werten in } K\} \leftrightarrow \{\text{lineare Abbildungen } V \to W\} \tag{21}$$

$$A \mapsto L_A \tag{22}$$

$$[T]_{\mathcal{C}}^{\mathcal{B}} \leftarrow\!\shortmid\; T \tag{23}$$

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 $\mathcal{B} = (v_1, \ldots, v_n)$. Es sei $T : V \to V$ 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 $\left([T]_{\mathcal{B}}^{\mathcal{B}}\right)^{-1}$.*

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

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

Dann gilt aufgrund von Satz 4.3.7, dass

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

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 Lecture 14

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

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

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

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

[22] Um (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 $T : V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen K-Vektorraeumen, 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 verschidene Basen von V, bzw. von W. Was ist die Beziehung zwischen den Matrizen $[T]^{\mathcal{B}}_{\mathcal{C}}$ und $[T]^{\mathcal{B}'}_{\mathcal{C}'}$?

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 $A = [\mathrm{id}]^{\mathcal{B}}_{\mathcal{B}'}$, d.h. $A = (a_{ij})$ ist die Matrix, deren Eintraege definiert sind durch die Gleichungen*

$$v_j = \sum_{i=1}^{n} a_{ij} v'_i \qquad \textit{fuer } 1 \le j \le n.$$

Die Matrix A heisst Basiswechselmatrix *von $\mathcal{B}$ nach $\mathcal{B}'$.*

Bemerkung 4.5.2. *Wir koennen mit Hilfe der Basiswechselmatrix $A = [\mathrm{id}]^{\mathcal{B}}_{\mathcal{B}'}$ eine Linearkombination bezueglich der Basis $\mathcal{B}$ als eine Linearkombination bezueglich der Basis $\mathcal{B}'$ ausdruecken: es sei $v = \alpha_1 v_1 + \cdots + \alpha_n v_n \in V$. Dann gilt*

$$v = \beta_1 v'_1 + \cdots + \beta_n v'_n,$$

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 $[\mathrm{id}]^{\mathcal{B}}_{\mathcal{B}'}$ ist invertierbar, mit Inversem*

$$\left([\mathrm{id}]^{\mathcal{B}}_{\mathcal{B}'}\right)^{-1} = [\mathrm{id}]^{\mathcal{B}'}_{\mathcal{B}}.$$

Proof. Von Theorem 4.3.7 folgt, dass

$$[\mathrm{id}]^{\mathcal{B}}_{\mathcal{B}} = \mathbf{1}_n = [\mathrm{id}]^{\mathcal{B}}_{\mathcal{B}'} \circ [\mathrm{id}]^{\mathcal{B}'}_{\mathcal{B}}$$
$$[\mathrm{id}]^{\mathcal{B}'}_{\mathcal{B}'} = \mathbf{1}_n = [\mathrm{id}]^{\mathcal{B}'}_{\mathcal{B}} \circ [\mathrm{id}]^{\mathcal{B}}_{\mathcal{B}'}.$$

□

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

$$[\mathrm{id}]^{\mathcal{B}}_{\mathcal{E}} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

Wir koennen natuerlich auch die Basiswechselmatrix von $\mathcal{B}'$ nach $\mathcal{B}$ betrachten: es gilt

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

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 $T : V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen K-Vektorraeumen, 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 verschidene Basen von V, bzw. von W. Dann gilt*

$$[T]^{\mathcal{B}'}_{\mathcal{C}'} = [\mathrm{id}_W]^{\mathcal{C}}_{\mathcal{C}'} \cdot [T]^{\mathcal{B}}_{\mathcal{C}} \cdot [\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{B}}.$$

Proof. Das Theorem folgt durch mehrfache Anwendung von Satz 4.3.7.

Alternativ koennen Sie ovn den Definitionen argumentieren: let $A = [T]^{\mathcal{B}}_{\mathcal{C}}$, $P = [\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{B}}$ and $Q = [\mathrm{id}_W]^{\mathcal{C}}_{\mathcal{C}'}$. Diese Matrizen sind durch folgende Eigenschaften definiert (vergl. Definition 4.3.1 und Beispiel 4.3.2 (3)):

$$v'_\ell = p_{1\ell} v_1 + \cdots + p_{n\ell} v_n,$$
$$T(v_k) = a_{1k} w_1 + \cdots + a_{mk} w_m,$$
$$w_j = q_{1j} w'_1 + \cdots + q_{mj} w'_m$$

fuer alle $1 \leq k, \ell \leq n$ und $1 \leq j \leq m$. Dann gilt

$$
\begin{aligned}
T(v'_\ell) &= \sum_{k=1}^{n} p_{k\ell} T(v_k) \\
&= \sum_{k=1}^{n}\sum_{j=1}^{m} p_{k\ell} a_{jk} w_j \\
&= \sum_{k=1}^{n}\sum_{j=1}^{m}\sum_{i=1}^{m} p_{k\ell} a_{jk} q_{ij} w'_i \qquad (24)
\end{aligned}
$$

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 (24) und (25) vergleichen, erhalten wir

$$b_{i\ell} = \sum_{k=1}^{n}\sum_{j=1}^{m} p_{k\ell} a_{jk} q_{ij}. \qquad (25)$$

□

Bemerkung 4.5.6. *Wir koennen Theorem 4.5.5 folgendermassen zusammenfassen: $T : V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen K-Vektorraeumen, und es seien $\mathcal{B}$ und $\mathcal{B}'$, bzw. $\mathcal{C}$ und $\mathcal{C}'$ zwei verschidene 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 $T : V \to V$ ein Endomorphismus, und es seien $\mathcal{B} = (v_1, \ldots, v_n)$ und $\mathcal{B}' = (v'_1, \ldots, v'_n)$ Basen von V. Es sei $P = [\mathrm{id}]_{\mathcal{B}'}^{\mathcal{B}}$. Dann gilt*

$$[T]_{\mathcal{B}'}^{\mathcal{B}'} = [\mathrm{id}]_{\mathcal{B}'}^{\mathcal{B}} \cdot [T]_{\mathcal{B}}^{\mathcal{B}} \cdot [\mathrm{id}]_{\mathcal{B}}^{\mathcal{B}'}.$$

Beispiel 4.5.8. Es sei $A = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$, und wir betrachten die Abbildung $T_A : \mathbb{R}^2 \to \mathbb{R}^2$. 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

$$[\mathrm{id}]_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \qquad \text{und} \qquad [\mathrm{id}]_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}.$$

Von Korollar 4.5.7 erhalten wir

$$[T_A]_{\mathcal{B}}^{\mathcal{B}} = [\mathrm{id}]_{\mathcal{B}}^{\mathcal{E}} \cdot [T_A]_{\mathcal{E}}^{\mathcal{E}} \cdot [\mathrm{id}]_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -3 & -10 \\ 2 & 5 \end{pmatrix}.$$

Diese Matrix bedeutet, dass

$$T_A : v_1 \mapsto -3v_1 + 2v_2 \qquad \text{und} \qquad T_A : v_2 \mapsto -10v_1 + 4v_2.$$

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 $S : V \to V$ 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

$$[\mathrm{id}]_{\mathcal{B}}^{\mathcal{C}} = \begin{pmatrix} 1 & 1 \\ \varphi & \varphi \end{pmatrix}$$

mit inverser Matrix

$$[\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}} = \frac{1}{\sqrt{5}} \begin{pmatrix} -\psi & 1 \\ \varphi & -1 \end{pmatrix}.$$

Daher gilt

$$[S]^{\mathcal{C}}_{\mathcal{C}} = [\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}} \cdot [S]^{\mathcal{B}}_{\mathcal{B}} \cdot [\mathrm{id}]^{\mathcal{C}}_{\mathcal{B}} = \frac{1}{\sqrt{5}} \begin{pmatrix} -\psi & 1 \\ \varphi & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \varphi & \varphi \end{pmatrix} = \begin{pmatrix} \varphi & 0 \\ 0 & \psi \end{pmatrix},$$

wie erwartet. (Rechnen Sie nach!)

Definition 4.5.10.

(1) Zwei Matrizen $A, B \in M_{n\times n}(K)$ sind aehnlich*, 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 aehnlich.

(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 aequivalent, 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) 'Aehnlichkeit' definiert eine Aequivalenzrelation auf $M_{n\times n}(K)$.

(2) 'Äquivalenz' definiert eine Aequivalenzrelation auf $M_{m\times n}(K)$.

Proof. Wir beweisen (1); der Beweis von (2) isteine Uebung. Schreibe $A \sim B$ wenn A aehnlich 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
$$B = P^{-1} \cdot A \cdot P \qquad \Rightarrow \qquad A = (P^{-1})^{-1} B\, P^{-1},$$
und daher $B \sim A$.
- Wenn $A \sim B$ und $B \sim C$, dann gibt es invertierbare Matrizen P, Q so dass
$$B = P^{-1} A\, P \qquad \text{und} \qquad C = Q^{-1} B\, Q,$$
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 Laenge n mit Koeffizienten in einem Koerper K ist ein Vektorraum, naemlich K^n. Allerdings gibt es noch viele andere Vektorraeume.

Ein *Vektor* ist zunaechst nur ein Element u eines Vektorraums V. **Nachdem** wir eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ gewaehlt haben, koennen wir u einen Spaltenvektor zuordnen: schreibe u als Linearkombination bezueglich der Basis,

$$u = \alpha_1 v_1 + \cdots + \alpha_n v_n \qquad \alpha_i \in K \,\forall\, 1 \leq i \leq n.$$

Dann sind die Eintraege des Vektors $\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \in K^n$ die *Koordinaten* von u bezueglich der gegebenen Basis. Wenn wir die Basis von V aendern, dann aendern sich natuerlich auch die Koordinaten von u: die neuen Koordinaten sind gegeben durch die Formel in 4.5.2.

Formal laesst sich das folgendermassen formulieren:

Satz 4.6.1. *Es sei $\mathcal{B}$ eine Basis von V. Definiere*

$$\Phi_{\mathcal{B}} : V \to K^n$$

als die Abbildung, die einen Vektor auf seinen Koordinaten bezueglich der Basis $\mathcal{B}$ abbildet. Dann ist $\Phi_{\mathcal{B}}$ ein nicht-kanonischer Isomorphismus.

Proof. Da $\mathcal{B}$ eine Basis ist, laesst 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

$$\Psi_{\mathcal{B}} : K^n \to V,$$
$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \mapsto \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

Daher ist $\Phi_{\mathcal{B}}$ surjektiv und daher ein Isomorphismus aufgrund von Lemma 4.2.19. □

Wenn wir nun Vektorraeume 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 koennen wir die lineare Abbildung $L_A : K^n \to K^m$ aus Bemerkung 4.4.1 folgendermassen mit der Abbildung $T_A : K^n \to K^m$ 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

$$\text{Zeilenrang}(A) = \text{Spaltenrang}(A). \tag{26}$$

Der Beweis erfordert etwas Vorbereitung. Wir beginninen mit folgenden Bemerkungen:

Bemerkung 4.7.1. *Von Definiton 4.1.4 wissen wir, dass wir von A eine lineare Abbildung $T_A : K^n \to K^m$ erhalten. Es seinen $\mathcal{E}$ und $\mathcal{F}$ die jeweiligen Standardbasen von K^n und K^m. Dann folgt von Lemma 4.4.3* [23]*, dass*

$$[T_A]^{\mathcal{E}}_{\mathcal{F}} = A.]$$

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 $[T_A]^{\mathcal{B}}_{\mathcal{C}} = (c_{ij})$, dann gilt per Definition

$$T_A(v_j) = c_{1j}w_1 + \cdots + c_{mj}w_m.$$

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,

$$[T_A]^{\mathcal{B}}_{\mathcal{C}} = \begin{pmatrix} \mathbf{1}_r & 0_{r\times s} \\ 0_{t\times r} & 0_{t\times s}\end{pmatrix}.$$

Nun wissen wir von Theorem 4.5.5, dass

$$[T_A]^{\mathcal{B}}_{\mathcal{C}} = [\text{id}_{K^m}]^{\mathcal{F}}_{\mathcal{C}} \cdot [T_A]^{\mathcal{E}}_{\mathcal{F}} \cdot [\text{id}_{K^n}]^{\mathcal{B}}_{\mathcal{E}}.$$

Da $[T_A]^{\mathcal{E}}_{\mathcal{F}} = A$, erhalten wir das Resultat fuer $Q = [\text{id}_{K^m}]^{\mathcal{F}}_{\mathcal{C}}$ und $P = [\text{id}_{K^n}]^{\mathcal{B}}_{\mathcal{E}}$. □

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.*

---

[23] weil T_A nichts anderes ist als die Abbildung $L_A : K^n \to K^m$ bezueglich der Basen $\mathcal{E}$ und $\mathcal{F}$.

Proof. $\Leftarrow$: Es sei r der gemeinsame Spaltenrang. By Theorem 4.7.3 gilt

$$a \sim \begin{pmatrix} \mathbf{1}_r & 0_{r\times s} \\ 0_{t\times r} & 0_{t\times s} \end{pmatrix} \qquad \text{und} \qquad B \sim \begin{pmatrix} \mathbf{1}_r & 0_{r\times s} \\ 0_{t\times r} & 0_{t\times s} \end{pmatrix}.$$

Da '$\sim$' eine Aequialenzrelation 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

$$\begin{aligned} \mathrm{im}(T_B) &= T_B(K^n) \\ &= T_P \circ T_A\left(T_Q(K^n)\right) \\ &= T_P \circ T_A(K^n) && \text{da } T_Q \text{ ein Isomorphismus ist} \\ &= T_P(\mathrm{im}(T_A)) \end{aligned}$$

Daraus folgt, dass

$$\dim \mathrm{im}(T_B) = \dim T_P\left(\mathrm{im}(T_A)\right),$$

Doch da T_P ein Isomorphismus ist, folgt von Korollar 4.2.12, dass

$$\dim T_P\left(\mathrm{im}(T_A)\right) = \dim \mathrm{im}(T_A).$$

□

Übung 4.7.5. *Koennen Sie die Aequivalenzklassen von aehnlichen Matrizen beschreiben?*

Um (26) 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) Zeilenrang(QA) = Zeilenrang(A)*;*
(2) Spaltenrang(QA) = 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

$$q_{i1}x_1 + \cdots + q_{im}x_m,$$

mit anderen Worten, die Zeilen von QA sind Linearkombinationen von den Zeilen von A und daher

$$\text{Zeilenrang}(QA) \le \text{Zeilenrang}(A).$$

Da Q invertierbar ist, koennen wir dieses Argument auch Multiplikation bei $Q^{-1}A$ anwenden und erhalten

$$\text{Zeilenrang}(Q^{-1}QA) \le \text{Zeilenrang}(QA),$$

das heisst Zeilenrang$(A) \le$ Zeilenrang(QA). Daraus folgt, dass

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

(2) Es seien nun $y_1, \dots, y_n$ die Spalten von A, und es sei $U = \langle y_1, \dots, y_n\rangle \le K^m$. 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 $L_Q : K^m \to K^m$ 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

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

□

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) Fuer einen 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

$$\text{Zeilenrang}(B) = \text{Spaltenrang}(B^t) \quad \textit{und} \quad \text{Zeilenrang}(B^t) = \text{Spaltenrang}(B)$$

Proof. Uebung. □

Wir koennen 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) Zeilenrang(AP) = Zeilenrang(A)*;*

(2) Spaltenrang(AP) = 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*

$$\text{Zeilenrang}(QAP) = \text{Zeilenrang}(A) \qquad \textit{und} \qquad \text{Spaltenrang}(QAP) = \text{Spaltenrang}(A).$$

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 Korollary 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 rk(A) *von A.*

Bemerkung 4.7.15. *Obwohl fuer $A \in M_{m\times n}(K)$ gilt*

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

sind die Raeume Spalten(A) *und* Zeilen(A) *verschieden! Zunaechst ist* Spalten$(A) \leq K^m$ *und* Zeilen$(A) \leq K^n$*. Aber auch wenn $m = n$ erhalten wir im allgemeinen verschiedene Unterraeume von K^n.*

4.8. **Zurueck zu linearen Gleichungssystemen.** Es sei $A \in M_{m\times n}$, und wir betrachten das homogene lineare Gleichungssystem

$$(S_A):\ A.x = 0_{K^m}.$$

Wir schreiben $L(S_A) \leq K^n$ fuer die Loesungen des Gleichungssystems.

