Lineare Algebra I

Malte Witte

Wintersemester 2014/15

Inhaltsverzeichnis

0 Einführung

0.1 Organisatorisches

- Vorlesung: Di 09:00-11:00 (d. h. Anfang 9:15, Ende 10:45) L1, Do 14:00-16:00 L1
- Zentralübung: Mi 16:00-18:00 L2 bei Maarten van Pruijssen: Hier können Sie Fragen stellen und Beispiele sehen.
- Übungsgruppen: Do, Fr: Hier werden die Übungsaufgaben besprochen. Einschreiben über PAUL bis 24:00 (Präferenzvergabe). Danach ab Mi ab 20:00: Vergabe der Restplätze. Die Übungen diese Woche finden statt.
- Studienbegleitseiten: Unbedingt Anmelden bei moodle

  `https://moodle.math.uni-paderborn.de`.

  Kurspasswort GAUSS. Dort stehen die Übungsaufgaben, wichtige Infos, Diskussionsforen.
- Sprechstunde: Di 13:00-14:00 D3 221
- Übungsblätter: erscheinen wöchentlich am Dienstag, 4 Aufgaben à 6P + 1 Zusatzaufgabe (6 Bonuspunkte). Abgabe am darauffolgenden Dienstag bis 9:00 Uhr in den vorgesehenen Zettelkästen. Abgabe mit Partner aus der selben Übungsgruppe, handschriftlich. NICHT ABSCHREIBEN!
- Klausurzulassung: 50 Prozent der regulären Punkte + Vorrechnen in Übung.
- Klausur: Voraussichtlich Do 19.02.15, 10-12 (2 Stunden), Genaueres wird noch bekanntgegeben.
- Skript: wird es nicht geben! (Lärmpegel, Von-Hand-ins-Hirn-Prinzip) Untereinander austauschen, falls was nicht mitbekommen.

Literatur:

- M. Artin, *Algebra*
- S. Bosch, *Lineare Algebra*
- G. Fischer, *Lineare Algebra*
- F. Lorenz, *Lineare Algebra I*

0.2 Überblick

Lineare Algebra:
Die lineare Algebra ist ein Teilgebiet der Mathematik, dass sich mit *Vektorräumen* und *linearen Abbildungen* zwischen diesen beschäftigt. Dies schließt insbesondere auch die Betrachtung von linearen Gleichungssystemen mit ein.

Beispiele (aus der Schule)**.**
Für Vektorräume:

- $$\mathbb{R} \text{ (die reellen Zahlen)},$$
- $$\mathbb{R}^2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x, y \in \mathbb{R} \right\},$$
- $$\mathbb{R}^3 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : x, y, z \in \mathbb{R} \right\},$$
- $$\dots$$

Für lineare Abbildungen:

- $$f \colon \mathbb{R}^2 \to \mathbb{R}^2, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ 3x + 4y \end{pmatrix},$$
- $$g \colon \mathbb{R}^2 \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = 3x + 4y,$$

Für ein lineares Gleichungssystem:

$$\begin{aligned} 3x + 4y &= 5 \\ 2x + 5y &= 3 \end{aligned}$$

für Unbestimmte x, y. Andere Schreibweise:

$$\begin{pmatrix} 3 & 4 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}.$$

Dies ist eine Gleichung für Elemente aus $\mathbb{R}^2$.

Rechenoperationen auf Vektorräumen:
Beispiel: $\mathbb{R}^2$. Seien $a, b, c, d, \alpha \in \mathbb{R}$.

$$\text{Addition:} \qquad \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a+c \\ b+d \end{pmatrix},$$

$$\text{Skalarmultiplikation:} \qquad \alpha \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \alpha a \\ \alpha b \end{pmatrix}.$$

Analog für $\mathbb{R}^n$.

Rechengesetze:

(V1) (Assoziativität) $(x + y) + z = x + (y + z)$ für beliebige $x, y, z \in \mathbb{R}^n$

(V2) (neutrales Element) Es gibt ein (eindeutiges) Element $o \in \mathbb{R}^n$ so dass $o+x = x$ für alle $x \in \mathbb{R}^n$.

(V3) (Kommutativität) $x + y = y + x$ für alle $x, y \in \mathbb{R}^n$

(V4) (Verträglichkeit der Multiplikation) $\alpha \cdot \underbrace{(\beta \cdot x)}_{\in \mathbb{R}^n} = \underbrace{(\alpha \cdot \beta)}_{\in \mathbb{R}} \cdot x$ für bel. $\alpha, \beta \in \mathbb{R}, x \in \mathbb{R}^n$

(V5) (erstes Distributivgesetz) $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$, $\alpha, \beta \in \mathbb{R}$, $x \in \mathbb{R}^n$

(V6) (zweites Distributivgesetz) $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$, $\alpha \in \mathbb{R}$, $x, y \in \mathbb{R}^n$

(V7) (Wirkung der $0 \in \mathbb{R}$) $0 \cdot x = o$ für $x \in \mathbb{R}^n$

(V8) (Wirkung der $1 \in \mathbb{R}$) $1 \cdot x = x$, $x \in \mathbb{R}^n$.

Inhalt der Vorlesung:

1. Mathematische Grundlagen: Aussagenlogik, Beweise, Mengentheorie.
2. Gruppen, Ringe, Körper: Definitionen, der Körper der komplexen Zahlen.
3. Vektorräume und lineare Abbildungen: Definition, Existenz von Basen, Dimension.
4. Matrizen: Definition, Matrizenprodukt, Darstellungsmatrix.
5. Lineare Gleichungssysteme: Gauß-Elimination, Lösungsverhalten.
6. Determinante und Eigenwerte: Zur Klassifikation von Endomorphismen, Diagonalisierbarkeit, Trigonalisierbarkeit.

Bevor wir mit Vektorräumen anfangen, müssen wir erst noch ein paar mathematische und logische Grundlagen legen. Mathematiker bedienen sich gerne der Sprache der Aussagenlogik. Wir werden mit einer Einführung dazu beginnen. Außerdem werden wir einige Beweisformen kennenlernen. Wichtig ist hier vor allem das Prinzip der vollständigen Induktion.

Beweiseführen ist das mathematische Tageswerk. Das unterscheidet uns ganz wesentlich von der Schulmathematik, wo das Rechnen, am besten durch Abarbeiten eines vorgegebenen Algorithmus, im Vordergrund steht. Beweise zu führen lernt man meiner Meinung nach allerdings am besten, wenn man es selbst versucht oder vorhandene Beweise nachvollzieht. In meiner Vorlesung werden Sie viele Beweise nachvollziehen dürfen und die Übungsaufgaben sind dazu da, dass Sie sich selbst daran versuchen.

Deshalb werden wir uns nicht lange damit aufhalten, abstrakt über Beweisführung zu reden. Stattdessen wenden wir uns dem nächsten Grundlagenthema zu, nämlich Mengentheorie. Den Begriff der Menge als eine Ansammlung von Dingen sollten Sie aus der Schule kennen. Auf diesem Begriff werden wir aufbauen und ein paar Konstruktionen einführen, die mit Mengen möglich sind.

Eine Menge, die Ihnen sicher ein Begriff ist, ist die Menge der ganzen Zahlen. Auf dieser Menge haben Sie aber noch zusätzliche Struktur, nämlich eine Addition und eine Multiplikation, die einer Reihe von Rechengesetzen unterliegen. Auch auf den Mengen der rationalen Zahlen und den reellen Zahlen haben Sie diese Operationen und die Rechengesetze gelten weiter. Es lohnt sich, diese Situation zu abstrahieren und in neue Begriffe zu gießen, indem wir die Rechengesetze axiomatisch beschreiben. Macht man das nur für die Addition, kommt man auf den Begriff der Gruppe, führt man zusätzlich eine Multiplikation ein, so erhält man den Begriff des Ringes. Der Vorteil von dieser abstrakten Betrachtungsweise ist, dass es noch weit mehr Beispiele für Gruppen und Ringe gibt, als die, die ich hier aufgezählt habe. Für all diese Beispiele kann man nun eine gemeinsame Theorie

entwickeln. Wir folgen damit der axiomatisch-deduktive Herangehensweise der modernen Mathematik: Ausgehend von einem System von Axiomen leiten wir allgemeingültige Gesetzmäßigkeiten ab.

Die Ringe der rationalen und reellen Zahlen haben dem Ring der ganzen Zahlen aber eine wichtige Eigenschaft voraus: In ihnen kann man durch alle Elemente ungleich Null dividieren. Fügt man diese Eigenschaft zu der Liste der Ringaxiome hinzu, kommt man auf den Begriff des Körpers. Auch Körper gibt es sehr viel mehr als nur den Körper der rationalen oder der reellen Zahlen. Wir werden in dieser Vorlesung sehen, dass es zum Beispiel für jede Primzahl p einen Körper mit genau p Elementen gibt. Ein wichtiges Beispiel ist der Körper der komplexen Zahlen, den wir erhalten, indem wir zu den reellen Zahlen formal eine Lösung der Gleichung $x^2 = -1$ hinzufügen.

Vektorräume und lineare Abbildungen können wir nun über einem beliebigen fixierten Körper K einführen:

Definition 0.1. Ein K-Vektorraum ist eine Menge V mit einer Addition und einer Skalarmultiplikation mit Elementen aus K, so dass $(V1) - (V8)$ gelten. Eine lineare Abbildung des K-Vektorraums V in einen K-Vektorraum W ist eine Abbildung der zugrundeliegenden Mengen, die mit der Addition und der Skalarmultiplikation auf V und W verträglich ist.

Wichtig ist, hier zu begreifen, dass V hier nicht notwendig aus mit Zahlen gefüllten Spaltenvektoren bestehen muss. Die Elemente können alles Mögliche sein, solange Sie eine geeignete Addition und Skalarmultiplikation darauf definieren können. Zum Beispiel könnte V aus reellwertigen Funktionen bestehen, etwa den Lösungsfunktionen einer Differentialgleichung.

Unser nächstes Ziel ist es, uns einen vollständigen Überblick über die Gesamtheit aller möglichen Typen von K-Vektorräumen zu verschaffen. Nun kann man aus einem K-Vektorraum V ganz einfach einen neuen Vektorraum V' machen, indem man die Elemente von V durch neue Symbole ersetzt und die Addition und Skalarmultiplikation von V entsprechend auf diese neuen Symbole überträgt. Die Vektoräume V und V' unterscheiden sich dann zwar formal, aber nicht in ihren mathematischen Eigenschaften. Wir sagen: V und V' sind isomorph (gleichgestaltig). Solche Unterschiede interessieren uns nicht, das heißt, wir wollen die Isomorphietypen von K-Vektorräumen klassifizieren. Das zentrale Resultat ist das folgende:

Theorem 0.2.