Lemma 4.8.1. *Es gilt*

$$\dim L(S_A) = n - \operatorname{Rang}(A).$$

Proof. Es sei T_A die durch A gegebene lineare Abbildung $K^n \to K^m$ bezueglich der Standardbasen. Beachte (Bemerkung 4.7.2), dass $L(S_A) = \ker(T_A)$ und $\operatorname{Rang}(A) = \dim \operatorname{im}(T_A)$. Aber aufgrund von Theorem 4.2.9 gilt

$$n = \dim\ker(T_A) + \dim\operatorname{im}(T_A).$$

□

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 $r = \operatorname{Rang}(A')$ die Anzahl der fuehrenden Einsen von A'; nimm an, dass sie in den Spalten

$$1 \leq j_1 < j_2 < \cdots < j_r \leq n.$$

auftreten. Die Loesungsmenge $L(S_{A'})$ hat dann $\ell = n - r$ freie Variablen, sagen wir $x_{i_1}, \ldots, x_{i_\ell}$ fuer

$$1 \leq i_1 < \cdots < i_\ell \leq n.$$

Bemerkung 4.8.2. *Fuer alle Werte der freien Variablen $x_{i_1}, \ldots, x_{i_\ell} \in K$ sind die Werte der Variablen $x_{j_1}, \ldots, x_{j_r}$ eindeutig bestimmt, und zwar durch*

$$x_{j_k} = -\sum_{1\leq q\leq \ell:\, i_q > j_k} a'_{k,q} x_{i_q}. \tag{27}$$

Wir erhalten daher eine Abbildung

$$\Phi:\ K^\ell \to L(S_A),$$

die ein ℓ-Tupel $(\lambda_1, \ldots, \lambda_\ell)$ 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

$$x_{i_1} = \lambda_1, \ldots, x_{i_\ell} = \lambda_\ell.$$

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

$$\begin{aligned} x_2 &= x_3 - 2x_6 \\ x_4 &= x_6 \\ x_5 &= -x_6 \end{aligned}$$

Mit anderen Worten, die Abbildung $\Phi : \mathbb{R}^3 \to L(S_{A'})$ 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 (27) zeigt, dass $x_{j_1}, \dots, x_{j_r}$ linear von den freien Variablen abhaengen; 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. □

Beispiel 4.8.5. Bezogen auf Beispiel 4.8.3 heisst das, dass jedes Element $x \in L(S_{A'})$ von der Form

$$a \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + c \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

fuer igendwelche $a, b, c \in \mathbb{R}$ 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

$$\Phi^{-1}(x) = \begin{pmatrix} x_{i_1} \\ \vdots \\ x_{i_\ell} \end{pmatrix}.$$

Lecture 17

Korollar 4.8.7. *Es seien $A \in M_{m \times n}(K)$, $b \in K^m$, und schreibe $(S_{A,b})$ fuer das lineare Gleichungssystem $A.x = b$.*

(1) Es gilt $L(S_{A,b}) \neq \emptyset$ genau dann, wenn $b \in$ 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 fuer einen beliebigen Vektor $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$:

$$A.x = x_1 y_1 + \dots + a_n y_n,$$

das heisst, $\mathrm{im}(L_A) = \langle y_1, \dots, y_n \rangle$.

(2) Es sei $y \in L(S_{A,b})$ und $z \in L(S_A)$. Dann gilt

$$A(y+z) = A.y + A.z = b + 0_{K^m} = b,$$

so dass $y + z \in L(S_{A,b})$.

Wenn umgekehrt $y' \in L(S_{A,b})$, dann gilt

$$A(y - y') = A.y - A.y' = b - b = 0_{K^m},$$

mit anderen Worten $y' \in y + L(S_A)$. □

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 Quotientenraeumen.*

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 aequivalent:*

(1) Es gilt $\mathrm{rk}(A|b) = \mathrm{rk}(A)$.

(2) Das lineare Gleichungssystem
$$S_{A,b} :\ A.x = b$$
hat eine Loesung.

Proof. Es seien $y_1, \dots, y_n$ die Spaltenvektoren von A. Dann gilt
$$\begin{aligned} & \mathrm{rk}(A|b) = \mathrm{rk}(A) \\ \Leftrightarrow\ & \text{Spaltenrang von } A|b = \text{Spaltenrang von } A \\ \Leftrightarrow\ & \langle y_1, \dots, y_n, b\rangle = \langle y_1, \dots, y_n\rangle \\ \Leftrightarrow\ & b \in \langle y_1, \dots, y_n\rangle \\ \Leftrightarrow\ & b \in \mathrm{Spalten}(A) \\ \Leftrightarrow\ & L(S_{A,b}) \neq \emptyset, \end{aligned}$$
wobei die letzte Aequivalenz 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 aequivalent:*

(1) Es gilt $\mathrm{rk}(A|b) = \mathrm{rk}(A) = n$.

(2) Das lineare Gleichungssystem
$$S_{A,b} :\ A.x = b$$
hat genau eine Loesung.

Proof. Uebung. □

5.1. **Gruppen.**

Definition 5.1.1. *Eine* Gruppe *ist eine Menge G zusammen mit einer Operation $\star : G \times G \to G$, die folgende Axiome erfuellt:*

(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 $K^\times = K \backslash \{0\}$. 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 sogenanne *dihedrale Gruppe D_8*.
(2) Es gibt 48 Symmetries 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 $\mathrm{GL}_n(K)$ die Menge aller invertierbaren Matrizen in $M_{n\times n}(K)$. Dann ist $\mathrm{GL}_n(K)$ eine Gruppe unter Matrix-Multiplikation; sie heisst die *allgemeine lineare Gruppe vom Grad n*. Wenn $n > 1$, dann ist $\mathrm{GL}_n(K)$ nicht abelsch. (Was ist $\mathrm{GL}_1(K)$?) Die Gruppe $\mathrm{GL}_n(K)$ had viele interessante Untergruppen: zum Beispiel die Untergruppe aller oberen Dreiecksmatrizen, oder die (abelsche) Untergruppe diagonaler 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 $\mathrm{GL}_n(\mathbb{F}_p)$?*

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.

[24]Man kann zeigen, dass das neutrale Element e eindeutig bestimmt ist.
[25]Man kann zeigen, dass g^{-1} eindeutig bestimmt ist; es heisst das Inverse von g.
[26]Ueber 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 klassifizert (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.* $8 \cdot 10^{53}$ *Elemente*[27]*) 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 Borcherds 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.*

[27]Zum 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 gilt*[28]
$$1_R \times a = a \times 1_R = a.$$
(4) *Multiplikation ist distributiv bezueglich der Addition, d.h.*
$$a \times (b + c) = a \times b + a \times c \qquad \text{und} \qquad (b + c) \times a = b \times a + c \times a.$$

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 Fontaine*[29]*, da er aeussert wichtige sogenannte 'Periodenringe' (mit den schoenen Namen $\mathbf{B}_{\mathrm{dR}}$, $\mathbf{B}_{\mathrm{HT}}$, $\mathbf{B}_{\mathrm{cris}}, \ldots$) in der p-adischen Hodge Theorie eingefuehrt hatte.*

[28] Das Element ist die multiplikative Identitaet.

[29] ganz egal, was Tolkien dazu sagt

Lecture 18

6.1. **Definition und erste Eigenschaften.**

Definition 6.1.1. *Es seien V, W K-Vektorraeume. Wir schreiben $\mathrm{Hom}_K(V, W)$ fuer die Menge aller linearen Abbildungen von V nach W.*

Satz 6.1.2. *Es seien V, W K-Vektorraeume. Dann hat $\mathrm{Hom}_K(V, W)$ ebenfalls die Struktur eines K-Vektorraums, mit den folgenden Operationen:*

(1) *Es seien $T_1, T_2 \in \mathrm{Hom}_K(V, W)$. Dann ist $T_1 + T_2$ definiert durch*
$$(T_1 + T_2)(v) = T_1(v) + T_2(v) \qquad \forall\, v \in V.$$
(2) *Es seien $T \in \mathrm{Hom}_K(V, W)$ und $\alpha \in K$. Dann ist*
$$(\alpha T)(v) = \alpha T(v) \qquad \forall\, v \in V.$$

Proof. Wir zeigen zunaechst, dass $T_1 + T_2$ und αT ebenfalls Elemente von $\mathrm{Hom}_K(V, W)$ 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 $\mathrm{Hom}_K(V, W)$ mit diesen Operationen alle Vektorraumaxiome erfuellt. Wir zeigen die Existenz des neutralen Elements; der Rest ist eine Uebung.

Es sei $0 : V \to W$ 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 $T \in \mathrm{Hom}_K(V, W)$. 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*
$$\Psi_{\mathcal{C}}^{\mathcal{B}} : \mathrm{Hom}_K(V, W) \to M_{m\times n}(K), \qquad T \mapsto [T]_{\mathcal{C}}^{\mathcal{B}}$$
linear und ein Isomorphismus.

Proof. Selbstverstanedlich 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 4.2.19 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 $\mathrm{Hom}(V, W)$, und es gilt*
$$\dim_K \mathrm{Hom}(V, W) = \dim_K V \cdot \dim_K W.$$

Proof. Es seien $n = \dim V$ und $m = \dim W$. Dank Theorem 6.1.3 wissen wir, dass $\mathrm{Hom}_K(V, W) \cong M_{m\times n}(K)$. 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* $\mathrm{Hom}(V,V)$ *ein Ring unter der Addition und Komposition von Funktionen. Weiterhin ist die Abbildung* $\Psi_{\mathcal{B}}^{\mathcal{B}} : \mathrm{Hom}_K(V,V) \to M_{n\times n}(K)$, *fuer eine Basis $\mathcal{B}$ von V, ein Ring-Isomorphismus.*

Proof. (Skizze) Die Ring-Struktur von $\mathrm{Hom}(V,V)$ 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. $\square$

6.2. **Der duale Vektorraum.** Ein sehr wichtiges Beispiel von $\mathrm{Hom}(V, W)$ ist der Fall, wenn $W = K$.

Definition 6.2.1. *Es sei V ein K-Vektorraum. Der* Dualraum von V *ist der Vektorraum*

$$V^* = \mathrm{Hom}_K(V, K).$$

Die Elemente von V^* sind lineare Abbildungen $V \to K$; sie heissen Linearformen.

Beispiele 6.2.2.

(1) Es sei $n \geq 1$. Dann ist die Spurabbildung

$$\mathrm{Tr} : M_{n\times n}(K) \to K, \qquad A = (a_{ij}) \mapsto a_{11} + \cdots + a_{nn}$$

eine Linearform.

(2) Es sei V ein endlich-dimensionaler Vektorraum mit Basis $v_1 \dots v_n$. Es sei $\ell : V \to K$ die Abbildung, die einen Vektor

$$v = a_1 v_1 + \cdots + a_n v_n$$

auf a_1 abbildet. Dann ist V eine Linearform.

(3) Es sei V der reelle Vektorraum aller stetigen Funktionen $[0, 1] \to \mathbb{R}$, uns 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 Functionen $(0, 1) \to \mathbb{R}$, 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 $[0, 1] \to \mathbb{R}$. 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 ℓ, $\lambda \in V^*$ genau dann gleich, wenn

$$\ell(v_i) = \lambda(v_i) \qquad \forall 1 \leq i \leq n$$

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 $\ell : V \to K$ eine Linearform; wir muessen zeigen, dass sich ℓ eindeutig als Linearkombination von $v_1^*, \dots, v_n^*$ schreiben laesst. Definiere die Linearform

$$f = \ell(v_1)v_1^* + \cdots + \ell(v_n)v_n^*. \tag{28}$$

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

$$\alpha_1 v_1^* + \cdots + \alpha_n v_n^* = 0_{V^*}; \tag{29}$$

hier bezeichnet 0_{V^*} die Linearform, die jeden Vektor $v \in V$ auf $0 \in K$ abbildet. Insbesondere koennen wir (29) 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 $v_1^*, \ldots, v_n^*$ sind linear unabhaengig 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 berachten folgende Linearform

$$\ell : V \to \mathbb{R}. \qquad \ell(f(x)) = Df(2),$$

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 $\mathcal{B}^* = (v_1^*, \ldots, v_n^*)$ ist die* duale Basis *von $\mathcal{B}$.*

Es seien nun $\mathcal{B} = (v_1, \ldots, v_n)$ und $\mathcal{C} = (w_1, \ldots, w_n)$ Basen von V. Wir erinnern uns: die Basiswechselmatrix $A = [\mathrm{id}]_{\mathcal{C}}^{\mathcal{B}}$ ist die Matrix $A = (a_{ij})$, deren Eintraege definiert sind durch die Gleichungen

$$v_j = \sum_{i=1}^{n} a_{ij} w_i. \tag{30}$$

Was gilt nun fuer die Basiswechselmatrizen $[\mathrm{id}]_{\mathcal{C}^*}^{\mathcal{B}^*}$ und $[\mathrm{id}]_{\mathcal{B}^*}^{\mathcal{C}^*}$?

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

$$[\mathrm{id}]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix},$$

das heisst

$$v_1 = e_1 + 2e_2 \qquad \text{und} \qquad v_2 = 3e_1 - e_2. \tag{31}$$

Betrachten wir nun die dualen Basen $\mathcal{B}^* = (v_1^*, v_2^*)$ und $\mathcal{C}^* = (e_1^*, e_2^*)$. Schreiben wir $[\mathrm{id}]_{\mathcal{C}^*}^{\mathcal{B}^*} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, so gilt

$$v_1^* = ae_1^* + ce_2^* \qquad \text{und} \qquad v_2^* = be_1^* + de_2^*. \tag{32}$$

Wenden wir (32) auf (31) an, so erhalten wir

$$[\mathrm{id}]_{\mathcal{C}^*}^{\mathcal{B}^*} = \begin{pmatrix} \frac{1}{7} & \frac{2}{7} \\ \frac{3}{7} & -\frac{1}{7} \end{pmatrix}.$$

Ok. Vielleicht haben wir mit $[\mathrm{id}]_{\mathcal{B}^*}^{\mathcal{C}^*}$ mehr Glueck? Schreiben wir $[\mathrm{id}]_{\mathcal{B}^*}^{\mathcal{C}^*} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, so gilt

$$e_1^* = \alpha v_1^* + \gamma v_2^* \qquad \text{und} \qquad e_2^* = \beta v_1^* + \delta v_2^*. \tag{33}$$

Wenden wir (33) auf (31) an, so erhalten wir

$$[\mathrm{id}]_{\mathcal{B}^*}^{\mathcal{C}^*} = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}.$$

Faellt Ihnen was auf? Lecture 19

Satz 6.2.9. *Es sei V ein endlich-dimensionaler K-Vektorraum mit Basen $\mathcal{B} = (v_1, \ldots, v_n)$ und $\mathcal{C} = (w_1, \ldots, w_n)$. Dann gilt*

$$[\mathrm{id}]_{\mathcal{B}^*}^{\mathcal{C}^*} = \left([\mathrm{id}]_{\mathcal{C}}^{\mathcal{B}}\right)^t.$$

[30] Anstatt die linere Unabhaengigkeit zu beweisen, haetten wir ebenfalls Korollar 6.1.4 benutzen koennen: $v_1^*, \ldots, v_n^*$ ist ein Erzeugendensystem mit n Elementen in einem n-dimensionalen Vektorraum und daher eine Basis.

Proof. Wir argumentieren wie in dem Beispiel. Schreibe $[\mathrm{id}]_{\mathcal{C}}^{\mathcal{B}} = (a_{ij})$ und $[\mathrm{id}]_{\mathcal{B}^*}^{\mathcal{C}^*} = (b_{ij})$, das heisst

$$v_j = \sum_{i=1}^{n} a_{ij} w_i, \tag{34}$$

$$w_\ell^* = \sum_{k=1}^{n} b_{k\ell} v_k^*. \tag{35}$$

Wenden wir nun (35) auf (34) 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

$$b_{j\ell} = a_{\ell j} \qquad \forall\, j, \ell.$$

□

Korollar 6.2.10. *Weiterhin gilt*

$$[\mathrm{id}]_{\mathcal{C}^*}^{\mathcal{B}^*} = \left([\mathrm{id}]_{\mathcal{B}}^{\mathcal{C}}\right)^t . = \left[\left([\mathrm{id}]_{\mathcal{C}}^{\mathcal{B}}\right)^{-1}\right]^t .$$

Proof. Folgt umittelbar 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

$$\left([\mathrm{id}]_{\mathcal{C}^*}^{\mathcal{B}^*}\right)^t = \begin{pmatrix} \frac{1}{7} & \frac{3}{7} \\ \frac{2}{7} & -\frac{1}{7} \end{pmatrix}$$

und weiterhin

$$\begin{pmatrix} \frac{1}{7} & \frac{3}{7} \\ \frac{2}{7} & -\frac{1}{7} \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{7} & \frac{3}{7} \\ \frac{2}{7} & -\frac{1}{7} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

6.3. **Die duale Abbildung.**

Definition 6.3.1. *Es seien V, W K-Vektorraeume, und es sei $T : V \to W$ eine lineare Abbildung. Definiere die* duale Abbildung

$$T^* : W^* \to V^*, \qquad \ell \mapsto \ell \circ T.$$

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 $V \to K$.*

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. $T^*(\ell) \in V^*$. □

Satz 6.3.4. *Die duale Abbildung $T^* : W^* \to V^*$ 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 $T^*(\ell_1 + \ell_2) = T^*(\ell_1) + T^*(\ell_2)$.

□

Satz 6.3.5. *Es seien U, V, W Vektorraeume ueber K, und es seien $T : U \to V$ und $S : V \to W$ lineare Abbildungen. Dann gilt*

$$(S \circ T)^* = T^* \circ S^*.$$

Proof. Uebung. □

Beispiele 6.3.6.

(1) Es gilt $(\mathrm{id}_V)^* = \mathrm{id}_{V^*}$.
(2) Es gilt $(0 : V \to V)^* = (0 : V^* \to V^*)$.

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 $T : V \to W$ eine neue lineare Abbildung $T^* : W^* \to V^*$, 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 $T : V \to W$ eine lineare Abbdildung zwischen endlich-dimensionalen K-Vektorraeumen, und es seien $\mathcal{B}$ und $\mathcal{C}$ Basen von V und W. Wir koennen dann die Abbildungsmatrix $[T]^{\mathcal{B}}_{\mathcal{C}}$ von T bezueglich $\mathcal{B}$ und $\mathcal{C}$ betrachten. Was koennen wir ueber die Abbdildungsmatrix $[T^*]^{\mathcal{C}^*}_{\mathcal{B}^*}$ sagen?

Beispiel 6.3.8. Es sei $V = \mathbb{R}^2$, und es sei $T : V \to V$ 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.

$$T(e_1) = e_1 + 3e_2 \qquad \text{und} \qquad T(e_2) = 2e_1 - e_2. \tag{36}$$

Es sei $B = [T^*]^{\mathcal{E}^*}_{\mathcal{E}^*}$, d.h. wenn $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, dann gilt

$$T^*(e_1^*) = ae_1^* + ce_2^* \qquad \text{und} \qquad T(e_2^*) = be_1^* + de_2^*. \tag{37}$$

[31] Wir benutzen hier, dass zwei Elemente von V^* gleich sind, wenn sie fuer alle $v \in V$ den gleichen Wert annehmen.