1. *Jeder K-Vektorraum V hat eine Basis, d. h. ein minimales System von Elementen, so dass sich alle Elemente von V mittels Addition und Skalarmultiplikation aus diesem System gewinnen lässt.*

2. *Die Anzahl der Elemente (Kardinalität) jeder Basis von V ist immer gleich, also nur von V abhängig. Sie wird* Dimension *von V genannt.*

3. *Zwei K-Vektorräume sind genau dann isomorph, wenn sie dieselbe Dimension haben, d. h. der Isomorphietyp eines Vektorraums ist durch seine Dimension vollständig festgelegt.*

Die linearen Abbildungen von einem K-Vektorraum V in einen K-Vektorraum W lassen sich nun wie folgt vollständig beschreiben: Angenommen, V hat die Dimension n und W hat die Dimension m. Indem wir willkürlich eine Basis von V und eine Basis von W auswählen, können wir V mit dem Raum der Spaltenvektoren K^n und W mit K^m identifizieren. Damit reicht es, alle linearen Abbildungen von K^n nach K^m anzugeben. Wir werden zeigen, dass sich jede solche Abbildung durch eine $m \times n$-Matrix, also einer Tabelle mit m Zeilen und n Spalten mit Einträgen aus K, eindeutig beschreiben läßt. Hier gilt es allerdings zu beachten, dass die Matrix, die wir auf diese Weise einer linearen Abbildung von V nach W zuordnen, ganz wesentlich von unserer Auswahl der Basen von V und W abhängt.

Als nächstes wenden wir uns dem Studium von linearen Gleichungssystemen zu. Ein lineares Gleichungsystem ist eine Gleichung $A \cdot x = b$, wobei $b \in K^m$ ein fest vorgegebener Spaltenvektor ist, A eine gegebene $n \times m$-Matrix bezeichnet und x für einen zu bestimmenden Spaltenvektor in K^n steht. Mit anderen Worten: Wir suchen nach allen Vektoren $x \in K^n$, die von der durch die Matrix A (bezüglich der Standardbasis) dargestellten linearen Abbildung auf den Vektor b abgebildet werden. Das wichtigste Verfahren zum Lösen solcher Gleichungssysteme, die Gauß-Elimination, haben Sie vielleicht schon in der Schule in der einen oder anderen Form kennengelernt. Wir wollen hier mathematisch streng beweisen, das dieses Verfahren tatsächlich funktioniert. Außerdem werden wir lineare Gleichungssysteme nach ihrem Lösungsverhalten klassifizieren.

Eine lineare Abbildung eines Vektorraumes V in sich selbst nennt man linearen Endomorphismus. Angenommen, V hat die Dimension n. Durch die Wahl einer Basis von V können wir einem fixierten Endomorphismus $f\colon V \to V$ eine $n \times n$-Matrix A zuordnen. Wie oben erwähnt, hängt A dabei wesentlich von der Wahl der Basis ab. Wie müssen wir die Basis wählen, damit A eine möglichst einfache Gestalt hat? Um der Antwort dieser Frage näher zu kommen, untersuchen wir Kennzahlen, die wir A zuordnen können, deren Wert aber nur vom Endomorphismus f und nicht von der Wahl der Basis abhängen soll. Die wichtigsten solcher „Invarianten" sind die Determinante und die Spur, sowie die Eigenwerte von f. Alle diese Invarianten werden gemeinsam durch das charakteristische Polynom kodiert. Mithilfe dieser Invarianten können wir ein Kriterium aufstellen, wann f durch eine Diagonalmatrix oder eine obere Dreiecksmatrix dargestellt werden kann. Diese Ergebnisse lassen sich noch weiter verfeinern: Tatsächlich kann man die Matrix von f immer in eine besonders übersichtliche Standardform bringen. Um das zu zeigen, müssen wir aber schärfere mathematische Werkzeuge entwickeln. Wir verschieben deshalb die endgültige Antwort auf diese Fragestellung auf das nächste Semester.

1 Mathematische Grundlagen

1.1 Aussagenlogik

Definition 1.1 (Aristoteles)**.** Eine Aussage ist eine sprachliches Gebilde, von dem es sinnvoll ist, zu fragen, ob es entweder wahr oder falsch ist.

Diese Definition hat ihre Tücken!

Beispiele.

- „Für jede natürliche Zahl x gibt es eine natürliche Zahl y mit $x < y$." (Wahre) Aussage
- $1 + 1 = 5$. (Falsche) Aussage
- „Es gibt unendlich viele Primzahlzwillinge, d. h. Paare von Primzahlen $(p, p + 2)$" Aussage, aber unbekannt ob wahr oder falsch.
- $3 + 5$. Keine Aussage
- $x < y$. Keine Aussage, da hier aus dem Kontext nicht ersichtlich ist, was diese Zeichenkette bedeuten soll.
- „Hier steht eine falsche Aussage." Keine Aussage, die Frage nach wahr oder falsch ist hier nicht sinnvoll.
- „Jeder Wuwu hat ein Huhu." Es mag Menschen geben, für die „Wuwu" und „Huhu" wohldefinierte Entitäten sind, die eine Relation des „Habens" zulassen. Für mich jedoch ist dies keine Aussage.

Logische Verknüpfungen: Seien A und B Aussagen. Dann sind auch die folgenden Ausdrücke Aussagen:

$\neg A$ „A ist nicht wahr" (Negation),

$A \wedge B$ „Sowohl A als auch B sind wahr" (Konjunktion),

$A \vee B$ „Mindestens eine der Aussagen A, B ist wahr" (Disjunktion),

$A \Rightarrow B$ „Falls A wahr ist, so auch B" (Implikation/Konditional),

$A \Leftrightarrow B$ „A ist genau dann wahr, wenn B wahr ist" (Äquivalenz/Bikonditional).

Beispiel. Die Aussage $A \Rightarrow B$ ist genau dann wahr, wenn entweder A falsch ist oder sowohl A als auch B richtig sind. Insbesondere ist die Aussage $1 + 1 = 3 \Rightarrow 1 = 0$ wahr.

Operatorenrangfolge: $\neg$ vor $\wedge$ vor $\vee$ vor $\Rightarrow$ vor $\Leftrightarrow$.
Besser: Klammern verwenden.

Quantoren: Wir benutzen Quantorensymbole hier nur als bequeme Abkürzungen, ohne uns ernsthaft mit ihrem prädikatenlogischen Hintergrund auseinanderzusetzen.

- $\forall$ „für alle", z. B.: $\forall x \in \mathbb{R}\colon x+1 > x$ „Für alle reellen Zahlen x gilt $x+1 > x$",
- $\exists$ „es gibt", z. B.: $\exists x \in \mathbb{R}\colon x > 0$ „Es gibt (mindestens) eine reelle Zahl x größer als 0",
- $\exists!$ „es gibt genau eins", z. B.: $\exists! x \in \mathbb{R}\colon x + 1 = 0$ „Es gibt genau eine reelle Lösung der Gleichung $x + 1 = 0$."

1.2 Beweisformen

Definition 1.2. Ein Beweis ist ein Diskurs, der einen imaginären Zweifler, ausgehend von als wahr vorausgesetzten Grundsätzen (Axiomen) und bereits als wahr erkannten Aussagen (Theoremen) restlos von der Gültigkeit der zu beweisenden Aussage überzeugen soll.

Einige wichtige Beweisformen:

Modus Ponens (direkter Beweis). Aus der Wahrheit von A und $A \Rightarrow B$ folgt die Wahrheit von B.

Modus Pollens (indirekter Beweis). Aus der Wahrheit von $\neg B$ und $\neg A \Rightarrow B$, folgt die Wahrheit von A.

Modus Barbara (Kettenschluss). Aus der Wahrheit von $A \Rightarrow B$ und $B \Rightarrow C$ folgt die Wahrheit von $A \Rightarrow C$.

Reductio ad adsurdum (Widerspruchsbeweis). Aus der Wahrheit von $\neg A \Rightarrow B \wedge \neg B$ folgt die Wahrheit von A.

Aufteilung der Äquivalenz. Aus der Wahrheit von $A \Rightarrow B$ und $B \Rightarrow A$ folgt die Wahrheit von $A \Leftrightarrow B$.

Vollständige Induktion. Für jede natürliche Zahl $n = 1, 2, \ldots$ sei eine Aussage A_n gegeben. Angenommen

Induktionsanfang: A_1 ist wahr,

Induktionsschritt: für alle n ist $A_n \Rightarrow A_{n+1}$ wahr.

Dann folgt die Wahrheit von A_n für alle n.

Beispiel. Sei A_n die Aussage $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$. Dann ist $A_1 \colon 1 = \frac{1(1+1)}{2}$ offensichtlich wahr. Angenommen, A_n gilt. Dann folgt

$$1 + 2 + \cdots + n + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2},$$

also A_{n+1}. Mit anderen Worten, $A_n \Rightarrow A_{n+1}$ ist wahr. Nach dem Prinzip der vollständigen Induktion gilt also A_n für alle n.

Beispiel. *Behauptung:* Seien n Zahlen $a_1 \leq 1, \ldots, a_n \leq 1$ gegeben, so dass $a_1 + \cdots + a_n \geq n$. Dann gilt $a_1 = \cdots = a_n = 1$.
Beweis (durch Induktion über n)
Induktionsanfang: Klar gilt $(a_1 \leq 1) \wedge (a_1 \geq 1) \Rightarrow a_1 = 1$.
Induktionsschritt: Sei die Aussage für n bewiesen. Wir betrachten die Aussage für $n+1$. Wegen $a_{n+1} \leq 1$ gilt $a_1 + \cdots + a_n \geq n + 1 - a_{n+1} \geq n$. Nach Induktionsannahme gilt $a_1 = \ldots a_n = 1$, also $a_{n+1} \geq n + 1 - (a_1 + \cdots + a_n) = 1$ und somit $a_{n+1} = \cdots = a_1 = 1$.

1.3 Mengen

Definition 1.3 (Cantor)**.** Eine Menge ist eine Zusammenfassung bestimmter, wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens zu einem Ganzen.

Diese naive Definition bringt Probleme mit sich, reicht aber für unsere Zwecke. Es gibt einen strikten, axiomatischen Zugang zur Mengenlehre (ZFC-Axiomensystem).

Schreibweise für Mengen
$M = \{a, b, c, \ldots\}$
z.B. $\{1, 2, 3\} = \{1, 3, 2\} = \{3, 2, 1, 1\}$
$a \in M$: a ist Element von M
$a \notin M$: a ist nicht Element von M
andere Schreibweise $M = \{A \mid B\}$ ist die Menge der Objekte der Form A, die der Bedingung B genügen, z.B. $M = \{x \mid x \in \mathbb{R},\ x \leq 5\}$ oder kürzer $M = \{x \in \mathbb{R} \mid x \leq 5\}$.
$\mathbb{N} = \{1, 2, 3, \ldots\}$ natürliche Zahlen
$\mathbb{N}_0 = \{0, 1, 2, \ldots\}$,
$\mathbb{Z} = \{\ldots, -2, 1, 0, 1, 2, \ldots\}$ ganze Zahlen
$\mathbb{Q} = \{\frac{m}{n} \mid m \in \mathbb{Z}, n \in \mathbb{N}\}$ rationale Zahlen
$\mathbb{R}$ reelle Zahlen.

Definition 1.4. Die *leere Menge* $\emptyset$ ist die Menge, die kein Element enthält.

Definition 1.5. Eine Menge N heißt Teilmenge der Menge M ($N \subset M$), wenn M alle Elemente aus N enthält.

Beispiele.

- $\{1, 2\} \subset \mathbb{N}$
- die leere Menge $\emptyset$ ist Teilmenge jeder Menge.

Bemerkung. Zwei Mengen M, N sind genau dann gleich, wenn $M \subset N$ und $N \subset M$. Das ist für Beweise sehr nützlich.

Bemerkung. Anstelle von $N \subset M$ wird oft auch $N \subseteq M$ oder $N \subseteqq M$ geschrieben. Diese drei Symbole sind vollkommen gleichwertig. Die Schreibweisen $N \subsetneq M$ oder auch $N \subsetneqq M$ oder auch $N \subsetneq M$ bedeuten: N ist Teilmenge von M aber nicht gleich M.

Definition 1.6. Die Menge aller Teilmengen einer Menge M heißt *Potenzmenge* von M und wird mit $\mathcal{P}(M)$ bezeichnet.

Beispiele.

- $M = \{0, 1\} \Longrightarrow \mathcal{P}(M) = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$
- Ist M eine endliche Menge mit n Elementen, so ist $\mathcal{P}(M)$ eine endliche Menge mit 2^n Elementen.

Endlich viele (nicht notwendig endliche) Mengen werden üblicherweise durch Indizes durchnummeriert $M_1, \ldots, M_n$. Allgemeiner:

Definition 1.7. Eine Familie von Mengen ist eine Vorschrift, die jedem Element i einer Menge I eine Menge M_i zuordnet. Schreibweise: $(M_i)_{i \in I}$. Die Menge I wird auch *Indexmenge* genannt.

Definition 1.8. Seien $K, L, (M_i)_{i \in I}$ Teilmengen einer Menge M. Dann bildet man die folgenden Mengen

(i) $\bigcup\limits_{i \in I} M_i = \{m \in M \mid \text{es gibt ein } i \in I \text{ mit } m \in M_i\}$ (Vereinigung)

(ii) $\bigcap\limits_{i \in I} M_i = \{m \in M \mid m \in M_i \text{ für alle } i \in I\}$ (Durchschnitt)

(iii) $K \setminus L = \{m \in K \mid m \notin L\}$ (Komplement, auch $K - L$)

Beispiel. $M = \mathbb{N}$

- $\{1, 2, 3\} - \{3, 4, 5\} = \{1, 2\}$,
- $\{1, 2\} - \{1, 2, 3\} = \emptyset$.

Definition 1.9. (Produktmenge, kartesisches Produkt). Es seien $M_1, \ldots, M_n$ Mengen. Die *Produktmenge*

$$M_1 \times \cdots \times M_n$$

besteht aus allen n-Tupeln $(m_1, \ldots, m_n)$ mit $m_1 \in M_1, \ldots, m_n \in M_n$.

Beispiel.

- $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$
- $\mathbb{R} \times \emptyset = \emptyset$

Definition 1.9 dehnt sich auf Familien aus.

Definition 1.10. Sei $(M_i)_{i \in I}$ eine Familie von Mengen. Wir schreiben $(m_i)_{i \in I}$ für die Vorschrift, die jedem Element i der Indexmenge I ein Element m_i der Menge M_i zuordnet. Die Produktmenge der Familie $(M_i)_{i \in I}$ ist die Menge dieser Vorschriften:

$$\prod_{i \in I} M_i = \{(m_i)_{i \in I} \mid m_i \in M_i\}.$$

Definition 1.11. Sei $(M_i)_{i \in I}$ eine Familie von Teilmengen einer Menge M. Man sagt, M ist die disjunkte Vereinigung der M_i und schreibt

$$M = \dot{\bigcup_{i \in I}} M_i \text{ oder } M = \bigsqcup_{i \in I} M_i$$

falls $M = \bigcup_{i \in I} M_i$ und $M_i \cap M_j = \emptyset$ für $i \neq j$. Ist $(M_i)_{i \in I}$ eine beliebige Familie von Mengen, so kann man jedes M_i als eine Teilmenge der Menge

$$M = \{(m, i) \mid i \in I, m \in M_i\}$$

ansehen, indem man $m \in M_i$ mit dem Paar (m, i) identifiziert. Auch hier sagt man, M ist die disjunkte Vereinigung der M_i.

1.4 Äquivalenzrelationen

Definition 1.12. Eine *(binäre) Relation* R auf einer Menge M ist eine Teilmenge $R \subset M \times M$.
Sprechweise: $x, y \in M$ stehen in Relation R, wenn $(x, y) \in R$.
Schreibweise $x \sim_R y$.

Beispiele.

1. $M = \mathbb{R},\ R = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x < y\}$

2. $M = \mathbb{Z},\ R = \{(m,n) \in \mathbb{Z} \times \mathbb{Z} \mid m - n \quad \text{ist gerade}\}$

3. $M =$ die Menge der Schüler einer Schule
   $R = \{(x,y) \in M \times M \mid x$ und y gehen in die gleiche Klasse$\}$

Definition 1.13. Sei M eine Menge und R eine Relation auf M. R heißt *Äquivalenzrelation* wenn die folgenden Bedingungen erfüllt sind

(Ä1) Reflexivität: $x \sim_R x$ für alle $x \in M$

(Ä2) Symmetrie: $x \sim_R y \Longrightarrow y \sim_R x$

(Ä3) Transitivität: $(x \sim_R y$ und $y \sim_R z) \Longrightarrow x \sim_R z$.

Beispiele.

- die Relation in Beispiel 1 ist nicht reflexiv, nicht symmetrisch, aber transitiv,
- die Relationen in Beispiel 2 und Beispiel 3 sind Äquivalenzrelationen.

Bemerkung. Auf jeder Menge existiert die (nutzlose) Äquivalenzrelationen "=", d.h. $R = \{(x,y) \in M \times M \mid x = y\}$.

Typischerweise existieren auf einer Menge verschiedene Äquivalenzrelationen. So kann man z.B. auf einer Menge von Bauklötzchen die Äquivalenzrelationen "gleiche Farbe", "gleiche Größe" oder "gleiche Form" einführen.

Definition 1.14. Sei M eine nichtleere Menge und R eine Äquivalenzrelation auf M. Eine nichtleere Teilmenge $A \subset M$ heißt *Äquivalenzklasse*, wenn sie den folgenden Bedingungen genügt:

(i) $\forall a, b \in A \colon a \sim_R b$

(ii) $\forall a, b \in M \colon (a \in A) \wedge (a \sim_R b) \Longrightarrow b \in A$.

Lemma 1.15. *Ist R eine Äquivalenzrelation auf einer Menge M, so gehört jedes Element $x \in M$ zu genau einer Äquivalenzklasse.*

Beweis. Wir zerlegen die Behauptung in zwei zu beweisende Teilaussagen:

Existenz: Es gibt eine Äquivalenzklasse, die x enthält.

Eindeutigkeit: $(x \in A$ und $x \in A') \Longrightarrow A = A'$.

(Das ist das Standardverfahren, um eine Aussage zu beweisen, die ein „es gibt genau ein" enthält.)

Beweis der Existenz: Definiere

$$A := \{a \in M \mid x \sim_R a\}.$$

Wegen $x \sim_R x$ (siehe Ä1) gilt $x \in A$, also $A \neq \emptyset$. Es verbleibt, die Bedingungen (i) und (ii) aus Definition 1.14 zu verifizieren.

(i) Seien $a, b \in A$. Dann gilt $x \sim_R a$ und $x \sim_R b$. Aus (Ä2) folgt $a \sim_R x$ und (Ä3) liefert $a \sim_R b$.

(ii) Sei $a \in A$ und $a \sim_R b$. Zu zeigen: $b \in A$. Nach Definition gilt $x \sim_R a$. (Ä3) liefert $x \sim_R b$, also $b \in A$.

Beweis der Eindeutigkeit: Wir zeigen $A \subset A'$. Der Nachweis von $A' \subset A$ ist aus Symmetriegründen derselbe, und wir haben die Implikation

$$(A \subset A') \wedge (A' \subset A) \Rightarrow A = A'.$$

Sei nun $a \in A$. Dann gilt wegen $x \in A$ $a \sim_R x$. Wegen $x \in A'$ folgt $a \in A'$, also $A \subset A'$. □

Bemerkung. Insbesondere gilt für zwei Äquivalenzklassen A, A' dass entweder $A = A'$ oder $A \cap A' = \emptyset$.

Mit anderen Worten: M zerfällt in die disjunkte Vereinigung der Restklassen bzgl. R.

Definition 1.16. Die Menge der Äquivalenzklassen einer Menge M bzgl. einer Äquivalenzrelation R wird mit M/R bezeichnet.

Beispiele.

- In Beispiel 2 nach Definition 1.12 gibt es zwei Äquivalenzklassen

  $$A_{\text{gerade}} = \{\ldots, -4, -2, 0, 2, 4, \ldots\}$$

  und

  $$A_{\text{ungerade}} = \{\ldots, -3, -1, 1, 3, \ldots\}$$

- in Beispiel 3 nach Definition 1.12 ist die Menge der Äquivalenzklassen die Menge der Schulklassen der Schule.

Hintergrund: Der Übergang zu Äquivalenzklassen soll (für ein jeweils gegebenes Problem) nicht relevante Information abstreifen. So ist für die Erstellung eines Stundenplans nur die Menge der Schulklassen relevant, nicht die (größere) Menge der Schüler.

Beispiel. Sei $n \in \mathbb{N}$. Wir betrachten die Relation auf $\mathbb{Z}$: $a \sim b \iff n|(a-b)$. Dies ist eine Äquivalenzrelation denn
(Ä1) $a - a = 0$, $n|0$, also $a \sim a$.
(Ä2) $a \sim b \implies n|(a-b) \implies n|(b-a) \implies b \sim a$.
(Ä3) ($a \sim b$ und $b \sim c$) $\implies n|(a-b)$ und $n|(b-c) \implies n|((a-b)+(b-c)) = (a-c) \implies a \sim c$.

Es gibt genau n verschiedene Äquivalenzklassen, die mit

$$\begin{aligned} \overline{0} &= \{kn \mid k \in \mathbb{Z}\}, \\ \overline{1} &= \{kn+1 \mid k \in \mathbb{Z}\}, \\ &\dots \\ \overline{n-1} &= \{kn+(n-1) \mid k \in \mathbb{Z}\} \end{aligned}$$

bezeichnet werden. Die Menge der Äquivalenzklassen heißt die Menge der *Restklassen modulo* n und wird mit

$$\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \dots, \overline{n-1}\}$$

bezeichnet. Man schreibt: $a \equiv b \bmod n$, wenn a und b in der gleichen Restklasse liegen.

1.5 Abbildungen

Definition 1.17. Eine Abbildung $f\colon M \to N$ einer Menge M in eine Menge N ist eine Vorschrift, die jedem Element $m \in M$ genau ein Element $f(m) \in N$ zuordnet.

Beispiele.

- $q\colon \mathbb{Z} \to \mathbb{Z}\ ,\ x \longmapsto x^2$

- Ist $M \subset N$ eine Teilmenge, so gibt es die *kanonische Inklusionsabbildung* $i\colon M \to N$,

  m (aufgefaßt als Element von M) $\longmapsto m$ (aufgefaßt als Element von N).

  z.B. $\{0,1,2\} \xrightarrow{i} \{0,1,2,3\}$. Ist $M = N$, so ist dies die sogenannte *Identitätsabbildung* $\mathrm{id}\colon M \to M$, $m \mapsto m$.