Um die Eintraege von B zu bestimmen, benutzen wir wieder die definierenden Eigenschaft der dualen Basis. Indem wir (37) 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 $\mathcal{B} = (v_1, \ldots, v_n)$, bzw. $\mathcal{C} = (w_1, \ldots, w_m)$, eine Basis for V, bzw. fuer W. Es sei $T : V \to W$ eine lineare Abbildung. Dann gilt*[32]

$$[T^*]_{\mathcal{B}^*}^{\mathcal{C}^*} = \left([T]_{\mathcal{C}}^{\mathcal{B}}\right)^t.$$

Proof. Es sei

$$A = [T]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

das heisst

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

Was ist $T^*(w_k^*)$? Wir benutzen (28): es gilt[33]

$$T^*(w_k^*) = [(T^*(w_k^*))(v_1)] \cdot v_1^* + \cdots + [(T^*(w_k^*))(v_n)] \cdot v_n^*,$$

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

$$T^*(w_k^*) = a_{k1} v_1^* + \cdots + a_{kn} v_n^*.$$

Daher ist die Abbildungsmatrix $[T^*]_{\mathcal{B}^*}^{\mathcal{C}^*}$ gegeben durch

$$[T^*]_{\mathcal{B}^*}^{\mathcal{C}^*} = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ a_{12} & a_{22} & \ldots & a_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{pmatrix} = A^t$$

□

Bemerkung 6.3.10. *Es sei $A \in M_{m \times n}(K)$, und es sei $T_A : K^n \to K^m$ die bezueglich der Standardbasen $\mathcal{E}$ und $\mathcal{F}$ von K^n und K^m. Dann ist $[T_A^*]_{\mathcal{E}^*}^{\mathcal{F}^*} = A^t$. Vielleicht hilft uns das, um einen neuen Beweis zu finden fuer den Satz, dass*

$$\text{Spaltenrang}(A) = \text{Spaltenrang}(A^t)?$$

Wir bemerken, dass $\text{Spaltenrang}(A) = \dim \operatorname{im}(T_A)$ *und* $\text{Spaltenrang}(A^t) = \dim \operatorname{im}(T_A^*)$. *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 erklaeren, 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.*

---

[32] Sie koennen sich das folgendermassen merken: die Abbildungsmatrix von T is 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.

[33] Beachten 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.** Lecture 20

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 Bedinungen

$$a - b + c = 0 \quad \text{und} \quad a + b = 0,$$

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 $u_1, \ldots, u_k$ von U; dann koennen wir sie laut Theorem 3.3.19 zu einer Basis $\mathcal{B} = (u_1, \ldots, u_k, v_1, \ldots, v_{n-k})$ von V erweitern. Es sei $\mathcal{B}^* = (u_1^*, \ldots, u_k^*, v_1^*, \ldots, v_{n-k}^*)$ 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, $v_j^* \in U^\circ$ $\forall\, 1 \leq j \leq n - k$ und daher $\langle v_1^*, \ldots, u_{n-k}^* \rangle \subseteq U^\circ$.

Wir behaupten, dass $v_1^* \ldots, v_{n-k}^*$ eine Basis von U° ist. Natuerlich sind die Vektoren linear unabhaengig, es ist daher zu zeigen, dass $\langle v_1^*, \ldots, v_{n-k}^* \rangle = U^\circ$. Es sei $\lambda \in U^\circ$. Da $\mathcal{B}^*$ eine Basis von V^* ist, gibt es Skalare $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_{n-k}$, so dass

$$\lambda = \alpha_1 u_1^* + \cdots + \alpha_k u_k^* + \beta_1 v_1^* + \cdots + \beta_{n-k} v_{n-k}^*.$$

Da $\lambda \in U^\circ$, gilt

$$\lambda(u_i) = \alpha_i = 0 \qquad \forall\, 1 \leq i \leq k,$$

und wir erhalten

$$\lambda = \beta_1 v_1^* + \cdots + \beta_{n-k} v_{n-k}^*,$$

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 $\iota : U \to V$ die Inklusionsabbildung und $\iota^* : V^* \to U^*$ die duale Abbildung. Dann gilt laut Theorem 4.2.9 die Formel

$$\dim(V) = \dim(V^*) = \dim\ker(\iota^*) + \dim\operatorname{im}(\iota^*) \tag{38}$$

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 $P : V \to U$ eine Projektion[34], 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

$$\begin{aligned} \ell \in \ker(\iota^*) \quad &\Leftrightarrow \quad \iota^*(\ell) == \ell \circ \iota = 0_{V^*} \\ &\Leftrightarrow \quad \forall u \in U: \ \ell(u) = \ell(\iota(u)) =)_{V^*}(u) = 0 \\ &\Leftrightarrow \quad \ell \in U^\circ. \end{aligned}$$

Von (38) und den beiden Behauptungen folgt daher, dass

$$\dim(V) = \dim(U^*) + \dim(U^\circ).$$

Doch da $\dim(U^*) = \dim(U)$, folgt das Theorem. □

Satz 6.4.6. *Es seien V, W endlich-dimensionale K-Vektorraeume und $T : V \to W$ 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

$$\lambda(T(v)) = 0 \qquad \forall\, v \in V$$

und daher $\lambda \in (\operatorname{im} T)^\circ$.

Umgekehrt sei $\lambda \in (\operatorname{im} T)^\circ$, das heisst $\lambda(T(v)) = 0\ \forall\, v \in V$. 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

$$\lambda(v) = T^*(\psi)(v) = \psi(T(v)) = 0 \qquad \forall\, v \in \ker(T),$$

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 $\operatorname{im}(T^*) = (\ker T)^\circ$.

□

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 $T_A : K^n \to K^m$ 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. □

[34]Wir 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):

$$\begin{array}{rcll} T \text{ ist injektiv} & \Leftrightarrow & \ker(T) = \{0_V\} & \\ & \Leftrightarrow & (\ker T)^\circ = V^* & \\ & \Leftrightarrow & \operatorname{im}(T^*) = V^* & \text{Satz 6.4.6} \\ & \Leftrightarrow & T^* \text{ ist surjektiv} & \end{array}$$

□

6.5. **Reflexivitaet.** Was passiert, wenn wir einen Vektorraum zweimal dualisieren? Lecture 21

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 $V^* \to K$. Koennen wir solche Abbdildungen konstruieren?*

Beachte 6.5.3. *Es sei $v \in V$. Dann definiert die* Auswertung an v *eine Abbildung $V^* \to K$: explizit ist sie definiert durch*

$$\tau_v : V^* \to K, \qquad \ell \mapsto \ell(v).$$

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 $\tau : V \to V^{**}$, $V \mapsto \tau_v$ 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 $V \to V^{**}$, 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:

$$\tau_{\lambda v}(\ell) = \ell(\lambda v) = \lambda \ell(v) = \lambda \tau_v(\ell) \qquad \forall\, \ell \in V^*,$$

das heisst $\tau_{\lambda v} = \lambda \tau_v$.

(2) Es seien $u, v \in V$. Dann gilt

$$\tau_{u+v}(\ell) = \ell(u+v) = \ell(u) + \ell(v) = \tau_u(\ell) + \tau_v(\ell) \qquad \forall\, \ell \in V^*,$$

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

$$\tau_v(\ell) = \ell(v) = 0 \qquad \forall \ell \in V^*.$$

Es sei $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V, und schreibe $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Dann gilt

$$v_i^*(v) = \alpha_i = 0 \qquad \forall 1 \le i \le n,$$

und daher $v = 0$. Daher ist τ injective und ein Isomorphismus. □

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.

$$\tau_v(\ell) = 0 \quad \forall\, \ell \in V^* \qquad \Leftrightarrow \qquad \ell(v) = 0 \quad \forall\, \ell \in V^*.$$

Dann folgt aus der Linearitaet von ℓ, dass

$$\ell(w) = 0 \qquad \forall w \in \langle v \rangle, \quad \forall \ell \in V^*,$$

das heisst $\ell \in \langle v \rangle^\circ\ \forall\, \ell \in V^*$ 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. □

Bemerkung 6.5.6. *Die Abbildung $\tau : V \to V^{**}$ 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 $\mathcal{B}^* = (v_1^*, \ldots, v_n^*)$ die duale Basis, und $\mathcal{B}^{**} = ((v_1^*)^*, \ldots, (v_n^*)^*)$ die duale Basis von $\mathcal{B}^*$. Dann gilt*

$$\tau_{v_i} = (v_i^*)^* \qquad \forall i$$

das heisst, τ bildet die Basis $\mathcal{B}$ auf die bidule Basis $\mathcal{B}^{**}$ ab.

Proof. Per Definition ist $\mathcal{B}^{**} = ((v_1^*)^*, \ldots, (v_n^*)^*)$ diejenige Basis von V^{**}, fuer die gilt

$$(v_i^*)^*(v_j^*) = \delta_{ij}.$$

Wir muessen also zeigen, dass $\{\tau_{v_1}, \ldots, \tau_{v_n}\}$ 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 $T : V \to W$ eine lineare Abbildung. Dann sind die Abbildungen τ^V und τ^W mit T kompatibel: wir haben ein kommutative Diagramm*

$$\begin{tikzcd} V \arrow[r, "\tau^V"] \arrow[d, "T"'] & V^{**} \arrow[d, "T^{**}"] \\ W \arrow[r, "\tau^W"'] & W^{**} \end{tikzcd}$$

$V \mapsto V^{**}$ ist ein kovarianter Funktor. *Nun seien $\mathcal{B}$ und $\mathcal{C}$ jeweilige Basen von V und W. Dann gilt*

$$[T^{**}]_{\mathcal{C}^{**}}^{\mathcal{B}^{**}} = \left([T^*]_{\mathcal{B}^*}^{\mathcal{C}^*}\right)^t = \left(\left([T]_{\mathcal{C}}^{\mathcal{B}}\right)^t\right)^t = [T]_{\mathcal{C}}^{\mathcal{B}},$$

7.1. **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 Aequivalenzklassen folgender Aequivalenzrelation 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*

$$[y_1] + [v_2] = [v_1 + v_2] \qquad \textit{und} \qquad \alpha\,[v] = [\alpha v].$$

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

$$[v_1] + [v_2] = [v_1'] + [v_2'] \quad \Leftrightarrow \quad [v_1 + v_2] = [v_1' + v_2'] \quad \Leftrightarrow \quad v_1 + v_2 - v_1' - v_2' \in U.$$

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 mussen zeigen, dass

$$\alpha[v] = \alpha[v'] \quad \Leftrightarrow \quad [\alpha v] = [\alpha v'] \quad \Leftrightarrow \quad \alpha(v - v') \in U.$$

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 $q_U : V \to V/U$ die Abbildung*

$$q_U(v) = [v];$$

sie heisst die kanonische Quotientenabbildung. *Dann ist q_U linear, und es gilt*[36] $\ker(q_U) = U$ *und* $\operatorname{im}(q_U) = V/U$.

Proof. Wir zeigen zunaechst, dass die Abbildung linear ist:

$$\begin{aligned} q_U(v_1 + \alpha v_2) &= [v_1 + \alpha v_2] \\ &= [v_1] + \alpha[v_2] \qquad \text{aufgrund der Definition von Addition und Skalar-Multiplikation} \\ &= q_U(v_1) + \alpha\, q_U(v_2). \end{aligned}$$

Es sei nun $x \in V/U$. Dann ist x eine Aequivalenzklasse der Aequivalenzrelation 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

$$\begin{aligned} q_U(v) = [v] = 0_{V/U} \quad &\Leftrightarrow \quad v \sim 0_V \\ &\Leftrightarrow v - 0_V \in U \\ &\Leftrightarrow v \in U, \end{aligned}$$

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:

[35] Was ist die additive Identitaet?

[36] Mit 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. Lecture 22
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 $\gamma(w) = [v] = x$. □

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 $w_1, \ldots, w_k$ eine Basis von W ist, dann ist $[w_1], \ldots, [w_k]$ eine Basis von W, 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 bestimmen Unterraeumen von V und Unterraeumen von V/U:

Satz 7.1.10. *Die Abbildung*

$$\begin{array}{ccc} \{W \subseteq V : U \leq W \leq V\} & \to & \{X \leq V/U\} \\ W & \mapsto & W/U \end{array}$$

ist eine $1:1$ Korrespondenz zwischen Unterraeumen von V, die U enthalten, und Unterraeumen von V/U.

Proof. Uebung. □

[37] Sie koennen alternativ mit Hilfe der Dimensionen argumentieren.

7.2. **Die Isomorphismensaetze.** Quotientenraeume tauchen in the linearen Algebra (und auch sonst) ueberall auf. Zum Beispiel koennen wir Theorem 4.2.9 folgendermassen verfeinern:

Theorem 7.2.1. *(Erster Isomorphiesatz) Es seien V, W K-Vektorraeume und $T : V \to W$ eine lineare Abbildung. Definiere eine neue Abbildung*

$$\overline{T} : V/\ker(T) \to \operatorname{im}(T)$$

wie folgt: fuer ein Element $x \in V/\ker(T)$ waehle ein representierendes Element $v \in V$ und definiere $\overline{T}(x) = T(v)$.[38] *Dann gilt*

(1) $\overline{T}$ eine wohldefinierte lineare Abbildung;
(2) $\overline{T} : V/\ker(T) \to \operatorname{im}(T)$ ist ein Isomorphismus;
(3) folgendes Diagram ist kommutativ:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "q_{\ker(T)}"'] & W \\ V/\ker(T) \arrow[r, "\overline{T}"] & \operatorname{im}(T) \arrow[u, "\subseteq"'] \end{tikzcd}$$

Proof. (1) Es seien v_1, v_2 representierende 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.

$\overline{T}$ ist linear: es seien $x = [v_1], y = [v_2]$ und $\alpha \in K$. Dann gilt

$$\begin{aligned}\bar{T}([v_1] + \alpha[v_2]) &= \bar{T}([v_1 + \alpha v_2]) \\ &= T(v_1 + \alpha v_2) \\ &= T(v_1) + \alpha\, T(v_2) \\ &= \overline{T}([v_1]) + \alpha\overline{T}([v_2]).\end{aligned}$$

(2) $\overline{T}$ ist injektiv: Es sei $x \in \ker(\overline{T})$, und es sei $v \in V$ ein representierendes Element von x. Dann gilt aufgrund der Definition

$$\overline{T}(x) = T(v) = 0_W,$$

das heisst $v \in \ker(T)$ und daher $x = [v] = 0_{V/\ker(T)}$.

$\overline{T}$ ist surjektiv: Aufgrund der Definition von $\overline{T}$ ist es offensichtlich, dass $\operatorname{im}(\overline{T}) \subseteq \operatorname{im}(T)$. Es sei $w \in \operatorname{im}(T)$. Dann gibt es ein $v \in V$ so dass $T(v) = w$. Aber dies bedeutet, dass

$$\overline{T}([v]) = T(v) = w,$$

und daher $w \in \operatorname{im}(\overline{T})$.

(3) Die Kommutativitaet des Diagrams folgt unmittelbar von den Definitionen.

□

Theorem 7.2.2. *(Zweiter Isomorphiesatz) Es sei V ein Vektorraum und $U, W \leq V$. Es sei*

$$i : U \hookrightarrow V \xrightarrow{q_W} V/W, \qquad u \mapsto q_W(u)$$

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 zunaechst, 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 muessen 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 repraesentierendes Element $u \in U$ von x, und es gilt $i(u) = x$. □

[38] Wir nennen $\overline{T}$ die *von T induzierte Abbildung*.

[39] Beachte, dass $(U + W)/W$ ein Unterraum von V/W ist.

Was passiert, wenn wir zwei Unterraeume 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 Unterraeume. Dann gibt es eine natuerliche lineare Abbildung $\varpi_{U,W} : V/U \mapsto V/W$, gegeben durch $v + U \mapsto v + W$, so dass folgendes Diagram kommutiert:*

$$\begin{tikzcd} & V \arrow[dl, "q_U"'] \arrow[dr, "q_W"] & \\ V/U \arrow[rr, "\varpi_{U,W}"] & & V/W \end{tikzcd}$$

Proof. Eine einfache Rechnung zeigt, dass $\varpi_{U,W}$ linear ist. Die Kommutativitaet des Diagrams ist eine Uebung. □

Theorem 7.2.4. *(Dritter Isomorphiesatz) Es sei V ein Vektorraum und $U \leq W \leq V$ zwei Unterraeume. 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

$$\varpi_{U,W} : V/U \longrightarrow V/W$$

am. Die Abbildung ist surjektiv: Es sei $x \in V/W$ und v ein repraesentierendes 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 repraesentierendes 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. □

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

$$\ell \circ q_U : V \to K$$

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

$$\overline{\lambda} : V/\ker(\lambda) \to K$$

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

$$\alpha^{-1}(\lambda) = \overline{\lambda} \circ \varpi_{U,\ker(\lambda)} : V/U \to K.$$

Eine einfache Rechnung zeigt, dass α^{-1} die inverse Abbildung von α ist. $\square$

Lecture 23

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 characteristisch 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* $\tau^2 = \mathrm{id}$, *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

$$\tau_n\tau_{n-1}\cdots\tau_1\sigma = \mathrm{id} \quad \Rightarrow \quad \sigma = \tau_1\tau_2\cdots\tau_n,$$

da $\tau_i^2 = \mathrm{id}\ \forall i$ 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 $\tau_2 \tau_1 \sigma = \mathrm{id}$ 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 Paritaet der Anzahl von Transposition 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
$$k \cong m \pmod 2.$$

Proof. Spaeter - 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 laesst. Wir definieren das* Vorzeichen *einer Permutation σ als*
$$\mathrm{sgn}(\sigma) = \begin{cases} 1 & \text{wenn } \sigma \text{ gerade ist} \\ -1 & \text{wenn } \sigma \text{ ungerade ist.} \end{cases}$$

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 aufloesbare Gruppe*: damit werden Sie sich beschaeftigen, wenn Sie im naechsten Jahr Galois Theorie lernen.*

8.3. **Determinantenfunktionen.**

Lecture 24

Notation 8.3.1. Es sei $A \in M_{n\times n}(K)$; wir schreiben $A = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, wobei $\mathbf{v}_i$ der $i$te Spaltenvektor ist.