- Für jede Menge M gibt es genau eine Abbildung $\emptyset \to M$, die *leere Abbildung*, die den nicht vorhandenen Elementen nichts zuordnet. Sie ist auch gleichzeitig die kanonische Inklusionsabbildung.

Definition 1.18. Zwei Abbildungen $f, g\colon M \to N$ heißen gleich, wenn $f(m) = g(m)$ für alle $m \in M$ gilt.

Beispiel. Die Abbildungen $f, g\colon \mathbb{R} \to \mathbb{R}$
$f(x) = x^2 + 2x + 1,\ g(x) = (x+1)^2$ sind gleich.

Definition 1.19. Sei $f\colon M \to N$ eine Mengenabbildung

(i) für $n \in N$ heißt die Teilmenge von M

$$f^{-1}(n) = \{m \in M \mid f(m) = n\}$$

die *Urbildmenge* von n. Die Menge

$$f(M) = \{f(m) \in N \mid m \in M\}$$

der $n \in N$ mit $f^{-1}(n) \neq \emptyset$ heißt das *Bild* von f. Andere Schreibweisen: Bild(f) oder
$im(f)$.

(ii) f heißt *injektiv*, wenn gilt $f(m) = f(m') \Rightarrow m = m'$ (äquivalent: für jedes $n \in N$ enthält $f^{-1}(n)$ höchstens ein Element).

(iii) f heißt *surjektiv*, wenn zu jedem $n \in N$ ein $m \in M$ mit $f(m) = n$ existiert (äquivalent: $f^{-1}(n) \neq \emptyset$ für alle $n \in N$).

(iv) f heißt *bijektiv* falls es surjektiv und injektiv ist. (äquivalent: $f^{-1}(n)$ enthält für jedes $n \in N$ genau ein Element).

Ist $f\colon M \to N$ bijektiv, so definiert man die Umkehrabbildung $f^{-1}\colon N \to M$ durch die Regel:

$f^{-1}(n) =$ DAS Element der Menge $f^{-1}(n)$.

(Diese Bezeichnungsdoppelung bringt in der Praxis typischerweise keine Probleme mit sich.)

Bemerkung. Sei $f : M \to N$ eine Mengenabbildung. Die Eigenschaften injektiv, surjektiv und bijektiv signalisiert man durch Modifikation des Pfeils:
$f : M \hookrightarrow N$ (injektiv), $f : M \twoheadrightarrow N$ (surjektiv), $f : M \xrightarrow{\sim} N$ (bijektiv)

Definition 1.20. Sei M eine Menge und R eine Äquivalenzrelation. Dann ist die *kanonische Projektion*

$$p\colon M \to M/R$$

definiert durch:
$m \in M$ geht auf die eindeutig bestimmte Klasse $A \in M/R$ mit $m \in A$.

Bemerkung. Es gilt $p^{-1}(A) = A$. Da Äquivalenzklassen per definitionem nicht-leer sind, ist die kanonische Projektion eine surjektive Mengenabbildung.

Beispiel. Die Klasse 12B bestehe aus den Schülern {Albert, Berta, ...}. Für den Mathematiker IST die Klasse 12B eine Menge, nämlich 12B = {Albert, Berta, ...}. Die kanonische Projektion

$$\begin{array}{ccc} p\colon M & \longrightarrow & M/R \\ \text{Menge der Schüler der Schule} & & \text{Menge der Schulklassen der Schule} \end{array}$$

ordnet jedem Schüler seine Klasse zu.

$$\begin{aligned} p^{-1}(12B) &= \text{die Menge der Schüler der Klasse 12B} \\ &= \{\text{Albert, Berta}, \dots\} \\ &= 12B. \end{aligned}$$

Definition 1.21. Sei M eine endliche Menge. Die Anzahl der Elemente von M bezeichnet man mit $\#M$ oder auch mit card (M).

Beispiel.
$\#\emptyset = 0$
$\#\{2, 7, 9\} = 3$

Lemma 1.22. *Sei $f\colon M \to N$ eine Abbildung endlicher Mengen*
(i) Ist f injektiv, so gilt $\#M \leq \#N$
(ii) ist f surjektiv, so gilt $\#M \geq \#N$
(iii) ist f bijektiv, so gilt $\#M = \#N$

Beweis. einfach. (ähnlich wie unten) □

Lemma 1.23. *Sei $f\colon M \to M$ eine Selbstabbildung einer endlichen Menge M. Dann sind die folgenden Aussagen (paarweise) äquivalent*
(i) f ist injektiv
(ii) f ist surjektiv
(iii) f ist bijektiv.

Beweis. Nach der Aufteilung der Äquivalenz und Kettenschluss reicht es, $(i) \Rightarrow (iii)$, $(ii) \Rightarrow (iii)$, $(iii) \Rightarrow (i)$ und $(iii) \Rightarrow (ii)$ zu zeigen.
(i) ⇒ (iii) Sei f injektiv. Dann gilt für jedes $m \in M$: $\#f^{-1}(m) \leq 1$.
Nun zerfällt M in die disjunkte Vereinigung der Urbildmengen: $M = \dot{\bigcup}_{m \in M} f^{-1}(m)$.
Daher gilt

$$\#M = \sum_{m \in M} \#f^{-1}(m) \leq \sum_{m \in M} 1 = \#M.$$

Daher gilt in der Mitte Gleichheit, also $\#f^{-1}(m) = 1$ für alle $m \in M$, d.h. f ist bijektiv.
(ii) $\Longrightarrow$ (iii) analog, hier haben wir $\#f^{-1}(m) \geq 1$ für alle m.
(iii) $\Longrightarrow$ (i) und (iii) $\Longrightarrow$ (ii) sind trivial. □

Die Gesamtheit aller Abbildungen einer Menge M in einer Menge N ist wieder eine Menge und wird mit $\mathrm{Abb}(M, N)$ bezeichnet.

Definition 1.24. Seien M, N, K Mengen und $f\colon M \to N$, $g\colon N \to K$ Abbildungen. Die Abbildung $g \circ f\colon M \to K$, $m \mapsto g(f(m))$ heißt die *Komposition* von f und g. Die Komposition kann man als Mengenabbildung auffassen

$$\begin{array}{rccc} \circ \ : & \mathrm{Abb}(N,K) \times \mathrm{Abb}(M,N) & \longrightarrow & \mathrm{Abb}(M,K) \\ & (g,f) & \longmapsto & g \circ f \end{array}$$

Die Menge der Abbildungen $\mathrm{Abb}(M, N)$ wird auch mit N^M bezeichnet. Der Grund dafür ist

Lemma 1.25. *Seien I und M Mengen und sei $(M_i)_{i\in I}$ die Familie von (immer gleichen) Mengen $M_i = M$ indiziert über $i \in I$. Dann existiert eine natürliche Bijektion*

$$\Phi : \quad M^I \xrightarrow{\sim} \prod_{i\in I} M_i.$$

Beweis. Die rechte Seite ist die Menge aller Tupel $(m_i)_{i\in I}$, $m_i \in M_i = M$. Die linke Seite ist die Menge der Abbildung $f\colon I \to M$. Eine solche Abbildung ist dadurch gegeben, dass man jedem $i \in I$ ein $m_i = f(i) \in M$ zuordnet. Wir definieren Φ durch die Zuordnung

$$f \in M^I \longmapsto (f(i))_{i\in I} \in \prod_{i\in I} M_i.$$

Da die Abbildung f durch den Wert $f(i) \in M$, $i \in I$ gegeben ist, ist Φ injektiv. Ist umgekehrt $(m_i)_{i\in I} \in \prod\limits_{i\in I} M_i$ gegeben, so ist die Abbildung

$$f\colon I \longrightarrow M,\ i \longrightarrow m_i \in M$$

ein Urbild unter Φ. Daher ist Φ auch surjektiv. □

Bemerkung. Das Adjektiv *natürlich/kanonisch* steht hier dafür, dass man die Abbildung Φ angeben kann, ohne die spezielle Gestalt der Elemente von I und M zu kennen. Die Abbildung Φ ergibt sich aus der Konstruktion der Objekte M^I und $\prod\limits_{i\in I} M_i$ heraus. Tatsächlich steht hinter dieser etwas laxen Definition ein präziser Begriff der Kategorientheorie, auf die wir hier aber nicht eingehen werden.

2 Gruppen, Ringe, Körper

2.1 Gruppen

Definition 2.1. Eine (binäre) *Verknüpfung* auf einer Menge M ist eine Abbildung

$$*\colon M \times M \to M,\ (m,n) \mapsto m * n.$$

Definition 2.2. Eine *Gruppe* $(G, *, e)$ ist eine Menge G mit einer Verknüpfung $*$ und einem ausgezeichneten Element $e \in G$, sodass

(G1) $g * (h * k) = (g * h) * k$ für alle $g, h, k \in G$ (Assoziativität)
(G2) $e * g = g$ für alle $g \in G$ (Existenz eines linksneutralen Elements)
(G3) für alle $g \in G$ existiert ein $h \in G$ mit $h * g = e$ (Ex. eines Linksinversen)

Eine Gruppe heißt kommutativ (oder abelsch), wenn zusätzlich gilt:

(G4) $g * h = h * g$ für alle $g, h \in G$.

Beispiele.
1. $(\mathbb{Z}, +, 0)$ ist eine abelsche Gruppe
2. $(\mathbb{Q}, +, 0)$, $(\mathbb{R}, +, 0)$ sind abelsche Gruppen
3. $(\mathbb{Q} \setminus \{0\}, \cdot, 1)$ ist eine abelsche Gruppe
4. Sei $\mathbb{R}_{>0} = \{x \in \mathbb{R} \mid ; x > 0\}$. Dann ist $(\mathbb{R}_{>0}, \cdot, 1)$ eine abelsche Gruppe

Beispiel. Sei $n \in \mathbb{N}$. Wir definieren wie folgt eine Verknüpfung

$$+\colon \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}\,:$$

Seien $A, B \in \mathbb{Z}/n\mathbb{Z}$. Wähle Repräsentanten $a \in A$, $b \in B$ und setze

$$A + B = \overline{a + b},$$

d. h. die Restklasse, zu der $a + b$ gehört.

Wir müssen zeigen, dass diese Abbildung *wohldefiniert* ist, d. h. nicht von der speziellen Wahl der Repräsentanten $a \in A$ und $b \in B$ abhängt: Seien $a' \in A$, $b' \in B$ weitere Repräsentanten. Dann gilt

$$n \mid a - a', \qquad n \mid b - b', \qquad \text{also auch } n \mid (a + b) - (a' + b').$$

Mit anderen Worten, $\overline{a' + b'} = \overline{a + b}$.

Es gilt: $(\mathbb{Z}/n\mathbb{Z}, +, \overline{0})$ ist eine abelsche Gruppe. Den Beweis führen wir später.

Definition 2.3. Sei $n \in \mathbb{N}$ eine natürliche Zahl. Die *symmetrische Gruppe* vom Grad n ist wie folgt gegeben:

$S_n \stackrel{df}{=}$ die Menge aller bijektiven Abbildungen

$$\pi\colon \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$$

(sogenannte Permutationen).
$* = \circ$ Komposition (d.h. Hintereinanderausführung) von Abbildungen
$e = \mathrm{id}_{\{1,\dots,n\}}$ die identische Abbildung.

Wir schreiben Permutationen in der Form

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(n) \end{pmatrix}$$

– oben stehen die Zahlen 1 bis n in der gewöhnlichen Reihenfolge.
– unten stehen die Zahlen 1 bis n in einer (evtl.) anderen Reihenfolge.

Umgekehrt definiert ein solches Diagramm eine Permutation. Wie viele gibt es?
Elementare Kombinatorik:
n Möglichkeiten für die 1
$(n-1)$ Möglichkeiten für die 2
$\vdots$
1 Möglichkeit für n
Daher gilt:

$$\#S_n = n(n-1)\dots 2 \cdot 1 = n! \ (n \text{ Fakultät})$$

Lemma 2.4. $(S_n, \circ, \mathrm{id})$ *ist eine Gruppe.*

Beweis. Wir verifizieren (G1) – (G3).

(G1) Sei $\rho, \sigma, \tau \in S_n$ und $k \in \{1, \dots, n\}$ Dann gilt

$$(\rho \circ (\sigma \circ \tau))(k) = \rho(\sigma(\tau(k))) = ((\rho \circ \sigma) \circ \tau)(k),$$

also

$$\rho * (\sigma * \tau) = \rho \circ (\sigma \circ \tau) = (\rho \circ \sigma) \circ \tau = (\rho * \sigma) * \tau$$

(G2) Sei $\sigma \in S_n$. Dann gilt $e * \sigma = \mathrm{id} \circ \sigma = \sigma$
(G3) Sei $\sigma \in S_n$. Da σ bijektiv ist, existiert die inverse Abbildung $\tau = \sigma^{-1}$. Diese ist wieder bijektiv und es gilt $\tau * \sigma = \tau \circ \sigma = \mathrm{id} = e$. □

Für $n \geq 3$ ist S_n nicht kommutativ:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & \dots \\ 2 & 1 & 3 & 4 & \dots \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & \dots \\ 1 & 3 & 2 & 4 & \dots \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots \\ 2 & 3 & 1 & 4 & \dots \end{pmatrix}$$

aber

$$\begin{pmatrix} 1 & 2 & 3 & 4 & \dots \\ 1 & 3 & 2 & 4 & \dots \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & \dots \\ 2 & 1 & 3 & 4 & \dots \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots \\ 3 & 1 & 2 & 4 & \dots \end{pmatrix}$$

Allgemeine Eigenschaften von Gruppen:

Satz 2.5. *Sei $G = (G, *, e)$ eine Gruppe. Dann gilt für alle $g, h, k \in G$*

(1) aus $g * h = g * k$ folgt $h = k$ (Linkskürzung).
(2) aus $g * h = k * h$ folgt $g = k$ (Rechtskürzung).
(3) $g * e = e * g = g$ (das (links)neutrale Element ist auch rechtsneutral).
(4) aus $g * h = g$ oder $h * g = g$ für ein $g \in G$ folgt $h = e$.
(5) für alle $g \in G$ gibt es ein eindeutig bestimmtes $g^{-1} \in G$ mit $g^{-1} * g = e = g * g^{-1}$.
(6) aus $h * g = e$ oder $g * h = e$ folgt $h = g^{-1}$.
(7) Es gilt $(g^{-1})^{-1} = g$.

Beweis.
(1) Sei $g * h = g * k$. Nach (G3) ex. ein $s \in G$ mit $s * g = e$. Daher gilt

$$s * (g * h) \overset{G1}{=} (s * g) * h = e * h \overset{G2}{=} h$$

analog

$$s * (g * k) = (s * g) * k = e * k = k$$

Somit folgt $h = k$.
(3) $e * g = g$ gilt nach (G2). Nach (G3) ex. ein $h \in G$ mit $h * g = e$. Daher gilt

$$h * (g * e) = (h * g) * e = e * e = e = h * g$$

Nach (1) folgt $g * e = g$.
(5, Existenz) Sei $h \in G$ mit $h * g = e$ (existiert nach G3). Dann gilt

$$h * (g * h) = (h * g) * h = e * h = h \overset{(3)}{=} h * e$$

Nach (1) folgt $g * h = e$, d.h. h ist auch ein Rechtsinverses.
(2) Sei $g * h = k * h$. Sei $s \in G$ (nach 5) so dass $h * s = e$. Dann gilt

$$(g * h) * s = g * (h * s) = g * e \overset{(3)}{=} g$$

analog

$$(k * h) * s = k * (h * s) = k * e = k.$$

Somit folgt $g = k$.
(4) $g * h = g \overset{(3)}{=} g * e \overset{(1)}{\Longrightarrow} h = e$, analog: $h * g = g = e * g \overset{(2)}{\Longrightarrow} h = e$.
(5, Eindeutigkeit) und (6) Seien $h, h' \in G$ mit $h * g = e = h' * g$. Mit (2) folgt $h = h'$ Daher ist $\Rightarrow g^{-1}$ ist eindeutig in G. Seien $h \in G$ mit $g * h = e$. Wegen $g * g^{-1} = e$ folgt mit (1) dass $h = g^{-1}$.
(7) aus $g * (g^{-1}) = e$ folgt $g = (g^{-1})^{-1}$ □

Bemerkung. Sind $g, g' \in G$ so gilt

$$(g * g')^{-1} = (g')^{-1} * g^{-1}.$$

Begründung: $((g')^{-1} * g^{-1}) * g * g' = (g')^{-1} * e * g' = e$.

Bemerkung. Man schreibt auch G für die Gruppe $(G, *, e)$, 1 für das neutrale Element e, gh für $g * h$. Falls G kommutativ ist, schreibt man 0 für e und $g + h$ für $g * h$, $-g$ für g^{-1}.

2.2 Ringe

Definition 2.6. Ein *Ring* $R = (R, +, \cdot, 0_R)$ ist eine Menge R mit zwei Verknüpfungen $+, \cdot\colon R \times R \to R$ und einem ausgezeichneten Element $0_R \in R$ so dass gilt:

(R1) $(R, +, 0_R)$ ist eine abelsche Gruppe,

(R2) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ für alle $a, b, c \in R$,

(R3) $a(b + c) = ab + ac$, $(a + b)c = ac + bc$ für alle $a, b, c \in R$.

Ein *Ring mit 1* (unitärer Ring) ist ein Tupel $R = (R, +, \cdot, 0_R, 1_R)$, so dass $(R, +, \cdot, 0_R)$ ein Ring ist und außerdem noch gilt:

(R4) $1_R \cdot a = a = a \cdot 1_R$ für alle $a \in R$.

Ein Ring heißt kommutativ, wenn die Multiplikation kommutativ ist, also wenn

(R5) $a \cdot b = b \cdot a$ für alle $a, b \in R$ gilt.

Bemerkung. Ein Inverses bzgl. der Multiplikation existiert i. A. nicht.

Beispiele.

1. $(\mathbb{Z}, +, \cdot, 0, 1)$ ist ein kommutativer Ring mit 1
2. $\mathbb{Q}$, $\mathbb{R}$ analog.
3. $\mathbb{Z}/n\mathbb{Z}$ ist ein kommutativer Ring mit 1. Multiplikationsvorschrift für $A \cdot B$.
   (i) Wähle Repräsentanten $a \in A$, $b \in B$.
   (ii) Bilde $a \cdot b$ in $\mathbb{Z}$.
   (iii) $A \cdot B :=$ Klasse von $a \cdot b$.

   Wieder muß nachgewiesen werden, dass das Ergebnis unabhängig von der Auswahl in (i) ist.
4. Die Menge der geraden ganzen Zahlen ist ein kommutativer Ring ohne 1.

Lemma 2.7. *In einem Ring $R = (R, +, \cdot, 0_R)$ gelten die folgenden Aussagen:*

(i) $0_R \cdot a = 0_R = a \cdot 0_R$ *für alle* $a \in R$

(ii) $a(-b) = -ab = (-a) \cdot b$ *für alle* $a, b \in R$.

Ist R unitär, so gilt

(iii) $-b = (-1_R)b$.

Beweis.
(i) $0_R \cdot a = (0_R + 0_R) \cdot a = 0_R a + 0_R a$. Nach Kürzen folgt $0_R = 0_R \cdot a$. Analoger Beweis für $a \cdot 0_R = 0_R$.

(ii) $0_R = a \cdot 0_R = a(b + (-b)) = ab + a(-b)$ also $-ab = a(-b)$. Die andere Aussage beweist man analog.
Ist R unitär, so setzt man in (ii) $a = 1_R$, und erhält (iii). □

Beispiel. $R = \{0\}$ mit den einzig möglichen Verknüpfungen $+$ und $\cdot$ heißt der *Nullring*. Der Nullring ist ein kommutativer Ring mit 1. (Es gilt $0_R = 0 = 1_R$). Dies ist der einzige Ring mit 1 in dem $1_R = 0_R$ gilt.
Grund: Gilt $0_R = 1_R$, so gilt für jedes $r \in R$: $r = 1_R r = 0_R r = 0_R$, d.h. R besteht aus genau einem Element.

Lemma 2.8. *Es sei $R = (R, +, \cdot, 0_R, 1_R)$ ein Ring mit 1 und $R^\times \subset R$ die Menge aller Elemente in R die sowohl ein Links- als auch ein Rechtsinverses bzgl. Multiplikation haben, d.h.*

$$R^\times = \{r \in R \mid \exists s, t \in R : sr = 1_R = rt\}$$

Dann ist $\{R^\times, \cdot, 1_R\}$ eine Gruppe. Man nennt $R^\times$ die Einheitengruppe *von R.*

Beweis. Wir müssen zunächst zeigen, dass die Multiplikation nicht aus $R^\times$ hinausführt. Seien also $r, r' \in R^\times$ und s, s', t, t' mit $sr = 1 = rt$, $s'r' = 1 = r't'$. Dann gilt

$$(s's)(rr') = s'(sr)r' = s'1r' = s'r' = 1 \quad \text{und} \quad (rr')(t't) = r(r't')t = r1t = rt = 1.$$

Daher gilt $rr' \in R^\times$. Wir weisen nun die Gruppenaxiome (G1)–(G3) nach. G1, also die Assoziativität der Multiplikation, folgt aus den Axiomen für Ringe (R2). $1 \in R^\times$ ist ein neutrales Element, also gilt G2. Bleibt zu zeigen, dass für $r \in R^\times$ ein $r' \in R^\times$ mit $r'r = 1$ existiert. Nach Definition existiert ein $s \in R$ mit $sr = 1$ und wir müssen einsehen, dass $s \in R^\times$ gilt. s hat offensichtlich ein Rechtsinverses, nämlich r. Aber r ist auch linksinvers zu s. Dies sieht man so. Wähle $t \in R$ mit $rt = 1$. Dann gilt:

$$s = s(rt) = (sr)t = t.$$

Hieraus folgt: $rs = rt = 1$. □

Bemerkungen.

1. Im Beweis haben wir gesehen, dass für $r \in R^\times$ das Links- und das Rechtsinverse übereinstimmen. Dies folgt auch aus Lemma 2.5.