Definition 8.3.2. *Eine Funktion $f : M_{n\times n}(K) \to K$ ist* n-linear *(in den Spaltenvektoren), wenn $\forall 1 \leq i \leq n$ f eine lineare Funktion der iten Spalte ist, wenn die die anderen Spalten fixiert werden. Mit anderen Worten, f ist n-linear, wenn*

$$\begin{aligned} f(\mathbf{v}_1, \ldots, \mathbf{v}_i + \mathbf{u}, \ldots, \mathbf{v}_n) &= f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_n) + f(\mathbf{v}_1, \ldots, \mathbf{u}, \ldots, \mathbf{v}_n), \\ f(\mathbf{v}_1, \ldots, \lambda\mathbf{v}_i, \ldots, \mathbf{v}_n) &= \lambda f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_n) \end{aligned}$$

fuer alle $A = (\mathbf{v}_1, \ldots, \mathbf{v}_n) \in M_{n\times n}(K)$, $u \in K^n$ und $\lambda \in K$.

Beispiele 8.3.3.

(1) $\det : M_{2\times 2}(K) \to K$ ist bilinear.
(2) Es sei $A = (a_{ij}) \in M_{n\times n}(K)$. Die Funktion
$$f : A \mapsto a_{11}a_{22}\ldots a_{nn}$$
ist n-linear.

Lemma 8.3.4. *Es seien $f, g : M_{n\times n}(K) \to K$ 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 $f : M_{n\times n}(K) \to K$ ist* alternierend*, wenn folgende Eigenschaft erfuellt ist: wenn $\mathbf{v}_i = \mathbf{v}_{i+1}$ fuer ein $1 \leq i < n$, dann gilt*

$$f(\mathbf{v}_1, \ldots, \mathbf{v}_n) = 0.$$

Lecture 25

Beispiel 8.3.6. $\det : M_{2\times 2}(K) \to K$ ist alternierend.

Lemma 8.3.7. *Es sei $f : M_{n\times n}(K) \to K$ n-linear und alternierend. Dann gilt*

(1) $f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n) = -f(\mathbf{v}_1, \ldots, \mathbf{v}_{i+1}, \mathbf{v}_i, \ldots, \mathbf{v}_n)$;
(2) wenn $\mathbf{v}_i = \mathbf{v}_j$ fuer $i \neq j$, dann gilt $f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_n) = 0$;
(3) $f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_n) = -f(\mathbf{v}_1, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_n)$.

Proof. (1) Es gilt

$$\begin{aligned} f(\mathbf{v}_1, \ldots, \mathbf{v}_i + \mathbf{v}_{i+1}, \mathbf{v}_i + \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n) &= f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_n) + f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n) \\ &\quad + f(\mathbf{v}_1, \ldots, \mathbf{v}_{i+1}, \mathbf{v}_i, \ldots, \mathbf{v}_n) + f(\mathbf{v}_1, \ldots, \mathbf{v}_{i+1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n) \\ &= f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n) + f(\mathbf{v}_1, \ldots, \mathbf{v}_{i+1}, \mathbf{v}_i, \ldots, \mathbf{v}_n) \\ &= 0 \end{aligned}$$

und daher
$$f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n) = -f(\mathbf{v}_1, \ldots, \mathbf{v}_{i+1}, \mathbf{v}_i, \ldots, \mathbf{v}_n).$$

(2) Nimm nun an, dass $A = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ 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 $D : M_{n\times n}(K) \to K$ 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*

$$\det : M_{2\times 2}(K) \to K$$

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 $D : M_{2\times 2}(K) \to K$ 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* $D : M_{n\times n}(K) \to K$*.*

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*

$$E_i : M_{n\times n}(K) \to K, \qquad A = (a_{ij}) \mapsto \sum_{j=1}^{n} (-1)^{i+j} a_{ij} f(A_{i,j}).$$

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 duch Induktion nach n. Fuer $n = 1$ sei D die Indentitaet, und fuer $n = 2$ die Determinante.

Nimm nun an, dass wir $D_{n-1} : M_{n-1,n-1}(K) \to K$ definiert haben. Es sei $A = (a_{ij})_{1 \leq i,j \leq n} \in M_{n \times n}(K)$. Waehle nunn 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

$$D_n^{(i)}(\mathbf{1}_n) = (-1)^{i+i} D_{n-1}(\mathbf{1}_{n-1}) = 1.$$

□

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

$$D_3^{(1)} : A \mapsto a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix},$$

$$D_3^{(2)} : A \mapsto -a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{22} \det \begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix} - a_{23} \det \begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix},$$

$$D_3^{(3)} : A \mapsto a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} - a_{32} \det \begin{pmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{pmatrix} + a_{33} \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

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

$$k \cong m \pmod 2.$$

Proof. Es sei $D : M_{n \times n}(K) \to K$ eine Determinantenfunktion, uns es sei $\mathbf{e}_1, \ldots, \mathbf{e}_n$ die Stadardbasis von K^n. Wenn σ durch die Verknuepfung von k Transpositionen erhalten werden kann, dann an die Einheitsmatrix $\mathbf{1}_n$ von der Matrix $(\mathbf{e}_{\sigma(1)}, \ldots, \mathbf{e}_{\sigma(n)})$ durch k Vertauschungen von Spalten konstruiert werden. Da das Vertauschen von Spalten das Vorzeichen des Wertes der Determinantenfunktion aendert, erhalten wir

$$D(\mathbf{e}_{\sigma(1)}, \ldots, \mathbf{e}_{\sigma(n)}) = (-1)^k D(\mathbf{1}_n) = (-1)^k.$$

Durch das gleiche Argument erhalten wir

$$D(\mathbf{e}_{\sigma(1)}, \ldots, \mathbf{e}_{\sigma(n)}) = (-1)^m,$$

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 $n = 2$.

[40] Dieses 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 eine*[41] *Determinantenfunktion*

$$D : M_{n\times n}(K) \to K.$$

Explizit ist D durch folgende Formel gegeben: es sei $A = (a_{ij})_{1\leq i,j\leq n}$. Dann gilt

$$D(A) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\, a_{\sigma(1),1}a_{\sigma(2),2}\cdots a_{\sigma(n),n}. \tag{39}$$

Proof. Es sei $\mathbf{e}_1, \ldots, \mathbf{e}_n$ die Standardbasis von K^n; fuer $A = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ schreiben wir

$$\mathbf{v}_i = a_{1i}\mathbf{e}_1 + \cdots + a_{ni}\mathbf{e}_n.$$

Es sei nun $D : M_{n\times n}(K) \to K$ eine Determinantenfunktion.[42] Da D alternierend und n-linear ist, erhalten wir durch Expansion

$$\begin{aligned} D(A) &= D(a_{11}\mathbf{e}_1 + \cdots + a_{n1}\mathbf{e}_n, \ldots, a_{1n}\mathbf{e}_1 + \cdots + a_{nn}\mathbf{e}_n) \\ &= \sum_{\sigma\in S_n} a_{\sigma(1),1}\cdots a_{\sigma(n),n} \cdot D(\mathbf{e}_{\sigma(1)}, \ldots, \mathbf{e}_{\sigma(n)}). \end{aligned}$$

Aber nun gilt

$$D(\mathbf{e}_{\sigma(1)}, \ldots, \mathbf{e}_{\sigma(n)}) = \operatorname{sgn}(\sigma) \cdot D(\mathbf{e}_1, \ldots, \mathbf{e}_n) = \operatorname{sgn}(\sigma),$$

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* (39) *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 $D : M_{n\times n}(K) \to K$ 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 Determinantenfuktion untersuchen, z.B. die Multiplikativitaet (Satz 8.5.2).

[41] Insbesondere sind die drei Formeln in Beispiel 8.3.14 gleich. Rechnen Sie nach!

[42] Wir wissen dank Theorem 8.3.10, dass eine solche Funktion existiert.

8.4. **Erste Eigenschaften.** Eine Eigenschaft der Determinante folgt unmittelbar von (39):

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 ueberraschendes Korollar:

Korollar 8.4.2. *Die Determinante ist ebenfalls n-linear und alternierend als eine Funktion in the Zeilenvektoren.*

Bemerkung 8.4.3. *Daher gilt auch das analoge Resultat von Lemma 8.3.7 fuer jede Funktion $M_{n\times n}(K) \to K$, die n-linear und alternierend in den Zeilenvektoren ist. Weiterhin ist jede n-lineare und alternierende Funktion in den Zeilenvektore 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 $1 \le r < s \le n$, dann gilt $\det(B) = -\det(A)$;
(2) wenn $X = M(r, \lambda)$ fuer $1 \le r \le n$ und $\lambda \in K^\times$, dann gilt $\det(B) = \lambda\det(A)$;
(3) wenn $X = S(r, s, \lambda)$ fuer $1 \le r, s \le n$, $r \ne s$ 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 bedeuted, 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:
$$\begin{aligned}\det(A) &= \det\begin{pmatrix} 0 & -9 & 3\\ 1 & 4 & 0\\ 2 & 6 & -1\end{pmatrix}\\ &= 0\cdot\det\begin{pmatrix} 4 & 0\\ 6 & -1\end{pmatrix} - (-9)\cdot\begin{pmatrix} 1 & 0\\ 2 & -1\end{pmatrix} + 3\cdot\begin{pmatrix} 1 & 4\\ 2 & 6\end{pmatrix}\\ &= -9-6\\ &= -15.\end{aligned}$$

(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*

$$M = \begin{pmatrix} A & B \\ 0_{s\times r} & C \end{pmatrix} \qquad \text{fuer } A \in M_{r\times r}(K),\ B \in M_{r\times s}(K) \text{ und } C \in M_{s\times s}(K).$$

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 Birkhäuser, Jakob Göhring und Marco Vaccaro erhalten.

Korollar 1. *Es sei $M \in M_{n\times n}(K)$ von der Form*

$$M = \begin{pmatrix} 0_{r\times s} & C \\ A & B \end{pmatrix} \qquad \text{fuer } A \in M_{s\times s}(K),\ B \in M_{s\times r}(K) \text{ und } C \in M_{r\times r}(K).$$

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 laesst. Da jede Vertauschung das Vorzeichen aendert und

$$r(s+r-1) \equiv rs \pmod 2,$$

folgt das Resultat von Theorem 8.4.6.

Es sei U die Verknuepfung 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'. $\square$

Bemerkung 8.4.7. *Es ist nicht richtig, dass wir einfach die Transposition auf M anwenden koennen, um eine Matrix der Blockform $\begin{pmatrix} \star\ast\star & \\ 0 & \star \end{pmatrix}$ 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

$$\begin{aligned} \det(B) &= \det \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix} \\ &= (-5) \cdot 4 \\ &= 20. \end{aligned}$$

(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 laesst 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
$$D : M_{n\times n}(K) \longrightarrow K,$$
$$B \mapsto \det(AB).$$

Behauptung. Die Funktion D ist n-linear und alternierend als eine Funktion in den Spalten.[43]

Beweis der Behaptung. Schreibe $A = (\mathfrak{a}_1, \ldots, \mathfrak{a}_n)$ und $B = (\mathfrak{b}_1, \ldots, \mathfrak{b}_n)$. Dann sind die Spalten der Matrix AB gegeben durch
$$AB = (A\mathfrak{b}_1, \ldots, A\mathfrak{b}_n).$$
Daher gilt
$$\begin{aligned} D(\mathfrak{b}_1, \ldots, \mathfrak{b}_i + \mathbf{u}, \ldots, \mathfrak{b}_n) &= \det\left(A\mathfrak{b}_1, \ldots, A(\mathfrak{b}_i + \mathbf{u}), \ldots, A\mathfrak{b}_n\right) \\ &= \det\left(A\mathfrak{b}_1, \ldots, A\mathfrak{b}_i + A\mathbf{u}, \ldots, A\mathfrak{b}_n\right) \\ &= \det\left(A\mathfrak{b}_1, \ldots, A\mathfrak{b}_i, \ldots, A\mathfrak{b}_n\right) + \det\left(A\mathfrak{b}_1, \ldots, A\mathbf{u}, \ldots, A\mathfrak{b}_n\right) \\ &= D\left(\mathfrak{b}_1, \ldots, \mathfrak{b}_i, \ldots, \mathfrak{b}_n\right) + D\left(\mathfrak{b}_1, \ldots, \mathbf{u}, \ldots, \mathfrak{b}_n\right), \end{aligned}$$
wobei die dritte Gleichheit von der n-Linearitaet von det folgt. Daher ist D n-linear.

Nimm nun an, dass $\mathfrak{b}_i = \mathfrak{b}_{i+1}$ fuer ein $1 \leq i < n$. Dann ist $A\mathfrak{b}_i = A\mathfrak{b}_{i+1}$, und daher
$$D(\mathfrak{b}_1, \ldots, \mathfrak{b}_n) = \det(A\mathfrak{b}_1, \ldots, A\mathfrak{b}_n) = 0,$$
da det alternierend ist. Das beweist die Behauptung.

Wir wissen daher von Korollar 8.3.19, dass
$$D(B) = \det(B) \cdot D(\mathbf{1}_n) \qquad \forall\, B \in M_{n\times n}(K).$$
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. *Korollary 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]

[43] Sie koennen alternativ auch die Funktion $A \mapsto \det(AB)$ betrachten; diese ist dann n-linear und alternierend in den Zeilen.

[44] Vergl. mit Korollar 4.7.4.

Proof. Uebung. ☐

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 $V \to V$ zu definieren. Wir kommen im naechsten Kapital 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

$$B = C^{-1}AC \qquad \text{mit } C = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

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

$$A = C^{-1} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \cdot C \qquad \text{mit } C = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

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 $A = (a_{ij})_{1 \le i,j \le n} \in M_{n\times n}(K)$, und es sei $1 \le i \le n$. Dann gilt

$$\det(A) = (-1)^{i+1} a_{i1} \det(A_{i,1}) + \cdots + (-1)^{i+n} a_{in} \det(A_{i,n}), \tag{40}$$

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 $1 \le i, j \le n$ schreiben wir*

$$c_{i,j} = (-1)^{i+j} \det(A_{i,j}). \tag{41}$$

Die Kofaktormatrix von A *ist die Matrix* $C = (c_{i,j}) \in M_{n\times n}(K)$. *Die* adjunkte Matrix $\mathrm{adj}(A)$ von A *ist die Transposition der Kofaktormatrix:*[45]

$$\mathrm{adj}(A) = C^t.$$

Beispiele 8.5.10. (1) Es sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Dann ist

$$\mathrm{adj}(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

(2) Es sei $B = \begin{pmatrix} 4 & 5 & 2 \\ -1 & 0 & 3 \\ 1 & 1 & 2 \end{pmatrix}$. Dann gilt

$$\mathrm{adj}(B) = \begin{pmatrix} -3 & -8 & 15 \\ 5 & 6 & -14 \\ -1 & 1 & 5 \end{pmatrix}.$$

Bemerkung 8.5.11. *Von Lemma 8.3.9 wissen wir, dass*

$$A \cdot \mathrm{adj}(A) = \mathrm{adj}(A) \cdot A = \det(A) \cdot \mathbf{1}_2.$$

Berechnen Sie $B \cdot \mathrm{adj}(B)$, $\mathrm{adj}(B) \cdot B$ *und* $\det(B)$. *Was faellt Ihnen auf?*

[45] Die adjunkte Matrix einer (1×1)-Matrix ist (1)!

Vermutung 8.5.12. *Es sei* $A = (a_{ij}) \in M_{n\times n}(K)$ *und* $\mathrm{adj}(A) = (b_{ij})$ *die adjunkte Matrix; per Definition gilt* $b_{ij} = c_{ji}$*, wobei* $c_{i,j}$ *durch* (41) *gegeben ist. Dann gilt*

$$A \cdot \mathrm{adj}(A) = \mathrm{adj}(A) \cdot A = \det(A) \cdot \mathbf{1}_n. \tag{42}$$

Mit anderen Worten, wenn $\det(A) \neq 0$*, dann ist* A *invertierbar, und es gilt*

$$A^{-1} = \det(A)^{-1} \cdot \mathrm{adj}(A).$$

Um diese Vermutung zu beweisen, muessen wir das Produkt $A \cdot \mathrm{adj}(A)$ ausrechnen. Schreibe

$$M = (m_{ik}) = A \cdot \mathrm{adj}(A).$$

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 (40). □

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

$$\det(A') = 0 = a'_{i1}c'_{i,1} + \cdots + a'_{in}c'_{i,n}. \tag{43}$$

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 (43) 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*

$$A \cdot \mathrm{adj}(A) = \det(A) \cdot \mathbf{1}_n.$$

Proof. Es seien $1 \leq i, k \leq n$. Dann ist der Eintrag der Matrix $A \cdot \mathrm{adj}(A)$ 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

$$\mathrm{adj}(A) \cdot A = \det(A \cdot \mathbf{1}_n.$$

Lemma 8.5.17. *Es gilt*

$$\mathrm{adj}(A) \cdot A = \det(A)\mathbf{1}_n.$$

Proof. Wir wenden Theorem 8.5.16 auf die transponierte Matrix A^t an:

$$A^t \cdot \mathrm{adj}(A^t) = \det(A^t)\mathbf{1}_n.$$

Aber $\det(A^t) = \det(A)$ aufgrund von Satz 8.4.1, und daher

$$A^t \cdot \mathrm{adj}(A^t) = \det(A)\mathbf{1}_n.$$

Wir wenden nun die Transposition auf diese Gleichung an und erhalten

$$(\mathrm{adj}(A^t))^t \cdot A = \det(A) \cdot \mathbf{1}_n.$$

Aber eine einfache Rechnung zeigt, dass $\mathrm{adj}(A^t) = (\mathrm{adj}(A))^t$. Das Lemma folgt. □

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*

$$A^{-1} = \det(A)^{-1} \cdot \mathrm{adj}(A).$$

Proof. Folgt unmittelbar von Lemma 8.5.17. □

Bemerkung 8.5.19. *Nimm an, wir haben ein lineares Gleichungssystem*

$$S:\ A.x = b,$$

wobei A eine invertierbare $(n \times n)$-Matrix ist. In diesem Fall ist die eindeutige Loesung $L(S)$ gegeben durch

$$A^{-1}.b = \det(A)^{-1} \cdot \mathrm{adj}(A).b.$$

In der Praxis ist es aber sehr muehsam, 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 $T : V \to V$ 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 muessen ueberpruefen, dass diese Definition sinnvoll ist, d.h. dass sie unabhaengig 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

$$[T]_{\mathcal{C}}^{\mathcal{C}} = [\mathrm{id}]_{\mathcal{C}}^{\mathcal{B}} \cdot [T]_{\mathcal{B}}^{\mathcal{B}} \cdot [\mathrm{id}_V]_{\mathcal{B}}^{\mathcal{C}}.$$

Weiterhin gilt, dass $[\mathrm{id}_V]_{\mathcal{B}}^{\mathcal{C}} = \left([\mathrm{id}_V]_{\mathcal{C}}^{\mathcal{B}}\right)^{-1}$. Daher gilt

$$\begin{aligned}\det[T]_{\mathcal{C}}^{\mathcal{C}} &= \det\left([\mathrm{id}]_{\mathcal{C}}^{\mathcal{B}} \cdot [T]_{\mathcal{B}}^{\mathcal{B}} \cdot \left([\mathrm{id}_V]_{\mathcal{C}}^{\mathcal{B}}\right)^{-1}\right) \\ &= \det[\mathrm{id}]_{\mathcal{C}}^{\mathcal{B}} \cdot \det[T]_{\mathcal{B}}^{\mathcal{B}} \cdot \det\left([\mathrm{id}_V]_{\mathcal{C}}^{\mathcal{B}}\right)^{-1} \\ &= \det[T]_{\mathcal{B}}^{\mathcal{B}}.\end{aligned}$$

□

Bemerkung 8.6.3. *Fuer unterschiedlich Basen $\mathcal{B}$ und $\mathcal{C}$ von V koennten wir natuerlich auch* $\det[T]_{\mathcal{B}}^{\mathcal{C}}$ *betrachten. Dieser Werte ist* nicht *wohldefiniert!*[46]

Satz 8.6.4. *Es sei V ein endlich-dimensionaler K-Vektorraum. Dann hat die Funktion*

$$\det : \mathrm{Hom}_K(V,V) \to K, \qquad T \mapsto \det(T)$$

folgende Eigenschaften:

[46] Warum nicht?

(1) Es gilt
$$\det(S \circ T) = \det(S) \cdot \det(T) \qquad \forall\, S, T \in \mathrm{Hom}_K(V, V).$$
(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 Uebung. □

Beispiel 8.6.5. Es sei V der Vektorraum aller Polynome ueber $\mathbb{Q}$ vom Grad ≤ 2 und $D : V \to V$ 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 due 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. □

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 $\partial(r) < 1$ (d.h. $r(x) \in K - \{0\}$), so dass
$$f(x) = g(x)q(x) + r(x). \tag{44}$$
Durch Auswertung von Gleichung (44) an $x = 0$ sehen wir, dass $r(0) = 0$ und daher $r = 0$. □

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 Nullstellen*[47] *in* $\mathbb{C}$*, das heisst, es gibt* $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$*, nicht notwendigerweise verschieden, so dass*
$$f(x) = (x - \lambda_1) \cdots (x - \lambda_n).$$

[47] Ein 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 Koerper in einem algebraisch abgeschlossenen Koerper 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 $n \geq 0$, fuer die es ein $q(x) \in K[x]$ gibt, so dass*
$$f(x) = (x - \lambda)^n \cdot q(x) \qquad \text{und} \qquad q(\lambda) \neq 0.$$
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 Koerper 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 eindeutig[48] in ein Produkt von unzerlegbaren Polynomen schreiben laesst. Unzerlgebare 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. Fuer Polynome vom Grad ≤ 4 gibt es explizite Formeln fuer die Nullstellen, aber bei Polynomen vom hoeren 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 loesbar ist oder nicht.

10. Eigenwerte und Eigenvektoren

10.1. Definitionen und erste Eigenschaften.

Definition 10.1.1. *Es sei V ein endlich-dimensionaler Vektorraum ueber K und $T : V \to V$ eine lineare Abbildung.*
(1) Ein Element $\lambda \in K$ ist ein Eigenwert[49] *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 $T_A : K^n \to K^n$.

Bemerkung 10.1.2. *Wenn v ein Eigenvektor von T ist, dann ist auch αv ein Eigenvektor (mit dem gleichen Eigenwert) fuer 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
$$A. \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{und} \qquad A \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$
Sind das alle Eigenwerte von A?

[48] abgesehen von Multiplikation mit Elementen in $K - \{0\}$.
[49] engl: 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 $B.v = \lambda v$, das heisst

$$\begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \quad \Leftrightarrow \quad \left[\begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wir erhalten daraus das lineare Gleichungssyste,

$$\begin{aligned} (1-\lambda)x - 2y &= 0 \\ x + (4-\lambda)y &= 0. \end{aligned}$$

Indem wir die zweite Gleichung mit $-(1-\lambda)$ multiplizieren und zur ersten Gleichung addieren, erhalten wir

$$\left(-(1-\lambda)(4-\lambda) - 2\right) y = 0 \quad \to \quad y = 0 \vee \lambda^2 - 5\lambda + 6 = 0.$$

- 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:

$$\begin{aligned} \lambda = 2 \quad &\Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \alpha \cdot \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \\ \lambda = 3 \quad &\Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \beta \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \end{aligned}$$

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 $T : V \to V$ 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.

$$\begin{aligned} \lambda \in K \text{ ist einEigenvalue von } T \quad &\Leftrightarrow \quad \exists v \in V,\, v \neq 0_V \text{ so dass } T.v = \lambda v \\ &\Leftrightarrow \quad \exists v \in V,\, v \neq 0_V \text{ so dass } (T - \lambda \mathbf{1}_V).v = 0_V \\ &\Leftrightarrow \quad v \in \ker(T - \lambda \mathbf{1}_V). \end{aligned}$$

□

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) ⇔ (ii) ist Satz 10.1.4, und (iii) ⇔ (iv) ist Satz 8.6.4 (ii). Die Aequivalenz (ii) ⇔ (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 $S : V \to V$ 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 characteristische 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 $T : V \to V$ definieren wir das characteristische 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

$$\begin{aligned} \det([T]_{\mathcal{B}}^{\mathcal{B}} - x \cdot \mathbf{1}_n) &= \det \left(A^{-1}[T]_{\mathcal{C}}^{\mathcal{C}} A - x \cdot \mathbf{1}_n \right) \\ &= \det \left[A^{-1} \left([T]_{\mathcal{C}}^{\mathcal{C}} A - x A \mathbf{1}_n \cdot A^{-1} \right) A \right] \\ &= \det(A^{-1}) \cdot \det([T]_{\mathcal{C}}^{\mathcal{C}} - x\mathbf{1}_n) \cdot \det(A) \\ &= \det([T]_{\mathcal{C}}^{\mathcal{C}} - x\mathbf{1}_n), \end{aligned}$$

das heisst, das characteristische Polynom von T ist wohl-definiert.

Wir koennen Korollar 10.1.6 folgendermassen zusammenfassen:

Theorem 10.2.5. *Es sei $T : V \to V$ linear. Dann ist $\lambda \in K$ genau dann ein Eigenwert von T, wenn $\chi_T(\lambda) = 0$. Mit anderen Worten:*

$$\{\textit{Eigenwerte von } T\} \quad = \quad \{\lambda \in K : \chi_T(\lambda) = 0\}.$$

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 $a_{11}, \ldots, a_{nn}$.

(2) Es sei $M \in M_{n\times n}(K)$ von der Form

$$M = \begin{pmatrix} A & B \\ 0_{s\times r} & C \end{pmatrix} \qquad \text{fuer } A \in M_{r\times r}(K),\ B \in M_{r\times s}(K) \text{ und } C \in M_{s\times s}(K).$$

Dann gilt

$$\chi_M(x) = \chi_A(x)\cdot\chi_C(x).$$

Proof. Uebung. □

Wir koennen auch einiges ueber die Koeffizienten des charakteristischen Polynoms sagen: wir erinnern uns an die Spurabbildung (Beispiel 4.1.3 (v))

$$\mathrm{Tr}: M_{n\times n}(K) \to K, \qquad A = (a_{ij}) \mapsto a_{11} + \cdots + a_{nn}.$$

Lemma 10.2.7. *Es gilt*

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA) \qquad \forall\, A, B \in M_{n\times n}(K).$$

Proof. Explizite Rechnung. □

Definition 10.2.8. *Es sei $T : V \to V$ linear. Wir definiere die Spur von T as*

$$\mathrm{Tr}(T) = \mathrm{Tr}[T]_{\mathcal{B}}^{\mathcal{B}}$$

fuer eine beliebige Basis $\mathcal{B}$ von V.

Satz 10.2.9. $\mathrm{Tr}(T)$ *ist wohl-definiert.*

Proof. Es seien $\mathcal{B}$ und $\mathcal{C}$ verschiedene Basen von V, und es sei $C = [\mathrm{id}_V]_{\mathcal{B}}^{\mathcal{C}}$, 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

$$\mathrm{Tr}\left([T]_{\mathcal{B}}^{\mathcal{B}}\right) = \mathrm{Tr}\left(C^{-1}\cdot[T]_{\mathcal{C}}^{\mathcal{C}}\cdot C\right),$$

d.h. $\mathrm{Tr}(T)$ ist wohldefiniert. □

Satz 10.2.10. *Es sei $A \in M_{n\times n}$. Dann gilt*

$$\chi_A(x) = (-1)^n x^n + (-1)^{n-1}\,\mathrm{Tr}(A)x^{n-1} + \cdots + \det(A).$$

Proof. Schreibe $A = (a_{ij})_{1\le i,j\le n}$. Dann gilt

$$\chi_A(x) = \det\begin{pmatrix} a_{11}-x & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22}-x & a_{23} & \ldots & a_{2n} \\ a_{31} & a_{32} & a_{33}-x & \ldots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nn}-x \end{pmatrix}$$

Wir benutzen nun (39), 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 $\{1, 2, \ldots, n\}$ 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 $\le n-2$. Daher gilt

$$\chi_A(x) = (-1)^n x^n + (-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})x^{n-1} + g(x).$$

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 dafuer erhalten, dass die Spur einer linearen Abbildung wohldefiniert ist: wir wissen von 10.2.4, dass $\chi_T(x)$ unabhaengig von der Wahl einer Basis von V ist; insbesondere gilt dies fuer den Koeffizienten von x^{n-1}. Nun wissen wir aber von Satz 10.2.10, dass dieser Koeffizient gegeben ist durch*

$$(-1)^{n-1} \cdot \mathrm{Tr}([T]_{\mathcal{B}}^{\mathcal{B}})$$

wobei $\mathcal{B}$ eine beliebige Basis von V ist; mit anderen Worten, der Wert $\mathrm{Tr}([T]_{\mathcal{B}}^{\mathcal{B}})$ ist unabhaengig von der Wahl der Basis. Dieses gibt einen neuen (und schoeneren) Beweis von Satz 10.2.9.

Korollar 10.2.12. *Es sei V an n-dimensionaler Vektorraum und $T : V \to V$ linear. Dann hat T hoechstens n Eigenwerte.*

Proof. Klar, da $\chi_T(x)$ ein Polynom vom Grad n ist und daher hoechstens n Nullstellen haben kann. □

Bemerkung 10.2.13. *Wenn λ ein Eigenwert von T ist, dann kann es sein, dass λ eine Nullstelle hoeherer Ordnung von $\chi_T(x)$ ist (e.g. fuer die Identitaetsabbildung* id_n *fuer $n > 1$).*

10.3. **Diagonalisierung.** Es sei $T : V \to V$ eine lineare Abbildung. Wir wollen nun die Idee von Eigenvektoren mit unserer Theorie des Basiswechsels zu verknuepfen: 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 $\varphi = \frac{1+\sqrt{5}}{2}$ and $\psi = \frac{1-\sqrt{5}}{2}$ die Eigenwerte der Abbildung S sind, mit dazu gehoerigen Eigenfolgen $\mathcal{F}_{1,\varphi}$ und $\mathcal{F}_{1,\psi}$. Eine einfache Rechnung zeigt, dass die beiden Eigenfolgen linear unabhaengig und daher einen Basis des Vektorraums aller Fibonacci-Folgen sind. Die Darstellungsmatrix von S bezueglich dieser Basis ist $\begin{pmatrix} \varphi & 0 \\ 0 & \psi \end{pmatrix}$: sie ist diagonal.

Satz 10.3.2. *Es sei $T : V \to V$ 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 unabhaengig.*

Proof. Nimm an, dass die Vektoren linear abhaengig sind. Es sei $1 \leq k \leq m$ der kleinste Index, so dass die Menge $\{v_1, \dots, v_k\}$ linear abhaengig ist. Dann gibt es Skalare $\alpha_1, \dots, \alpha_k \in K$, alle $\neq 0$, so dass

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0_V.$$

Wende die Abbildung $T - \lambda_k \mathbf{1}_V$ auf diese Gleichung an. Dann erhalten wir

$$\begin{aligned}(T - \lambda_k \mathbf{1}_V)(\alpha_1 v_1 + \cdots + \alpha_k v_k) &= \alpha_1 (T - \lambda_k \mathbf{1}_V) v_1 + \cdots + \alpha_k (T - \lambda_k \mathbf{1}_V) v_k \\ &= \alpha_1(\lambda_1 - \lambda_k) v_1 + \cdots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k) v_{k-1} + \alpha_k(\lambda_k - \lambda_k) v_k \\ &= \alpha_1(\lambda_1 - \lambda_k) v_1 + \cdots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k) v_{k-1} \\ &= 0_V.\end{aligned}$$

Da $\lambda_i \neq \lambda_j$ forall $i \neq j$, gilt $\alpha_i(\lambda_i - \lambda_k) \neq 0$ for all $1 \leq i < k$, mit anderen Worten, die Vektoren $v_1, \dots, v_{k-1}$ sind linear abhaengig. Dies ist ein Widerspruch zu der Wahl von k. □

Korollar 10.3.3. *Es sei V ein n-dimensionaler K-Vektorraum und $T : V \to V$ 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 unabhaengig aufgrund von Satz 10.3.2 und daher eine Basis (Satz 3.3.23). □

Definition 10.3.4. *Eine lineare Abbildung $T : V \to V$ ist* diagonalisierbar*, wenn V eine Basis besitzt, die aus Eigenvektoren von T besteht; in diesem Fall ist die Abbildungsmatrix von T bezueglich 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.*

[50] Wir 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

$$C = [\mathrm{id}]^{\mathcal{B}}_{\mathcal{E}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

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)*

$$\chi_A(x) = (x - \lambda_1) \cdots (c - \lambda_n).$$

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 characteristische 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 Matrix[51]. Wir berechnen den dazughoerigen Eigenvektor (oder sind es mehrere?):

$$\begin{aligned} & D. \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \\ \Rightarrow \quad & x + y = x \quad \text{und} \quad y = y, \end{aligned}$$

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?

[51] Beachte, dass $\chi_D(x) = \chi_{\mathbf{1}_2}(x)$!

Beispiel 10.3.9. Beispiel 10.3.8 (iv) laesst sich folgendermassen verallgemeinern: es sei $n \geq 1$ und $\lambda \in K$. Definiere die Jordan'sche Blockmatrix $J_n(\lambda) \in M_{n\times n}(K)$,

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}.$$

Dann gilt

$$\chi_{J_n(\lambda)}(x) = (\lambda - x)^n,$$

das heisst, λ ist der einzige Eigenwert. Andererseits is 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 $T : V \to V$ 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

$$E_\lambda = \ker(T - \lambda \cdot \mathrm{id}_V),$$

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. □

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 Unterraeumen eines Vektorraums:

Definition 10.4.4. *Es sei V ein Vektorraum, und es seien $U_1, \dots, U_k \leq V$. Es sei $W = U_1 + \cdots + U_k$. Dann ist W die direkte Summe von $U_1, \dots, U_k$ wenn jedes Element $w \in W$* in genau einer Weise *als Summe*

$$w = u_1 + \cdots + u_k \qquad \text{mit } u_i \in U_i \, \forall 1 \leq i \leq k$$

geschrieben werden kann. In diesem Fall schreiben wir

$$W = U_1 \oplus \cdots \oplus U_k.$$

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 laesst 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*

$$0_V = u_1 + \cdots + u_k \qquad \text{mid } u_i \in U_i \, \forall i$$

nur die Loesung

$$u_1 = u_2 = \cdots = u_k = 0_V$$

hat.