2. Es gibt Ringe mit 1 mit Elementen, die ein Rechtsinverses, aber kein Linksinverses besitzen (oder umgekehrt). Diese Elemente sind dann keine Einheiten.

3. Angenommen es gilt $0_R \in R^\times$. Dann existiert ein $r \in R$ mit $0_R r = 1_R$ und es folgt $0_R = 0_R r = 1_R$, d. h. R ist der Nullring.

Wir werden es zunächst nur mit kommutativen Ringen zu tun haben. Besonders einfache Ringe sind Körper. (siehe nächster Abschnitt)

2.3 Körper

Definition 2.9. Ein Körper (field) K ist ein kommutativer Ring mit 1 $(K, +, 0_K, 1_K)$, in dem gilt:

$$K^\times = K \setminus \{0_R\}.$$

In Worten: K ist nicht der Nullring (sonst wäre $0_K \in K^\times$) und jedes von Null verschiedene Element besitzt ein Inverses bzgl. Multiplikation.

Beispiele.

1. $\mathbb{Q}$, $\mathbb{R}$ mit den üblichen Operationen sind Körper.

2. $\mathbb{Z}$ ist kein Körper (nur 1 und -1 haben ein Inverses).

Lemma 2.10. *In einem Körper K gilt*

$$a \cdot b = 0_K \Longrightarrow (a = 0_K \text{ oder } b = 0_K)$$

Beweis. Angenommen $a \neq 0_K$. Dann existiert ein $a^{-1} \in K$ mit $a^{-1}a = 1_K$. Es folgt $b = 1_K \cdot b = a^{-1}ab = a^{-1}0_K = 0_K$. □

Bemerkung. Ein kommutativer Ring R mit 1, der die Eigenschaft aus Lemma 2.10 erfüllt, nennt man auch nullteilerfrei oder Integritätsbereich (integral domain). $\mathbb{Z}$ ist nullteilerfreier Ring, aber es gibt auch Ringe, die diese Eigenschaft nicht haben (siehe unten).

Wir wollen nun zeigen, dass der Restklassenring $\mathbb{Z}/p\mathbb{Z}$ ein Körper ist, falls p eine Primzahl ist.

Definition 2.11. Eine Primzahl ist ein $p \in \mathbb{Z}$ mit $p \notin \mathbb{Z}^\times$, so dass für alle $a, b \in \mathbb{Z}$ gilt:

$$p \mid a \cdot b \Rightarrow (p \mid a) \vee (p \mid b)$$

Zusätzlich fordern wir $p > 0$.

Bemerkung. Jede Primzahl in $\mathbb{Z}$ ist irreduzibel: $p \notin \mathbb{Z}^\times$ und für alle $a, b \in \mathbb{Z}$ gilt:

$$ab = p \Rightarrow (a \in \mathbb{Z}^\times) \vee (b \in \mathbb{Z}^\times)$$

Aus der Schule bekannt: Jedes irreduzible $p \in \mathbb{Z}$ mit $p > 0$ ist auch eine Primzahl (äquivalent zur Existenz einer Primfaktorzerlegung). (Die Definition von *irreduzibel* und *prim* ist in jedem Integritätsbereich sinnvoll, aber nicht immer sind irreduzible Elemente auch prim.)

Lemma 2.12. *Ist p eine Primzahl, so ist $\mathbb{Z}/p\mathbb{Z}$ ein Körper.*

Beweis. $\mathbb{Z}/p\mathbb{Z}$ ist ein kommutativer Ring mit 1. Es genügt zu zeigen, dass jede von 0 verschiedene Restklasse ein Inverses bzgl. Multiplikation besitzt. In anderen Worten ist zu zeigen:
Behauptung: Für jedes $A \in \mathbb{Z}/p\mathbb{Z}$, $A \neq \overline{0}$, ist die $\overline{1}$ im Bild der Abbildung

$$A\cdot\colon \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}, \qquad B \mapsto A \cdot B.$$

Wir zeigen sogar, dass die Abbildung surjektiv ist. Da $\mathbb{Z}/p\mathbb{Z}$ endlich ist, genügt es nach Lemma 1.23 zu zeigen, dass die Abbildung injektiv ist.
Angenommen es gäbe Restklassen $B, C \in \mathbb{Z}/p\mathbb{Z}$ mit $AB = AC$. Zu zeigen: $B = C$. Seien $a, b, c \in \mathbb{Z}$ Vertreter von A, B und C. Wegen $A \neq 0$ gilt $p \nmid a$. Wegen $AB = AC$ gilt $ab \equiv ac \bmod p \Rightarrow a(b-c) \equiv 0 \bmod p \Rightarrow p \mid a(b-c)$.
Weil p Primzahl ist und $p \nmid a$ folgt $p \mid (b-c)$, also $b \equiv c \bmod p$ also $B = C$. Also ist die Abbildung: $A\cdot\colon \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ injektiv, also surjektiv, also liegt $\overline{1}$ im Bild. □

Lemma 2.13. *Ist $n \in \mathbb{N}$ keine Primzahl, so ist $\mathbb{Z}/n\mathbb{Z}$ kein Körper.*

Beweis. Für $n = 1$ ist $\mathbb{Z}/n\mathbb{Z}$ der Nullring. Nun sei $n > 1$ und keine Primzahl. Dann existieren $a, b \in \mathbb{N}$ mit $1 < a, b < n$ und $ab = n$. Für die Restklassen bedeutet dies $\bar{a} \neq \bar{0}$, $\bar{b} \neq \bar{0}$ aber $\bar{a}\bar{b} = \overline{ab} = \bar{n} = \bar{0}$. Dies steht im Widerspruch zu Lemma 2.10. Also ist $\mathbb{Z}/n\mathbb{Z}$ kein Körper. □

Wie man am Beispiel $\mathbb{Z}/p\mathbb{Z}$ sieht, kann es in einem Körper passieren, dass $\underbrace{1_K + \cdots + 1_K}_{p-\text{mal}} = 0_K$ gilt.

Definition 2.14. Sei K ein Körper. Die kleinste natürliche Zahl n mit

$$\underbrace{1_K + \cdots + 1_K}_{n-\text{mal}} = 0_K \text{ (in } K)$$

heißt die *Charakteristik* von K. Notation $\operatorname{char}(K)$. Gibt es eine solche Zahl nicht, setzt man $\operatorname{char}(K) = 0$.

Bemerkung.

1. Offensichtlich gilt immer $\mathrm{char}(K) = 0$ oder $\mathrm{char}(K) \geq 2$ (wegen $1_K \neq 0_K$).
2. die Körper $\mathbb{Q}$, $\mathbb{R}$ haben die Charakteristik 0.
3. $\mathbb{Z}/p\mathbb{Z}$ hat die Charakteristik p.

Satz 2.15. *Die Charakteristik eines Körpers ist entweder gleich* 0 *oder eine Primzahl.*

Beweis. Sei $\mathrm{char}(K) \neq 0$, also $\mathrm{char}(K) = n \geq 2$. Wäre n keine Primzahl, so gäbe es $a, b \in \mathbb{N}$, $1 < a, b < n$ mit $ab = n$. Dann gilt

$$\underbrace{(1_K + \cdots + 1_K)}_{a-\text{mal}} \cdot \underbrace{(1_K + \cdots + 1_K)}_{b-\text{mal}} = \underbrace{(1_K + \cdots + 1_K)}_{n-\text{mal}} = 0.$$

Nach Lemma 2.10 ist entweder $\underbrace{(1_K + \cdots + 1_K)}_{a-\text{mal}}$ oder $\underbrace{(1_K + \cdots + 1_K)}_{b-\text{mal}}$ gleich 0_K, im Widerspruch zur Minimalität von n. □

Bemerkung. Für uns wird (außer recht spät) die Charakteristik keine Rolle spielen. Solange es keine Arbeit macht, werden wir jede Annahme an die Charakteristik vermeiden.

2.4 Homomorphismen

Homomorphismen = strukturerhaltende Abbildungen.

Definition 2.16. Seien $(G, *_G, e_G)$ und $(H, *_H, e_H)$ Gruppen. Eine Abbildung $f\colon G \to H$ heißt *Gruppenhomomorphismus*, wenn für alle $g, g' \in G$ gilt

$$f(g *_G g') = f(g) *_H f(g').$$

Sind $(R, +_R, \cdot_R, 0_R)$ und $(S, +_S, \cdot_S, 0_S)$ Ringe, so heißt eine Abbildung $f\colon R \to S$ *Ringhomomorphismus*, wenn für alle $a, b \in R$ gilt

$$f(a +_R b) = f(a) +_S f(b),\ f(a \cdot_R b) = f(a) \cdot_S f(b).$$

Ein Ringhomomorphismus $f : R \to S$ von Ringen mit 1 $(R, +_R, \cdot_R, 0_R, 1_R)$ und $(S, +_S, \cdot_S, 0_S, 1_S)$ heißt unitär (oder Homomorphismus von Ringen mit 1), wenn zusätzlich gilt: $f(1_R) = 1_S$.
Eine Abbildung von Körpern heißt *Körperhomomorphismus*, wenn sie ein unitärer Ringhomomorphismus ist.

Definition 2.17. Ein Gruppen-(Ring-, Körper-)homomorphismus heißt *injektiv* (auch *Monomorphismus*) bzw. *surjektiv* (auch *Epimorphismus*) wenn er als Mengenabbildung injektiv bzw. surjektiv ist. Er heißt Gruppen-(Ring-, Körper-) *Isomorphismus* wenn er *bijektiv*, d.h. injektiv und surjektiv ist.

Bemerkung. Die inverse Abbildung f^{-1} zu einem Gruppen-(Ring-, Körper-)Isomorphismus ist wieder ein Gruppen-(Ring-, Körper-)Isomorphismus. Zwei Gruppen (Ringe, Körper) heißen *isomorph*, wenn es einen Isomorphismus zwischen ihnen gibt.

Bemerkung. Es ist hilfreich, mathematische Strukturen (Objekte) mit ihren strukturerhaltenden Abbildungen (Morphismen) selbst wieder zu Objekten des mathematischen Denkens zusammenzufassen. Diese Objekte nennt man Kategorien, etwa

- Kategorie der Mengen: Objekte: Mengen, Morphismen: Mengenabbildungen,
- Kategorie der Gruppen: Objekte: Gruppen, Morphismen: Gruppenhomomorphismen,
- Kategorie der Ringe: Objekte: Ringe, Morphismen: Ringhomomorphismen.

Eine interessante mathematische Fragestellung ist immer, die Objekte einer Kategorie bis auf Isomorphie (deutsch: Gleichgestaltigkeit) zu klassifizieren, d. h. eine Möglichkeit zu finden, sämtliche Isomorphietypen von Objekten der Kategorie aufzulisten. Für die Kategorie der (endlichen) Mengen ist der Isomorphietyp einer Menge durch die Kardinalität der Menge eindeutig bestimmt. Das Klassifikationsproblem für die Kategorie der Gruppen oder Ringe ist deutlich schwieriger.

Lemma 2.18. *Sei $f\colon (G, *_G, e_G) \to (H, *_H, e_H)$ ein Gruppenhomomorphismus. Dann gilt*

(i) $f(e_G) = e_H$,

(ii) $f(g^{-1}) = f(g)^{-1}$ *für alle* $g \in G$.