Proof. Uebung. □

Beispiele 10.4.7.

(1) $\mathbb{R}^3$ is 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 $W = U_1 \oplus \cdots \oplus U_k$ ist, dann gilt*

$$\dim(W) = \dim(U_1) + \cdots + \dim(U_k).$$

Weiterhin gilt: wenn $\mathcal{B}_i$ eine Basis von U_i ist $\forall\, 1 \leq i \leq k$, dann ist

$$\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$$

einen Basis von W.

Zurueck zu den Eigenraeumen:

Satz 10.4.9. *Es sei $T : V \to V$ eine lineare Abbildung, und es seien $\lambda_1, \dots, \lambda_k$ die Eigenwerte von T. Dann ist der Unterraum*

$$W = E_{\lambda_1} + \cdots + E_{\lambda_k}$$

die direkte Summe $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$.

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, fuer die $v_{i_j} \neq 0_V$. Da $\lambda_i \neq \lambda_j$ fuer $i \neq j$, sind $v_{i_1}, \dots, v_{i_r}$ linear unabhaengig aufgrund von Satz 10.3.2; daher gilt[52]

$$v_{i_1} + \cdots + v_{i_r} \neq 0_V.$$

□

Korollar 10.4.10. *Es sei $T : V \to V$ eine lineare Abbildung, und es seien $\lambda_1, \dots, \lambda_k$ die Eigenwerte von T. Dann ist T genau dann diagonalisierbar, wenn*

$$\dim(V) = \dim E_{\lambda_1} + \cdots + \dim E_{\lambda_k}. \tag{45}$$

[52]Sie koennen alternativ auch mit Hilfe von Basen von Eigenraeumen 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 $a_1, \ldots, a_k \geq 1$, so dass*

$$\chi_T(x) = (\lambda_1 - x)^{a_1} \cdots (\lambda_k - x)^{a_k},$$

und insbesondere

(46) $$n = a_1 + \cdots + a_n.$$

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 $T : V \to V$ 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

$$g_\lambda = 1 \qquad \text{und} \qquad a_\lambda = n.$$

Faellt Ihnen was auf?

Satz 10.5.4. *Es sei $T : V \to V$ 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 $v_1, \ldots, v_k$ des Eigenraums E_λ. Dann koennen wir $\{v_1, \ldots, v_k\}$ zu einer Basis $\mathcal{B} = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ 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

$$\begin{aligned} \chi_T(x) &= \det\left([T]_\mathcal{B}^\mathcal{B} - x\mathbf{1}_n\right) \\ &= \det\left((\lambda - x) \cdot \mathbf{1}_k\right) \cdot \det\left(B - x\mathbf{1}_{n-k}\right) \\ &= (\lambda - x)^k \cdot \chi_B(x) \end{aligned}$$

mit $\chi_B(x) \in K[x]$. Daher gilt $k \leq a_\lambda$. ☐

Korollar 10.5.5. *Es sei $T : V \to V$ eine lineare Abbildung, und es seien $\lambda_1, \ldots, \lambda_k$ die Eigenwerte von T. Dann ist T genau dann diagonalisierbar, wenn*

$$g_{\lambda_i} = a_{\lambda_i} \qquad \forall\, 1 \leq i \leq n.$$

Proof. Wir wissen von Korollar 10.4.10, dass T genau dann diagonalisierbar, wenn Lecture 30

(47) $$\dim(V) = \dim E_{\lambda_1} + \cdots + \dim E_{\lambda_k}.$$

Von (46) erhalten wir, dass

$$g_{\lambda_1} + \cdots + g_{\lambda_k} \leq a_{\lambda_1} + \cdots + a_{\lambda_k} = \dim(V).$$

Da $g_{\lambda_i} \leq a_{\lambda_i}\ \forall i$, 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 $T : V \to V$ 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 $T : V \to V$ eine lineare Abbildung definieren wir T^k, $k \geq 1$, als die k-fache Verknuepfung von T mit sich selbst:*

$$T^k = T \circ \cdots \circ T.$$

Weiterhin sei $T^0 = \mathrm{id}_V$.
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 $T \in \mathrm{End}_K(V)$ und*

$$g(x) = a_d x^d + a_{d-1} x^{d-1} + \cdots + a_1 x + a_0 \in K[x].$$

Dann definieren wir

$$g(T) = a_d T^d + a_{d-1} T^{d-1} + \cdots + a_1 T + a_0 \,\mathrm{id}_V \in \mathrm{End}_K(V).$$

Entsprechend definieren wir $g(A)$ fuer $A \in M_{n\times n}(K)$ als

$$g(A) = a_d A^d + a_{d-1} A^{d-1} + \cdots + a_1 A + a_0 \mathbf{1}_n \in M_{n\times n}(K).$$

Bemerkung 11.1.3. *Die Auswertung eines Polynoms an einer Matrix oder eines Endomorphismus respektiert die Korrepsondenz 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 $T_A : K^n \to K^n$ 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 $T \in \mathrm{End}_K(V)$[54] gibt es $g(x) \in K[x]$, $g(x) \neq 0$, so dass $g(T) = 0_V$.*

Proof. $\mathrm{End}_K(V) \cong M_{n\times n}(K)$ 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

$$a_{n^2} T^{n^2} + \dots + a_1 T + a_0 \,\mathrm{id}_V = 0_V.$$

□

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 $T : V \to V$ 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 Abbdilungsmatrix 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 $T : V \to V$ linear. Das* minimale Polynom *von T ist das monische*[55] *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 $g(x) = m(x) - m'(x)$ vmom Grad $< d$, und es gilt $g(T) = 0$. Das ist ein Widerspruch zur Definition des minimalen Polynoms. □

[53] Eine 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*.

[54] und entsprechend fuer alle $A \in M_{n\times n}(K)$

[55] d.h. Leitkoeffizient ist gleich 1

Satz 11.1.10. *Es sei $T : V \to V$ linear, und es sei $g(x) \in K[x]$, $g \neq 0$, so dass $g(T) = 0_V$. Dann gilt $m_T(x) \,|\, g(x)$.*

Proof. Mithilfe von Polynomdivison finden wir Polynom $q(x), r(x) \in K[x]$, mit $r(x) = 0$ oder $\partial(r) < \partial(m)$, so dass

$$g(x) = q(x) \cdot m(x) + r(x). \tag{48}$$

- Wenn $r(x) = 0$, dann gilt $g(x) = q(x)m(x)$ und daher $m(x)|g(x)$.
- Wenn $r(x) \neq 0$, dann werten wir (48) an T aus und erhalten $r(T) = 0_V$. Doch dies gibt einen Widerspruch zu der Minimalitaet 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 $m_X(x) = x - \lambda$.
 (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
$$m_A(A) = (\lambda \cdot \mathbf{I}_2 - A)(\mu \cdot \mathbf{I} - A) = 0_{2\times 2},$$
 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 $T : V \to V$ diagonalisierbar ist, mit Eigenwerten $\lambda_1, \ldots, \lambda_k$, $\lambda_i \neq \lambda_j$ 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 beidem Matrizen haben also unterschiedlichen minimale Polynome aber das gleiche characteristische 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*
$$f(x) = (x - \lambda_1)^{a_1} \cdots (x - \lambda_m)^{a_n},$$
$$g(x) = (x - \lambda_1)^{b_1} \cdots (x - \lambda_n)^{b_n}$$
wobei $a_i, b_i \geq 0$ fuer all 1. *Das* kleinste gemeinsame Vielfache von $f(x)$ und $g(x)$ *ist definiert als*
$$\operatorname{lcm}(f, g) = (x - \lambda_1)^{\max\{a_1, b_1\}} \cdots (x - \lambda_n)^{\max\{a_n, b_n\}}.$$

Beispiel 11.1.13. Es seien
$$f(x) = x(x-3)^4(x+1)^2, \qquad g(x) = (x-3)(x+1)^3(x-2)^2.$$
Dann ist
$$\operatorname{lcm}(f, g) = x(x-3)^4(x+1)^3(x-2)^2.$$

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 $m_B(x)$:
$$m_C(x) = \operatorname{lcm}(m_A(x), m_B(x)).$$

Proof. Wir beachten zunaechst, 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 $f(x) = \mathrm{lcm}(m_A(x), m_B(x))$. 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 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. □

Korollar 11.1.15. *Es sei K algebraisch abgeschlossen und $C \in M_{n\times n}(K)$. Nimm an, dass*

$$C = \begin{pmatrix} A_1 & 0 & 0 & 0 & \dots \\ 0 & A_2 & 0 & 0 & \dots \\ & & \dots & & \\ & & \dots & & \\ 0 & 0 & 0 & 0 & A_k \end{pmatrix}$$

mit $A_i \in M_{n_i\times n_i}(K)$. Dann gilt

$$m_C(x) = \mathrm{lcm}\left\{m_{A_1}(x), \dots, m_{A_k}(x)\right\}.$$

Proof. Induktion. □

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 $T \in \mathrm{End}_K(V)$. Dann gilt $\chi_T(T) = 0_{\mathrm{End}_K(V)}$.*

Korollar 11.2.2. *$\forall T \in \mathrm{End}_K(V)$ gilt $m_T(x)|\chi_T(x)$.*

Proof. Folgt unmittelbar von Satz 11.1.10 und Theorem 11.2.1. □

Lemma 11.2.3. *Es sei $f(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0$, und es sei $A \in M_{k\times k}(K)$ gegeben durch*[56]

$$A = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 & -a_0 \\ 1 & 0 & 0 & \dots & 0 & 0 & -a_1 \\ 0 & 1 & 0 & \dots & 0 & 0 & -a_2 \\ 0 & 0 & 1 & \dots & 0 & 0 & -a_3 \\ & \dots & & \dots & & & \\ 0 & 0 & 0 & \dots & 1 & 0 & -a_{k-2} \\ 0 & 0 & 0 & \dots & 0 & 1 & -a_{k-1} \end{pmatrix}.$$

Dann ist $\chi_A(x) = (-1)^k f(x)$.

Proof. Uebung. □

Lecture 33

Lemma 11.2.4. *Es sei $T : V \to V$ 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

$$\begin{aligned} \chi_T(x) &= \det\left([T]_{\mathcal{B}}^{\mathcal{B}} - x\mathbf{1}_n\right) \\ &= \det \begin{pmatrix} A - x\mathbf{1}_k & B \\ 0_{(n-k)\times k} & C - x\mathbf{1}_{n-1} \end{pmatrix} \\ &= \det(A - x\mathbf{1}_k) \cdot \det(C - x\mathbf{1}_{n-k}) \\ &= \chi_{T'}(x) \cdot \det(C - x\mathbf{1}_{n-k}). \end{aligned}$$

Daraus folgt, dass $\chi_{T'}(x) \mid \chi_T(x)$. □

Satz 11.2.5. *Es sei $T : V \to V$ eine lineare Abbildung, und es sei $w \in V$. Es sei $W = \langle w, Tw, T^2 w, \dots\rangle$. 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

$$T^k w + a_{k-1} \cdot T^{k-1} w + \cdots + a_1 \cdot Tw + a_0 \cdot w = 0_V. \tag{49}$$

Dann ist $[T']_{\mathcal{B}}^{\mathcal{B}}$ die Begleitmatrix des Polynoms

$$f(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0,$$

und es folgt von Lemma 11.2.3, dass $\chi_{T'}(x) = f(x)$. Daher ist $\chi_{T'}(T).w$ gleich (49), 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

$$\chi_T(T).w = 0_V \qquad \forall w \in V.$$

□

[56]Wir 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*

$$(A - x\mathbf{1}_n)\mathrm{adj}(A - x\mathbf{1}_n) = \chi_A(x) \cdot \mathbf{1}_n. \tag{50}$$

Warum koennen wir A nicht direkt in diese Gleichung einsetzen und erhalten $\chi_A(A) = 0_{n\times n}$? Erinnern Sie sich, dass $\mathrm{adj}(A - x\mathbf{1}_n)$ *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* $\mathrm{adj}(A - x\mathbf{1}_n)$ *schreiben als*

$$\mathrm{adj}(A - x\mathbf{1}_n) = B_{n-1}x^{n-1} + \dots + B_1 x + B_0 \tag{51}$$

mit $B_0, B_1, \dots, B_{k-1} \in M_{n\times n}(K)$. Koennten wir hier nicht einfach A einsetzen? Die Antwort ist NEIN, weil wir nicht wissen, ob A mit den B_i kommutiert! Es gilt ja

$$B_{n-1}x^{n-1} + \dots + B_1 x + B_0 = x^{n-1} \cdot B_{n-1} + \dots + x \cdot B_1 + B_0,$$

aber wir wissen nicht, ob

$$B_{n-1}A^{n-1} + \dots + B_1 A + B_0 = A^{n-1} \cdot B_{n-1} + \dots + A \cdot B_1 + B_0!$$

Das Einsetzen von A in die Gleichung ist also a priori[57] *nicht wohldefiniert.*

Beweis von Theorem 11.2.1: Wir schreiben $\mathrm{adj}(A - x\mathbf{1}_n) = (p_{ij}(x))_{1\leq i,j\leq n}$, 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 $\mathrm{adj}(A - x\mathbf{1}_n)$ schreiben als Lecture 34

$$\mathrm{adj}(A - x\mathbf{1}_n) = B_{n-1}x^{n-1} + \dots + B_1 x + B_0$$

mit $B_0, B_1, \dots, B_{k-1} \in M_{n\times n}(K)$. Indem wir dies in (50) einsetzen, erhalten wir

$$(A - x\mathbf{1}_n)(B_{n-1}x^{n-1} + \dots + B_1 x + B_0) = \chi_A(x) \cdot \mathbf{1}_n. \tag{52}$$

Wenn wir $\chi_A(x)$ schreiben als

$$\chi_A(x) = (-1)^n(x^n + \dots + a_1 x + a_0),$$

dann koenne wir in (52) $\forall 0 \leq i \leq n$ die Koeffizienten von x^i vergleichen und erhalten folgendes System von Matrixgleichungen

$$\begin{aligned} AB_0 &= (-1)^n a_0 \mathbf{1}_n, \\ -B_0 + AB_1 &= (-1)^n a_1 \mathbf{1}_n, \\ &\vdots \\ -B_{n-2} + AB_{n-1} &= (-1)^n a_{n-1} \mathbf{1}_n, \\ -B_{n-1} &= (-1)^n \mathbf{1}_n. \end{aligned}$$

[57] siehe Bemerkung 11.3.5

Indem wir diese Gleichungen von links mit $\mathbf{1}_n, A, A^2, \ldots, A^n$ 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 $0_{n\times n}$, waehrend 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 laessst: 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

$$(A - x\mathbf{1}_n)\mathrm{adj}(A - x\mathbf{1}_n) = \mathrm{adj}(A - x\mathbf{1}_n)(A - x\mathbf{1}_n),$$

erhalten wir noch ein zweites System von Matrrixgleichungen

$$\begin{aligned} AB_0 &= (-1)^n a_0 \mathbf{1}_n, \\ -B_0 + B_1 A &= (-1)^n a_1 \mathbf{1}_n, \\ &\vdots \\ -B_{n-2} + B_{n-1} A &= (-1)^n a_{n-1} \mathbf{1}_n, \\ -B_{n-1} &= (-1)^n \mathbf{1}_n. \end{aligned}$$

Wir folgern daraus, dass $AB_i = B_i A$ fuer alle $1 \le i < n$, d.h. das Einsetzen von A in (51) *ist letztendlich* doch *wohldefiniert!*

12. Die Jordansche Normalenform einer Matrix

Annahme: Fuer den Rest dieses Kapitels sei K algebraisch abgeschlossen.

12.1. **Definition und Theorem.** Wir erinnern uns an die Definition der Jordanschen Blocks: es sei $n \ge 1$ und $\lambda \in K$. Dann ist der Jordanblock der Laenge n und Eigenwert λ die Matrix $J_n(\lambda) \in M_{n\times n}(K)$,

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & 0 & \ldots & 0 \\ 0 & 0 & \lambda & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 1 \\ 0 & 0 & 0 & 0 & \ldots & \lambda \end{pmatrix}$$

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 $T : V \to V$ linear. Dann gibt es eine Basis $\mathcal{B}$ von V, so dass*

$$[T]^{\mathcal{B}}_{\mathcal{B}} = \begin{pmatrix} J_{n_1}(\alpha_1) & & & & \\ & J_{n_2}(\alpha_2) & & & \\ & & J_{n_3}(\alpha_3) & & \\ & & & \vdots & \\ & & & & J_{n_k}(\alpha_k) \end{pmatrix} \tag{53}$$

fuer $n_1, \ldots, n_k \geq 1$ so dass $n_1 + \cdots + n_k = n$ und $\alpha_1, \ldots, \alpha_k \in K$[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* (54) *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 verraet; 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
$$E_{\lambda_1} = \langle e_1, e_4\rangle \qquad E_{\lambda_2} = \langle e_6, e_8, e_9\rangle;$$
- 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 Unterraeume $U = \langle e_1, \ldots, e_\ell\rangle$ und $W = \langle e_{\ell+1}, \ldots, e_n\rangle$. 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 gilt[59]

$$E_\lambda(T_C) = E_\lambda(T_A) \oplus E_\lambda(T_B)$$

[58] nicht notwendigerweise verschieden!