Beweis.
(i) Es gilt $e_G *_G e_G = e_G$, also

$$f(e_G) *_H f(e_G) = f(e_G *_G e_G) = f(e_G) = f(e_G) *_H e_H.$$

Kürzen ergibt $f(e_G) = e_H$.
(ii) $e_H \overset{(i)}{=} f(e_G) = f(g * g^{-1}) = f(g) \cdot f(g^{-1})$. Daher gilt $f(g)^{-1} = f(g^{-1})$. □

Beispiele.

- Ist $(G, *, e)$ eine Gruppe, so ist die Identität $\mathrm{id}\colon G \to G$ ein Gruppenisomorphismus.
- Sind $(G, *_G, e_G)$, $(H, *_H, e_H)$ Gruppen, so ist der *triviale* Homomorphismus $f\colon G \to H$, $f(g) = e_H$ für alle $g \in G$ ein Gruppenhomomorphismus. Er ist genau dann injektiv, wenn $G = \{e_G\}$ gilt und genau dann surjektiv, wenn $H = \{e_H\}$ gilt.

- Sei $n \in \mathbb{N}$. Die Restklassenabbildung (kanonische Projektion)

$$\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}\ ,\ a \longmapsto \overline{a}$$

  ist ein surjektiver, unitärer Ringhomomorphismus.

- Die Inklusion $\mathbb{Q} \hookrightarrow \mathbb{R}$ ist ein Körperhomomorphismus.

- Die Exponentialabbildung

$$\begin{array}{ccc} (\mathbb{R},+,0) & \longrightarrow & (\mathbb{R}_{>0},\cdot,1) \\ t & \longmapsto & e^t \end{array}$$

  ist ein Gruppenisomorphismus.

- Sei $n \in \mathbb{N}$. Die Abbildung

$$S_n \longrightarrow S_{n+1}, \begin{pmatrix} 1 & \dots & n \\ \pi(1) & \dots & \pi(n) \end{pmatrix} \longmapsto \begin{pmatrix} 1 & \dots & n, & n+1 \\ \pi(1) & \dots & \pi(n), & n+1 \end{pmatrix}$$

  ist ein injektiver Gruppenhomomorphismus.

Definition 2.19. Eine Teilmenge H einer Gruppe $G = (G, *, e)$ heißt *Untergruppe*, wenn sie mit der von G ererbten Struktur eine Gruppe ist, d.h.

(i) $e \in H$

(ii) $h, h' \in H \Rightarrow h * h' \in H$

(iii) $h \in H \Rightarrow h^{-1} \in H$.

Lemma 2.20. *Sei H eine Untergruppe von G. Die Relation*

$$g \sim_H g' \Longleftrightarrow g^{-1} * g' \in H$$

ist eine Äquivalenzrelation auf G.

Beweis.
Refl.: $g \sim_H g$ weil $g^{-1} * g = e \in H$.
Symm.: $g \sim_H g' \Rightarrow g^{-1} * (g') \in H \Rightarrow (g')^{-1} * g = (g^{-1} * g')^{-1} \in H \Rightarrow g' \sim_H g$
Trans.: $g \sim g'$ und $g' \sim g'' \Rightarrow g^{-1} * g'' = g^{-1} * g' * (g')^{-1} * g'' \in H \Rightarrow g \sim_H g''$ □

Bemerkung. Die Äquivalenzklasse eines Elements $g \in G$ besteht aus allen $g' \in G$ der Form $g*h$ mit $h \in H$. Bezeichnung: gH. Die Menge aller Äquivalenzklassen wird mit G/H bezeichnet (die "Linksnebenklassen zu H").

Definition 2.21. Sei $f\colon G \to H$ ein Gruppenhomomorphismus. Der *Kern* von f ist die Teilmenge

$$\ker(f) = \{g \in G \mid f(g) = e_H\}$$

Lemma 2.22. *Sei $f\colon G \to H$ ein Gruppenhomomorphismus.*
(i) $\ker(f)$ *ist eine Untergruppe von* G
(ii) $\mathrm{im}(f)$ *ist eine Untergruppe von* H
(iii) f *ist injektiv* $\iff \ker(f) = \{e_G\}$
(iv) f *ist surjektiv* $\iff \mathrm{im}(f) = H$.

Beweis.
(i) $f(e_G) = e_H \Rightarrow e_G \in \ker(f)$.
$g, g' \in \ker(f) \Rightarrow f(g * g') = f(g) * f(g') = e_H$ also $g * g' \in \ker(f)$
aus 2.18 (ii) folgt für $g \in \ker(f)$, dass $f(g^{-1}) = f(g)^{-1} = e_H^{-1} = e_H$, also $g^{-1} \in \ker(f)$.

(ii) $f(e_G) = e_H \Rightarrow e_H \in \mathrm{im}(f)$.
Seien $h, h' \in \mathrm{im}(f)$ und $g, g' \in G$ mit $f(g) = h$, $f(g') = h'$. Dann gilt $f(g * g') = f(g) * f(g') = h * h'$, also $h * h' \in \mathrm{im}(f)$.
Ist $h = f(g) \in \mathrm{im}(f)$, so gilt $h^{-1} = f(g^{-1}) \in \mathrm{im}(f)$.

(iii) ($\Longrightarrow$) Sei f injektiv. Für $g \in \ker(f)$ gilt $f(e_G) = e_H = f(g) \Rightarrow g = e_G$, also $g = e_G$, d.h. $\ker(f) = \{e_G\}$
($\Longleftarrow$) Sei nun $\ker(f) = \{e_G\}$ und $g, g' \in G$ mit $f(g) = f(g')$. Dann gilt

$$f(g * (g')^{-1}) = f(g) * f(g')^{-1} = e_H,$$

also $g * (g')^{-1} \in \ker(f) = \{e_G\}$. Es folgt $g = g'$.

(iv) ist trivial. □

Bemerkung. Jeder Gruppenhomomorphismus $f\colon G \to H$ induziert einen surjektiven Gruppenhomomorphismus $F\colon G \twoheadrightarrow \mathrm{im}(f)$ durch $F(g) = f(g) \in \mathrm{im}(f)$.

Lemma 2.23. *Sei G eine* kommutative *Gruppe und $H \subset G$ eine Untergruppe.*

(i) die Menge G/H der Linksnebenklassen zu H wird durch die Verknüpfung

$$(gH)(g'H) = (gg')H$$

zu einer kommutativen Gruppe.

(ii) Die kanonische Projektion $p\colon G \to G/H$ ist ein surjektiver Gruppenhomomorphismus und es gilt

$$\ker(p) = H.$$

Bemerkung. G/H heißt die *Faktorgruppe* von G nach H.

Beweis.
(i) Wir müssen nachweisen, dass die Verknüpfung wohldefiniert ist, d.h. gilt $g_1H = g_2H$ und $g_1'H = g_2'H$, so folgt $(g_1g_1') = (g_2g_2')H$.

Wir wissen: $g_1^{-1}g_2 \in H$ und $(g_1')^{-1}g_2' \in H$. Es ist H Untergruppe in G und G ist kommutativ. Also gilt

$$\begin{aligned}(g_1g_1')^{-1}(g_2g_2') &= ((g_1')^{-1}g_1^{-1})(g_2g_2') \\ &= ((g_1')^{-1}g_2')(g_1^{-1}g_2) \in H\end{aligned}$$

Die Gültigkeit der Gruppenaxiome wird von G ererbt, z.B. gilt $e_{G/H} = e_GH$.

(ii) Die kanonische Projektionen ist immer surjektiv. Dass p ein Homomorphismus ist, folgt direkt aus der Definition der Verknüpfung auf G/H. Schließlich gilt

$$\begin{aligned}\ker(p) &= \{g \in G \mid p(g) = e_{G/H}\} \\ &= \{g \in G \mid g \sim_H e_G\} \\ &= \{g \in G \mid g^{-1}e_G \in H\} = \{g \in G \mid g \in H\} = H.\end{aligned}$$

□

Bemerkung. Ist G nicht kommutativ, so ist die Verknüpfung auf G/H nur unter bestimmten Bedingungen an H wohldefiniert.

Definition 2.24. Sei $(R, +_R, 0_R, \cdot_R)$ ein Ring und $S \subset R$ eine Teilmenge. S heißt *Unterring* (oder *Teilring*), wenn S mit den von R ererbten Strukturen ein Ring ist, d. h.

- $0_R \in S$ und $(S, +_R, 0_R)$ ist eine Untergruppe von $(R, +_R, 0_R)$,
- mit $s_1, s_2 \in S$ liegt $s_1 \cdot s_2$ in S.

Ist R unitär, so heißt S *unitärer Unterring* von R wenn S ein Unterring ist, unitär ist, und es gilt $1_S = 1_R$.

Beispiele.

- $\mathbb{Z}$ ist ein unitärer Unterring von $\mathbb{Q}$.
- $2\mathbb{Z} = \{a \in \mathbb{Z} \mid a \text{ ist gerade}\}$ ist ein (nicht-unitärer) Unterring in $\mathbb{Z}$,
- $\mathbb{R} \times \mathbb{R}$ mit komponentenweiser Addition und Multiplikation ist ein unitärer Ring. $\mathbb{R} \times \{0\} \subset \mathbb{R} \times \mathbb{R}$ ist ein Unterring, ist unitär, aber kein unitärer Unterring weil: $1_{\mathbb{R}\times\mathbb{R}} = (1, 1)$, aber $1_{\mathbb{R}\times\{0\}} = (1, 0)$.

Definition 2.25. Ein *Unterkörper* eines Körpers ist ein unitärer Teilring, der selbst Körper ist.

Beispiel. $\mathbb{Q}$ ist Teilkörper von $\mathbb{R}$

Lemma 2.26. *Ist*

$$f\colon (R, +_R, 0_R, \cdot_R) \longrightarrow (S, +_S, 0_S, \cdot_S)$$

ein Ringhomomorphismus, so gilt

$$f(0_R) = 0_S,\ f(-a) = -f(a),\ a \in R.$$

Beweis. Der Ringhomomorphismus f induziert einen Homomorphismus der "unterliegenden" Gruppen $f\colon (R, +_R, 0_R) \to (S, +_S, 0_S)$. Das Ergebnis folgt aus 2.18. □

Bemerkungen.

- $\ker(f)$ ist ein Unterring in R, der i. A. nicht unitär ist, auch wenn R unitär ist.
- $\mathrm{im}(f)$ ist ein Unterring in S. Sind R und S unitär, und f ein unitärer Ringhomomorphismus, so ist $\mathrm{im}(f)$ ein unitärer Teilring.

Lemma 2.27. *Sei $f : (R, +_R, 0_R, \cdot_R, 1_R) \longrightarrow (S, +_S, 0_S, \cdot_S, 1_S)$ ein unitärer Ringhomomorphismus. Dann gilt $f(R^\times) \subset S^\times$ und die induzierte Abbildung*

$$(R^\times, \cdot, 1_R) \longrightarrow (S^\times, \cdot, 1_S)$$

zwischen den Einheitengruppen ist ein Gruppenhomomorphismus.