[59] Hier schreiben wir fuer eine lineare Abbildung $T : V \to V$ 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. Uebung. $\square$

Theorem 12.2.3. *Es sei* $T : V \to V$ *linear. Nimm an, dass es eine Basis* $\mathcal{B}$ *von* V *gibt, so dass*

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} J_{n_1}(\alpha_1) & & & & \\ & J_{n_2}(\alpha_2) & & & \\ & & J_{n_3}(\alpha_3) & & \\ & & \vdots & & \\ & & & & J_{n_k}(\alpha_k) \end{pmatrix} \tag{54}$$

Es sei λ *ein Eigenwert von* T*. Dann ist die geometrische Vielfachheit von* λ *die Anzahl der Jordanbloecke mit Eigenwert* λ*, d.h.*

$$g_\lambda = \#\{i : 1 \leq i \leq k, \, \alpha_i = \lambda\}.$$

Das minimale Polynom von T *is gegeben durch*

$$m_T(x) = \prod_{\lambda \text{ Eigenwert}} (x - \lambda)^{s(\lambda)},$$

wobei $s(\lambda)$ *die Laenge des groessten 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

$$\mathcal{B} = \left\{ b_1^{(1)}, \dots, b_{n_1}^{(1)}, b_1^{(2)}, \dots, b_{n_2}^{(2)}, \dots, b_1^{(k)}, \dots, b_{n_k}^{(k)} \right\},$$

und es sei $W_i = \left\langle b_1^{(i)}, \dots, b_{n_i}^{(i)} \right\rangle$, d.h. W_i ist der Unterraum von V mit Basis

$$\mathcal{B}^{(i)} = \{b_1^{(i)}, \dots, b_{n_i}^{(i)}\}$$

Dann gilt

$$V = W_1 \oplus \cdots \oplus W_k,$$

und die Abbildungsmatrix von $T|_{W_i}$ bezueglich der Basis $\mathcal{B}^{(i)}$ ist gleich dem Jordanblock $J_{n_i}(\alpha_i)$. Dann folgt von Lemma 12.2.2, dass

$$g_\lambda(T) = g_\lambda(T|_{W_1}) + \cdots + g_\lambda(T|_{W_k}).$$

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

$$m_T(x) = \operatorname{lcm} \left\{ m_{J_{n_1}}(\alpha_1), \dots, m_{J_{n_k}}(\alpha_k) \right\}.$$

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

$$m_T(x) = \prod_{\lambda \, \text{Eigenwert}} (x - \lambda)^{s(\lambda)}.$$

$\square$

12.3. **Verallgemeinerte Eigenraeume.**

Es sei V ein n-dimensionaler K-Vektorraum.

Definition 12.3.1. *Es sei $T: V \to V$ 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 $S: V \to V$ 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*

$$v, Sv, \ldots, S^{k-1}v$$

linear unabhaengig.

Proof. Nimm an, dass es Skalare $a_0, \ldots, a_{k-1}$ gibt, so dass

$$a_0 v + a_1 Sv + \cdots + a_{k-1}S^{k-1}v = 0_V. \tag{55}$$

Indem wir S^{k-1} auf die Gleichung (55) anwenden, erhalten wir $a_0 S^{k-1}v = 0_V$ und daher $a_0 = 0$. Dann wenden wir S^{k-2} auf (55) an und erhalten $a_1 S^{k-1}v = 0_V$ und daher $a_1 = 0$. Ueber Induktion erhalten wir daher $a_i = 0$ $\forall i$. □

Lecture 35

Lemma 12.3.3. *Es gilt $\tilde{E}_\lambda = \ker(T - \lambda\mathbf{1}_V)^n$.*

Proof. Es sei $S = T - \lambda\mathbf{1}_n$. 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 $m, Sv, \ldots, S^{m-1}v$ linear unabhaengig. Daher gilt $m \leq n$. □

Definition 12.3.4. *Es sei $v \in \tilde{E}_\lambda$, $v \neq 0_V$, und es sei $k \geq 1$ minimal, so dass $(T - \lambda\mathbf{1}_n)v = 0_V$. Dann ist die Menge*

$$\{v, (T - \lambda\mathbf{1}_V)v, \ldots, (T - \lambda\mathbf{1}_V)^{k-1}v\}$$

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 $T_A : K^n \to K^n$ 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 $T_A : K^9 \to K^9$. Dann ist

$$\tilde{E}_\lambda = \langle e_1, \ldots, e_6 \rangle \qquad \text{und} \qquad \tilde{E}_\mu = \langle e_7, e_8, e_9 \rangle.$$

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

$$\tilde{E}_{\lambda_1} = \langle e_1, \ldots, e_5 \rangle \qquad \text{und} \qquad \tilde{E}_{\lambda_2} = \langle e_6, \ldots, e_9 \rangle.$$

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

$$T\left(\tilde{E}_\lambda\right) \subseteq \tilde{E}_\lambda.$$

Nimm an, dass $(T - \lambda\mathbf{1}_V)^n v = 0_V$. Dann gilt

$$(T - \lambda\mathbf{1}_V)^n Tv = T(T - \lambda\mathbf{1}_V)^n v = 0_V,$$

und daher $Tv \in \tilde{E}_\lambda$. □

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 $(T-\lambda\,\mathrm{id}_V)^k v = 0$, und es sei $w = (T-\lambda\,\mathrm{id}_V)^{k-1}v$. Dann gilt

$$(T-\lambda\,\mathrm{id}_V)w = (T-\lambda\,\mathrm{id}_V)^k v = 0_V = (\mu-\lambda)w$$

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 $\lambda_1, \ldots, \lambda_k$ voneinander verschiedene Eigenwerte von T. Dann ist der Unterraum $W = \tilde{E}_{\lambda_1} + \cdots + \tilde{E}_{\lambda_k}$ die direkte Summe*

$$W = \tilde{E}_{\lambda_1} \oplus \cdots \oplus \tilde{E}_{\lambda_k}.$$

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 $T : V \to V$ linear, und es sei λ ein Eigenwert von T. Es sei $g(x) = (x-\lambda)^n$. Dann ist $\mathrm{im}(g(T))$ ein T-invarianter Unterraum von V, und es gilt

$$V = \tilde{E}_\lambda \oplus \mathrm{im}(g(T)).$$

(2) Es sei μ ein von λ verschiedener Eigenwert von T. Dann gilt

$$\tilde{E}_\mu \subseteq \mathrm{im}(g(T)).$$

Proof. (1) Die Invarianz von $\mathrm{im}(g(T))$ ist eine Uebung. Von Theorem 4.2.9 wissen wir, dass

$$\dim V = \dim\ker(g(T)) + \dim\mathrm{im}(g(T));$$

es ist daher ausreichend zu zeigen, dass $\ker(g(T)) \cap \mathrm{im}(g(T)) = \{0\}$. Es sei $w \in \mathrm{im}(g(T))$, d.h. es gibt $v \in V$ so dass $g(T)v = w$. Wenn $w \in \tilde{E}_\lambda$, dann gilt

$$(T-\lambda\mathbf{1}_V)^n w = (T-\lambda\mathbf{1}_V)^{2n} v = 0_V,$$

das heisst $v \in \tilde{E}_\lambda$. Nun wissen wir aber von Lemma 12.3.3, dass $\tilde{E}_\lambda = \ker(T-\lambda\mathbf{1}_n)^n$, und daher gilt

$$w = (T-\lambda\mathbf{1}_V)^n v = 0_V.$$

(2) Es sei $S = (T-\lambda\mathbf{1}_V)^n$. Eine einfache Rechnung zeigt, dass $\tilde{E}_\mu$ ein S-invarianter Unterraum von V ist. Es sei $S' = S|_{\tilde{E}_\mu}$. Dann ist

$$\ker(S') = \tilde{E}_\lambda \cap \tilde{E}_\mu = \{0\} \qquad \text{(aufgrund von Lemma 12.3.11)},$$

das heisst, S' ist ein Isomorphismus. Insbesondere gilt $\tilde{E}_\mu = \mathrm{im}(S') \subseteq \mathrm{im}(g(T))$, quod erat demonstrandum. □

Theorem 12.3.15. *Es sei $T : V \to V$ linear. Dann gilt*

$$V = \bigoplus_{\eta \text{ Eigenwert}} \tilde{E}_\eta :$$

V ist die direkte Summer seiner verallgemeinerten Eigenraeume.[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 Lecture 36

$$V = \tilde{E}_\lambda \oplus \operatorname{im}\left((T - \lambda \mathbf{1}_V)^n\right),$$

und dass $U = \operatorname{im}\left((T - \lambda \mathbf{1}_V)^n\right)$ 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 $T : V \to V$ 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*

$$V = \ker\left((T - \lambda_1 \mathbf{1}_V)^{n_1}\right) \oplus \cdots \oplus \ker\left((T - \lambda_k \mathbf{1}_V)^{n_k}\right).$$

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 $S : V \to V$ nilpotent ist.

Theorem 12.4.1. *Es sei V ein n-dimensionaler Vektorraum ueber K und $S : V \to V$ nilpotent. Dann gibt es $k \geq 1$, $n_1, \dots, n_k \geq 1$, so dass $n_1 + \cdots + n_k = n$, und eine Basis $\mathcal{B}$ von V, die aus Jordanketten besteht, da.h. von der Form*

$$S^{n_1} u_1, \dots, S u_1, u_1, S^{n_2 - 1} u_2, \dots, u_2, \dots, S^{n_k - 1} u_k, \dots, u_k, \tag{56}$$

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* (56) *von V heisst Jordanbasis.*

Bemerkung 12.4.3. *Die Abbildungsmatrix von S bezueglich 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 $S(V) \subsetneq V$ (warum?). Wenn $S(V) = \{0_V\}$, dann ist S die Nullabbildung, und das Theorem haelt trivialerweise. Wir koennen daher annehmen, dass

$$\{0_V\} \subsetneq S(V) \subsetneq V.$$

[60]Dieses 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 $b_1 + \cdots + b_\ell = \dim S(V)$ und

$$\mathcal{B}' = \{S^{b_1-1}v_1, \dots, v_1, \dots, S^{b_\ell-1}v_\ell, \dots, v_\ell\} \tag{57}$$

eine Basis von $S(V)$ ist, mit $S^{b_i}v_i = 0_V \; \forall i$.

Fuer $1 \leq i \leq \ell$ waehle $u_i \in V$, 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 unabhaegig sind; wir koennen 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

$$\left\{S^{b_1}u_1, Su_1 \dots, u_1, S^{b_2}u_2, \dots, u_2, \dots, S^{b_\ell}u_\ell, \dots, u_\ell, w_1, \dots, w_m\right\}$$

sind eine Basis $\mathcal{B}$ von V.[61]

Beweis der Behauptung. Die lineare Unabhaengigkeit ist eine Uebung. Um zu zeigen, dass die Vektoren auch ein Erzeugendensystem sind, berechnen wir die Dimension von V. Wir wissen bereis, dass $\dim(\ker(S)) = \ell + m$ und dass

$$\operatorname{rk}(S) = \dim \operatorname{im}(S) = \dim S(V) = b_1 + \cdots + b_\ell.$$

Daher folgt von Theorem 4.2.9, dass

$$\begin{aligned}\dim V &= \dim \ker(S) + \operatorname{rk}(S) \\ &= (\ell + m) + (b_1 + \cdots + b_\ell) \\ &= m + (b_1 + 1) + \cdots + (b_\ell + 1).\end{aligned}$$

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 $T : V \to V$ eine lineare Abbildung und λ ein Eigenwert mit dazugehoerigem verallgemeinerten Eigenraum $U = \tilde{E}_\lambda$. Dann hat U eine Basis, in der die Abbildungsmatrix von der Beschraenkung $T|_U$ in Jordanscher Normalenform ist.*

Proof. Es sei $S = (T - \lambda)\mathbf{1})|_U$. Dann ist $S : U \to U$ 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 Abbildugnsmatrix der Indentitaet in jeder Basis die Identitaetsmatrix 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 wichtigesten Resultate der linearen Algebra.

[61] Wenn sie die Jordanketten anersrun anordnen, d.h. $u_1, \dots, S^{b_1}u_1$ etc., dann erhalten Sie Einsen unterhalb der Hauptdiagronalen anstatt darueber.

[62] Sie koennten auch mit einem anderen Unterraum $U \lneq V$ arbeiten, um den Induktionsschritt durchzufuehren. 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.*

Frage. Wir haben gesehen, dass die JNF einer Matrix eindeutig ist; eine Jordanbasis hingegen ist *nicht* eindeutig. (Z.B. ist fuer die Identitaetsabbildung jede Basis eine Jordanbasis.) Wie haengen zwei verschiedene Jordanbasen miteinander zusammen? Mit anderen Worten, es sei $A \in M_{n\times n}(K)$ eine Matrix in Jordanscher Normalenform. Fuer welche Matrizen $C \in \mathrm{GL}_n(K)$ ist $C^{-1}AC = A$?

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 beschraenkt ist. Daher gilt $n_1 = 3$ und $n_2 = 1$, und die Jordansche Normalenfom 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

$$J(B) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

We gibt daher eine Matrix $C \in \mathrm{GL}_4(\mathbb{C})$, 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 $C^{-1}BC = J(B)$.

Lecture 37

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 Jordanbloecke, 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 Jordanbloecke der Laengen 1 und 3, und $y = e_1 + e_2 + e_3 + e_4$ gehoert zu einer Jordankette der Laenge 3. Wir nrechnen nach:

$$w = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} \quad \text{ist eine Loesung der Gleichung } (D - \mathbf{1}_4)w = y,$$

$$u = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \quad \text{ist eine Loesung der Gleichung } (D - \mathbf{1}_4)u = w,$$

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 Jordanbloecke hat. Die moeglichen Laengen sind 2 und 3, oder 1 und 4. Eine einfache Rechung 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 aehnlich.

Beweis der Behauptung. Wir analysieren die verschiedenen Faelle, die auftreten koennen:

- Wenn
$$\chi(x) = (\lambda_1 - x)(\lambda_2 - x)(\lambda_3 - x)$$
fuer drei verschiedene Eigenwerten $\lambda_1, \lambda_2, \lambda_3$, 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 verschiedene 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 Jordanbloecke die Laenge 1, das heisst, A und B sind aehnlich zu
$$J = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$
Wenn $j = 2$, dann gehoert λ_1 zu einem Jordanblock der Lange 2, d.h. A und B sind aehnlich 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
$$m(x) = (\lambda - x)^j \qquad \text{fuer } j \in \{1, 2, 3\},$$
und es folgt von Theorem 12.2.3, dass
$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \quad \text{oder} \quad J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \quad \text{oder} \quad J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix},$$
abhaengig davon, ob $j = 1$ oder $j = 2$ oder $j = 3$.

Bemerkung 12.5.6. *Fuer (4×4) Matrizen ist die Verallgemeinerung des vorheriges Beispiels falsch: Die Matrizen*
$$A = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{pmatrix} \quad \textit{und} \quad B = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$
haben das gleiche charakteristische und minimale Polynom, aber sie sind nicht aehnlich.

13. Euklidische und Hermitesche Raeume

Lecture 38

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*
$$|| \sim || : V \to \mathbb{R}_{\geq 0}$$
mit folgenden Eigenschaften:
- $||u + v|| \leq ||u|| + ||v|| \ \forall \, u, v \in V$ *(Dreieceksungleichung);*
- $||\alpha v|| = |\alpha| \cdot ||v||$ *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 $|| \sim ||$ in Beispiel 13.1.2 von den Normen $|| \sim ||'$ und $|| \sim ||''$ 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 $|| \sim ||_p$ eine Norm; sie heisst die p-*Norm.

Beispiel 13.1.5. Es sei V der Vektorraum aller stetigen Funktionen $f : [0, 1] \to \mathbb{R}$. Definiere
$$|f|_{\max} = \max\{|f(x)| : x \in [0, 1]\}.$$
Dann ist $| \sim |_{\max}$ eine Norm: sie heisst die *Maximumsnorm.*

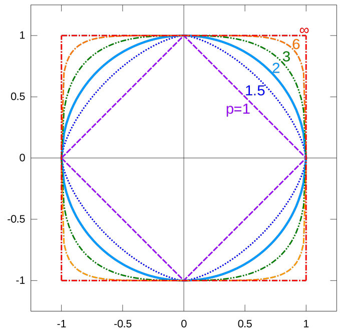
Definition 13.1.6.

(1) *Ein Vektor $v \in V$ ist ein Einheitsvektor wenn $||u|| = 1$.*
(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 und nun auf Vektorraeume ueber $\mathbb{R}$ oder $\mathbb{C}$ mit einer besonderen geometrischen Struktur: einem *inneren Produkt.* Wir erhalten

$$\begin{cases} \text{Euklidische Raeume ueber } \mathbb{R} \\ \text{Hermitesche Raeume ueber } \mathbb{C} \end{cases}$$

Definition 13.2.1. *Es sei V ein $\mathbb{R}$-Vektorraum. Ein* inneres Produkt *ist eine Funktion*

$$\langle\ ,\ \rangle : V \times V \to \mathbb{R}$$

mit folgenden Eigenschaften:

(1) *Linearitaet in der erste Variablen:*

$$\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle \quad \forall\, v_1, v_2, w \in V,$$
$$\langle \alpha v, w\rangle = \alpha\langle v, w\rangle \quad \forall\, \alpha \in \mathbb{R}, \forall\, v, w \in V.$$

(2) *Linearitaet in der zweiten Variablen:*[63]

$$\langle v, w_1 + w_2\rangle = \langle v, w_1\rangle + \langle v, w_2\rangle \quad \forall\, v, w_1, w_2 \in V,$$
$$\langle v, \alpha w\rangle = \alpha\langle v, w\rangle \quad \forall\, \alpha \in \mathbb{R}, \forall\, v, w \in V.$$

(3) *Symmetrie: $\langle v, w\rangle = \langle w, v\rangle\ \forall, v, w \in V$;*
(4) *Positivitaet: $\langle v, v\rangle > 0\ \forall\, v \in V,\ v \neq 0_V$.*

Das Paar $(V, \langle\, ,\rangle)$ ist ein Euklidischer Raum.

Beispiel 13.2.2. Es sei $V = \mathbb{R}^n$. Fuer $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ und $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ definiere

$$\langle u, v\rangle = \sum_{i=1}^{n} u_i v_i.$$

Dann ist $(V, \langle\ ,\ \rangle)$ ein Euklidischer Raum; das innere Proukt ist das Standardprodukt auf $\mathbb{R}^n$. Lecture 39

Beispiel 13.2.3. Es sei $V = \mathbb{R}^2$. Fuer $u_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$, $u_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ definiere

$$\langle u_1, u_2\rangle' = 2x_1x_2 - x_1y_2 - y_1x_2 + y_1y_2.$$

[63] Eine Abbildung $V \times V \to \mathbb{R}$, die linear in jeder der Variablen ist, heisst *bilinear*.

Dann ist $\langle\quad,\quad\rangle'$ ein inneres Produkt: Eigenschaften (1) - (3) sind klar; fuer Eigenschaft (4), beachte, dass fuer $u = \begin{pmatrix} x \\ y \end{pmatrix}$

$$\langle u, u\rangle' = x^2 + (x-y)^2 \geq 0,$$

mit Gleichheit genau dann, wenn $u = 0$.

Bemerkung 13.2.4. *Wir koennen 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) *laesst sich als Matrixmultiplikation schreiben:*

$$\langle u_1, u_2\rangle' = u_1^t A u_2 \qquad \textit{mit } A = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

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 $f : [a, b] \to \mathbb{R}$. Wir definieren

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx.$$

Die erfuellt 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*

$$\langle\quad,\quad\rangle : V \times V \to \mathbb{C}$$

mit folgenden Eigenschaften:

(1) Linearitaet in der erste Variablen:

$$\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle \quad \forall\, v_1, v_2, w \in V,$$
$$\langle \alpha v, w\rangle = \alpha\langle v, w\rangle \quad \forall\, \alpha \in \mathbb{C}, \forall\, v, w \in V.$$

*(2) **Sesquilinearitaet** in der zweiten Variablen:*

$$\langle v, w_1 + w_2\rangle = \langle v, w_1\rangle + \langle v, w_2\rangle \quad \forall\, v, w_1, w_2 \in V,$$
$$\langle v, \alpha w\rangle = \overline{\alpha}\langle v, w\rangle \quad \forall\, \alpha \in \mathbb{C}, \forall\, v, w \in V.$$

(3) Hermetische Eigenschaft: $\langle v, w\rangle = \overline{\langle w, v\rangle}$ $\forall, v, w \in V$;

(4) Positivitaet: $\langle v, v\rangle > 0$ $\forall\, v \in V$, $v \neq 0_V$.

Das Paar $(V, \langle\quad,\quad\rangle)$ ist ein Hermetischer *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 $(V, \langle\quad,\quad\rangle)$ 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) $\langle 0_V, v\rangle = \langle v, 0_V\rangle = 0\ \forall v \in V$*;*
(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* $(V, \langle \quad , \quad \rangle)$ *ein Euklidischer oder Hermitescher Raum. Fuer alle* $v \in V$*, definiere*

$$||v|| = \sqrt{\langle v, v\rangle} \in \mathbb{R}_{\geq 0}.$$

Dann ist $|| \sim ||$ *eine Norm.*

Um diesen Satz zu beweisen, brauchen wir folgendes Lemma:

Lemma 13.2.11 (Cauchy-Schwartz Ungleichung)**.** *Es sei* $(V, \langle \quad , \quad \rangle)$ *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 abhaengig sind.*

Proof. Wenn $u = 0_V$, dann haelt 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

$$
\begin{aligned}
&(58) & 0 &\leq ||w||^2 \\
&(59) & &= \langle v - \lambda u, v - \lambda u\rangle \\
& & &= ||v||^2 - \lambda\bar{\lambda}||u||^2 - \bar{\lambda}\lambda||u||^2 + \lambda\bar{\lambda}||u||^2 \\
& & &= ||v||^2 - |\lambda|^2 \cdot ||u||^2 \\
& & &= ||v||^2 - \frac{|\langle u, v\rangle|^2}{||u||^2} \\
& & \Rightarrow &\quad |\langle u, v\rangle| \leq ||u|| \cdot ||v||.
\end{aligned}
$$

In (59) erhalten wir Gleichheit genau dann, wenn $w = 0_V$, d.h. wenn $v = \frac{\langle v, u\rangle}{||u||^2} \cdot u$.

Andererseits, wenn $v = \lambda \cdot u$ fuer $\lambda \in K$ und $u \neq 0_V$, dann gilt $\lambda = \frac{\langle v, u\rangle}{||u||^2}$. Quod erat demonstrandum.

□

von Satz 13.2.10. Wir beweisen den Satz fuer Hermitesche Raeume; der Beweis fuer Euklidische Raeume ist sehr aehnlich. Wir ueberpruefen 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 $\Re(z) \leq |z|$ fuer 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 Uebung.

□

Lemma 13.2.12. *Es sei $(V, \langle \quad , \quad \rangle)$ ein Euklidischer Raum. Dann gilt*

$$\langle u, v\rangle = \frac{1}{2}\left(||u+v||^2 - ||u||^2 - ||v||^2\right)$$

fuer alle $u, v \in V$. Mit anderen Worten, das innere Produkt $\langle \quad , \quad \rangle$ ist durch die Norm $|| \sim ||$ vollstaendig bestimmt.

Proof. Explizite Rechnung. □

Uebung. Formulieren und beweisen Sie das Analog von Lemma 13.2.12 fuer Hermetische Raeume.

Bemerkung 13.2.13. *Nicht jede Norm kann ueber ein inneres Produkt definiert werden! Beispiele dafuer 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

$$\langle \quad , \quad \rangle_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \qquad \langle u, v\rangle_A = v^T \cdot A \cdot v.$$

Dann ist $\langle \quad , \quad \rangle_A$ 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*

$$\langle u, v\rangle_A = \langle v, u\rangle_A \qquad \forall\, u, v \in \mathbb{R}^n.$$

Proof. Es gilt

$$\langle v, u\rangle_A = v^tAu = v^tA^tu = (u^tAv)^t = u^tAv = \langle u, v\rangle_A.$$

□

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 $v^tAv > 0$ 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 $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$, $\alpha_i > 0$ fuer alle i. Dann ist die diagonale Matrix

$$A = \begin{pmatrix} \alpha_1 & 0 & \ldots & 0 \\ 0 & \alpha_2 & \ldots & 0 \\ 0 & 0 & \ldots & \alpha_n \end{pmatrix}$$

positiv definit.

Satz 13.3.7. *Es sei $A \in M_{n\times n}(\mathbb{R})$. Dann ist $\langle \quad , \quad \rangle_A$ 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 $\langle \quad , \quad \rangle$ 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*

$$\langle \quad , \quad \rangle_B : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}, \quad \langle u, v \rangle_B = u^t B \bar{v}.$$

Definition 13.3.10. *Es sei $B \in M_{n \times n}(\mathbb{C})$.*

(1) Schreibe $\bar{B} = (\overline{b_{ij}})_{ij}$. Die adjungierte Matrix *von B ist die Matrix $B^* = (\bar{B})^t$.*

(2) Die Matrix B ist Hermitesch*, wenn $B = B^*$, das heisst, wenn $b_{ij} = \overline{b_{ji}}$ $\forall 1 \le i, j \le n$.*

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 erfuellt $\langle \quad , \quad \rangle_B$ Eigenschaften (1) - (3) des Hermetischen 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 $v^t B v \in \mathbb{R}$ fuer alle $v \in \mathbb{C}^n$.

Mit Hilfe dieser Beobachtung koennen wir deine positiv definite Hermitesche Matrix definieren.

Definition 13.3.13. *Eine Hermitesche Matrix $BM_{n \times n}(\mathbb{C})$ ist positiv definit, wenn*

$$v^t B \bar{v} > 0 \qquad \forall v \in \mathbb{C}^n, \, v \neq 0.$$

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 $\langle \quad , \quad \rangle_B$ 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 $\langle \quad , \quad \rangle_B$ die Axiome eines inneren Produktes erfuellt.

Nimm nun an, dass $\langle \quad , \quad \rangle_B$ ein inneres Produkt ist. Wir zeigen zunaechst, dass B Hermitesch ist. Da $\langle \quad , \quad \rangle_B$ 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 Fuer $1 \le i, j \le n$ gilt

$$b_{ij} = e_i^t B e_j = \langle e_i, e_j \rangle_B = \overline{\langle e_j, e_i \rangle_B} = \overline{e_j}^t \bar{B} \bar{e_i} = e_j^t \bar{B} e_i = \overline{b_{ji}},$$

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. □

Wir koennen also die Vektorraeume $\mathbb{R}^n$ und $\mathbb{C}^n$ mit vielen verschiedenen inneren Produkten ausstatten; die Geometrie der Raeume (z.B. die Frage, ob zwei Vektoren orthogonal sind oder nicht) haengt von der Wahl eines inneren Produktes ab.

[64] Auch im komplexen Fall gibt es wieder eine 1:1 Aequivalenz zwischen IPs und Hermiteschen positiv definiten Matrizen: einem IP $\langle \quad , \quad \rangle$ ordnen wir die Matrix $B = (\langle e_i, e_j \rangle)$ zu. Dann gilt $\langle \quad , \quad \rangle = \langle \quad , \quad \rangle_B$.

13.4. **Orthogonalitaet.** Mit Hilfe eines inneren Produkts 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 $u \perp v$ fuer alle $u, v \in S$.*
(3) Ein orthogonales System $S \subset V$ ist orthonormal*, wenn $||v|| = 1$ fuer alle $v \in S$.*

Beispiele 13.4.2.

(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 $(\mathbb{R}^2, \langle\ \ ,\ \ \rangle')$ 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 Skalarprodukts 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 $(V, \langle\ \ ,\ \ \rangle)$ 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 $(V, \langle\ \ ,\ \ \rangle)$ ein Vektorraum mit einem inneren Produkt, und es sei $v \in V$, $v \neq 0_V$. Definiere*

$$\mathrm{proj}_v : V \to V, \quad u \mapsto \frac{\langle u, v\rangle}{\langle v, v\rangle} \cdot v.$$

Bemerkung 13.4.5. *Ein Vektor $u \in V$ ist genau dann orthogonal zu $v \neq 0_V$, wenn* $\mathrm{proj}_v(u) = 0_V$.

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 $\mathrm{proj}_{e_i}(v) = a_i \cdot e_i$ fuer alle $1 \leq i \leq n$.
(2) Es sei $(\mathbb{R}^2, \langle\ \ ,\ \ \rangle')$ 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 $\mathrm{proj}_v(u) = v$.

Folgendes Lemma wird wichtig sein, wenn wir orthogonale Basen von Vektorraeumen konstruieren. Lecture 41

Lemma 13.4.7. *Es sei $(V, \langle\ \ ,\ \ \rangle)$ 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*

$$w = u - \mathrm{proj}_v(u)$$

orthogonal zu v.

Proof. Uebung. □

Beispiel 13.4.8. Wir ueberpruefen das in Beispiel 13.4.6 (2):

$$\langle u - \mathrm{proj}_v(u), v\rangle' = \left\langle \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle' = 0.$$

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 $v \in V$.

Wir beginnen mit ein paar einfachen Beobachtungen:

Satz 13.5.1. *Es sei $(V, \langle \quad , \quad \rangle)$ ein K-Vekorraum mit einem inneren Produkt.*

(1) Wenn $S \subseteq V$ ein orthogonales System ist, das nicht den Nullvektor enthaelt, dann ist S linear unabhaengig.
(2) Wenn $v_1, \ldots, v_n$ ein orthogonales System ist, $v_i \neq 0$ fuer alle $1 \leq i \leq n$, und $v = a_1 v_1 + \ldots a_n v_n$, dann gilt

$$a_i = \frac{\langle v, v_i\rangle}{\langle v_i, v_i\rangle} \qquad \forall\, 1 \leq i \leq n.$$

Proof. (1) Nimm an, es gibt $v_1, \ldots, v_n \in S$ und $a_1, \ldots, a_n \in K$, so dass

$$a_1 v_1 + \cdots + a_n v_n = 0_V.$$

Dann gilt

$$\langle a_1 v_1 + \cdots + a_n v_n, v_i\rangle = a_i \langle v_i, v_i\rangle = 0.$$

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) Uebung.

□

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* $\mathrm{LH}(\sim)$ *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 $(V, \langle \quad , \quad \rangle)$ ein Vektorraum ueber K ($K = \mathbb{R}$ oder $\mathbb{C}$) mit einem inneren Produkt. Es sei $n = \dim_K V$, und es sei $v_1, \ldots, v_n$ eine beliebige Basis von V. Es sei $w_1, \ldots, w_n$ gegeben durch*

- *es sei $w_1 = v_1$;*
- *fuer $2 \leq j \leq n$ sei*

$$w_j = v_j - \sum_{i=1}^{j-1} \mathrm{proj}_{w_i}(v_j).$$

Dann ist $w_1, \ldots, w_n$ eine orthogonale Basis von V.

Proof. Wir zeigen ueber Induktion nach j, dass $w_1, \ldots, w_j$ folgende Eigenschaften hat

(1) $w_i \neq 0_V$ fuer alle $1 \leq i \leq j$;
(2) $w_1, \ldots, w_j$ ist ein orthogonales System;
(3) $\mathrm{LH}(w_1, \ldots, w_j) = \mathrm{LH}(v_1, \ldots, v_j)$.

Da jedes orthogonale System linear unabhaengig ist, folgt daraus, dass $w_1, \ldots, w_n$ 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

$$v_j = \sum_{i=1}^{j-1} \frac{\langle v_j, w_i\rangle}{\langle w_i, w_i\rangle} \cdot w_i \in \mathrm{LH}(w_1, \ldots, w_{j-1}).$$

Doch $\mathrm{LH}(w_1, \ldots, w_{j-1}) = \mathrm{LH}(v_1, \ldots, v_{j-1})$, und daher

$$v_j \in \mathrm{LH}(v_1, \ldots, v_{j-1}),$$

was ein Widerspruch zur linearen Unabhaengigkeit von $v_1, \ldots, v_n$ ergibt. Daher gilt (1).

Um (2) zu ueberpruefen, berechnen wir $\langle w_j, w_i \rangle$ fuer $1 \leq i < j$:

$$\langle w_j, w_k \rangle = \left\langle v_j - \sum_{i=1}^{j-1} \mathrm{proj}_{w_i}(v_j), w_k \right\rangle$$

Doch $\mathrm{proj}_{w_i}(v_j)$ is ein Vielfaches von w_i und aufgrund der Induktionshypothese $\langle w_i, w_k \rangle = 0$ wenn $i \neq k$. Daher gilt

$$(13.5) = \langle v_j - \mathrm{proj}_{w_k}(v_j), w_k \rangle = 0$$

aufgrund von Lemma 13.4.7.

Nun muessen wir nur noch zeigen, dass $\mathrm{LH}(w_1, \ldots, w_j) = \mathrm{LH}(v_1, \ldots, v_j)$. Aufgrund der Hypothese wissen wir, dass $\mathrm{LH}(w_1, \ldots, w_{j-1}) = \mathrm{LH}(v_1, \ldots, v_{j-1})$. Doch $w_j = v_j - u$ fuer $u \in \mathrm{LH}(w_1, \ldots, w_{j-1})$. Das beendet den Beweis. □

Korollar 13.5.5. *Es sei $(V, \langle\ \ ,\ \ \rangle)$ ein Vektorraum ueber K ($K = \mathbb{R}$ oder $\mathbb{C}$) mit einem inneren Produkt. Dann besitzt V eine orthonormale Basis. Weiterhin gilt: wenn $v_1, \ldots, v_n$ eine orthonormale Basis von V ist und $v = \sum_{i=1}^n a_i v_i$, dann gilt*

$$||v||^2 = |a_1|^2 + \cdots + |a_n|^2.$$

Proof. Es folgt von Theorem 13.5.4, dass V eine orthogonale Basis $w_1, \ldots, w_n$ besitzt. Es sei

$$v_i = \frac{1}{||w_i||} \cdot w_i \qquad \forall\, 1 \leq i \leq n.$$

Dann ist $\{v_1, \ldots, v_n\}$ eine orthonormale Basis.

Wenn $v = \sum_{i=1}^n a_i v_i$, dann gilt

$$\begin{aligned} ||v||^2 &= \langle v \rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i \bar{a_j} \langle v_i, v_j \rangle \\ &= \sum_{i=1}^n |a_i|^2. \end{aligned}$$

□

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 $V = \mathrm{LH}(v_1, v_2) \leq \mathbb{R}^3$ mit dem Standard IP. Es sei $w_1 = v_1$ und

$$\begin{aligned} w_2 &= v_2 - \mathrm{proj}_{w_1}(v_2) \\ &= \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} - \frac{3}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}. \end{aligned}$$

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 $f : [-1, 1] \to \mathbb{R}$ 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 $v_1, v_2, \ldots, v_n$ eine Basis. Dann gibt es ein IP $(\ \ ,\ \ )$ on V, bezueglich dessen die gegebene Basis orthonormal ist: wir definieren einfach $(v_i, v_j) = \delta_{ij}$. Dieses bestimmt aufgrund von Bilinearitaet (bzw. Sesquilinearitaet) ein IP auf V mit der geforderten Eigenschaft.*