Beweis. Sei $r \in R^\times$ und $s \in R^\times$ sei (Rechts-, wie Links-) Inverses. Dann gilt

$$f(s)f(r) = f(sr) = f(1_R) = 1_S, \quad \text{und} \quad f(r)f(s) = f(rs) = f(1_R) = 1_S.$$

Also gilt $f(r) \in S^\times$. Die induzierte Abbildung $R^\times \to S^\times$ ist ein Gruppenhomomorphismus, weil f ein unitärer Ringhomomorphismus ist. □

Satz 2.28. *Seien $K = (K, +_K, \cdot_K, 0_K, 1_K)$ und $L = (L, +_L, \cdot_L, 0_L, 1_L)$ Körper und $f\colon K \to L$ ein Körperhomomorphismus. Dann gilt*

(i) *f ist injektiv,*

(ii) *$\mathrm{char}(K) = \mathrm{char}(L)$,*

(iii) *$\mathrm{im}(f)$ ist ein Teilkörper von L.*

Beweis.
(i) Nach 2.22 genügt es zu zeigen, dass $\ker(f) = \{0\}$ gilt. Sei $a \in \ker(f),\ a \neq 0$. Dann existiert $a^{-1} \in K$ und es gilt

$$1_L = f(1_K) = f(a \cdot a^{-1}) = f(a) \cdot f(a^{-1}) = 0_L.$$

Dieser Widerspruch zeigt, dass ein solches a nicht existiert, also $\ker(f) = \{0\}$.
(ii) Aus (i) folgt

$$\underbrace{1_K + \cdots + 1_K}_{n-\text{mal}} = 0_K \iff f(1_K) + \cdots + f(1_K) = 0_L$$
$$\iff \underbrace{1_L + \cdots + 1_L}_{n-\text{mal}} = 0_L.$$

Direkt nach der Definition von Charakteristik folgt $\text{char}(K) = \text{char}(L)$.
(iii) $\text{im}(f)$ ist ein unitärer Teilring in L. Zu zeigen: ist $y \in \text{im}(f)$, $y \neq 0$, so gilt $y^{-1} \in \text{im}(f)$. Sei nun $y = f(x)$. Nach (i) folgt $x \neq 0$. Daher liegt y in der Untergruppe $f(K^\times) \subset L^\times$ und das gleiche gilt für y^{-1}.

□

Bemerkung. Die induzierte Abbildung $F\colon K \to f(K)$, $x \mapsto f(x) \in f(K)$, ist also ein Körperisomorphismus und man identifiziert K mit $f(K)$.
Sprechweise: Der Körper K ist über f in L eingebettet. I.A. kann es mehrere Einbettungen von K nach L geben (!)

Sind $f\colon G_1 \to G_2$ und $g\colon G_2 \to G_3$ Gruppenhomomorphismen, so auch die Verknüpfung

$$g \circ f\colon G_1 \longrightarrow G_3.$$

Gleiches gilt für Ring- und Körperhomomorphismen. Spezialfall $G_1 = G_2 = G_3$.

Definition 2.29. Sei G eine Gruppe. Ein Gruppenhomomorphismus $f\colon G \to G$ heißt *Gruppenendomorphismus*. Ist f bijektiv, so heißt f *Gruppenautomorphismus*.
Die analoge Sprechweise benutzt man für Ringe und Körper.
Bezeichnung: $\text{End}(G)$, $\text{End}(R)$, $\text{End}(K)$ bzw. $\text{Aut}(G)$, $\text{Aut}(R)$, $\text{Aut}(K)$.

Lemma 2.30. *Sei G eine Gruppe (R ein Ring, K ein Körper). Dann ist $\text{Aut}(G)(\text{Aut}(R), \text{Aut}(K))$ mit der Verknüpfung*

$$\begin{array}{ccc} \text{Aut}(G) \times \text{Aut}(G) & \longrightarrow & \text{Aut}(G) \\ (g, f) & \longmapsto & g \circ f \end{array}$$

(analog für $\text{Aut}(R)$, $\text{Aut}(K)$) eine Gruppe.

Beweis. Wohldefiniertheit: Mit f und g ist auch $g \circ f$ bijektiv.
- Assoziativität: $h \circ (g \circ f) = h \circ g \circ f = (h \circ g) \circ f$
- neutrales Element: id_G
- Inverses. Mit f ist auch f^{-1} ein Gruppenautomorphismus und es gilt

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

□

Bemerkung. Ist $R = (R, +_R, \cdot_R, 0_R)$ ein Ring, so muß man zwischen den Gruppen $\text{Aut}(R, +_R, \cdot_R, 0_R)$ (Ringautomorphismen) und $\text{Aut}(R, +_R, 0_R)$ (Gruppenautomorphismen) unterscheiden. Die erste Gruppe ist eine Untergruppe der zweiten.

2.5 Die komplexen Zahlen

Die komplexen Zahlen bilden den kleinsten Körper, der die reellen Zahlen umfasst und in dem -1 ein Quadrat ist.

Definition 2.31. Der Körper der komplexen Zahlen $\mathbb{C} = (\mathbb{C}, +_{\mathbb{C}}, \cdot_{\mathbb{C}}, 0_{\mathbb{C}}, 1_{\mathbb{C}})$ ist wie folgt gegeben:

- $\mathbb{C} = \mathbb{R} \times \mathbb{R} = \{(a, b) \mid a, b \in \mathbb{R}\}$,
- $(a, b) +_{\mathbb{C}} (a', b') = (a + a', b + b')$ für $a, a', b, b' \in \mathbb{R}$,
- $(a, b) \cdot_{\mathbb{C}} (a', b') = (aa' - bb', ab' + a'b)$ für $a, a', b, b' \in \mathbb{R}$.
- $0_{\mathbb{C}} = (0, 0)$, $1_{\mathbb{C}} = (1, 0)$.

Lemma 2.32. $\mathbb{C}$ *ist ein Körper.*

Beweis.

- $\mathbb{R} \times \mathbb{R}$ ist mit der komponentenweisen Addition eine Gruppe.
- Kommutativgesetz der Multiplikation: klar.
- Distributivgesetz: Nachrechnen:

$$\begin{aligned}(a, b) \cdot_{\mathbb{C}} (a' + a'', b' + b'') &= (a(a' + a'') - b(b' + b''), a(b' + b'') + (a' + a'')b) \\ &= (aa' - bb', ab' + a'b) +_{\mathbb{C}} (aa'' - bb'', ab'' + a''b).\end{aligned}$$

  für $a, a', a'', b, b', b'' \in \mathbb{R}$.

- $(1, 0) \cdot_{\mathbb{C}} (a, b) = (1a - 0b, 1b + 0a) = (a, b)$ für $a, b \in \mathbb{R}$.
- Existenz des Inversen: Sei $(a, b) \in \mathbb{C}, (a, b) \neq 0_{\mathbb{C}}$. Dann gilt

$$(a, b)^{-1} = (\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}).$$

□

Bemerkung. Die Abbildung $\mathbb{R} \to \mathbb{C}, \quad a \mapsto (a, 0)$ ist ein Körperhomomorphismus, insbesondere also injektiv. Wir identifizieren $\mathbb{R}$ mit seinem Bild unter dieser Abbildung. Damit wird $\mathbb{R}$ zu einem Unterkörper von $\mathbb{C}$.

Schreibweise: $i := (0, 1)$ (*imaginäre Einheit*). Es gilt $i^2 = -1$ und

$$(a, b) = (a, 0) + (0, b) = a + bi \text{ für } (a, b) \in \mathbb{C}$$

Beispiele.

- $(3+2i)(2-i) = (6+2) + (4-3)i = 8+i$,
- $(1+i)^{-1} = \frac{1}{2}(1-i)$.

Definition 2.33. Für $\lambda = a + bi \in \mathbb{C}$ nennt man

1. $\operatorname{Re}\lambda := a \in \mathbb{R}$ den *Realteil* von λ,
2. $\operatorname{Im}\lambda := b \in \mathbb{R}$ den *Imaginärteil* von λ,
3. $\bar{\lambda} := a - bi \in \mathbb{C}$ die zu λ *konjugiert komplexe Zahl*,
4. $|\lambda| = \sqrt{\lambda\bar{\lambda}} = \sqrt{a^2+b^2} \in \mathbb{R}$ den *Absolutbetrag* von λ.

Bemerkung. $a^2+b^2 \geq 0$, so dass die Quadratwurzel $\sqrt{a^2+b^2}$ als nicht negative reelle Zahl x mit $x^2 = a^2+b^2$ wohldefiniert ist.

Lemma 2.34. *Für alle $\lambda, \mu \in \mathbb{C}$ gilt:*

1. $\lambda + \bar{\lambda} = 2\operatorname{Re}\lambda$
2. $\frac{1}{i}(\lambda - \bar{\lambda}) = 2\operatorname{Im}\lambda$.
3. $\overline{\lambda+\mu} = \bar{\lambda} + \bar{\mu}$,
4. $\overline{\lambda\cdot\mu} = \bar{\lambda}\cdot\bar{\mu}$,
5. $\lambda \in \mathbb{R} \Leftrightarrow \lambda = \bar{\lambda}$,
6. $|\lambda\cdot\mu| = |\lambda|\cdot|\mu|$

Beweis. Einfaches Nachrechnen. □

Lemma 2.35. *Für alle $\lambda, \mu \in \mathbb{C}$ gilt $|\lambda+\mu| \leq |\lambda| + |\mu|$ (Dreiecksungleichung).*

Beweis. Für alle $\tau = a + bi \in \mathbb{C}$ gilt

$$(\operatorname{Re}\tau)^2 = a^2 \leq a^2+b^2 = |\tau|^2$$

und somit

$$\operatorname{Re}\tau \leq |\operatorname{Re}\tau| \leq |\tau|.$$

Setze $\tau = \lambda\bar{\mu}$. Dann gilt $\overline{\lambda\bar{\mu}} = \bar{\lambda}\mu$ und somit

$$\lambda\bar{\mu} + \bar{\lambda}\mu = 2\operatorname{Re}\tau \leq 2|\lambda\bar{\mu}| = 2|\lambda||\mu|,$$

also

$$|\lambda+\mu|^2 = (\lambda+\mu)(\bar{\lambda}+\bar{\mu}) = |\lambda|^2 + |\mu|^2 + \lambda\bar{\mu} + \bar{\lambda}\mu \leq (|\lambda|+|\mu|)^2$$

und somit die Behauptung. □

3 Vektorräume und lineare Abbildungen

3.1 Definitionen

Sei R ein unitärer Ring.

Definition 3.1. Ein (unitärer Links-)*Modul* über R ist eine abelsche Gruppe $(M, +_M, 0_M)$ mit einer Operation

$$R \times M \to M, \quad (a, m) \mapsto a \cdot m,$$

so dass für alle $a, b \in R$, $v, w \in M$ gilt:

(M1) $a \cdot (b \cdot v) = (a \cdot b) \cdot v$

(M2) $(a + b) \cdot v = a \cdot v + b \cdot v$

(M3) $a \cdot (v + w) = a \cdot v + a \cdot w$

(M4) $1_R \cdot v = v$.

Bemerkungen.

(i) (M2) und (M3) sind so zu interpretieren, dass Punkt- vor Strichrechnung geht.

(ii) Schreibweise: av statt $a \cdot v$.

Definition 3.2. Ein Modul über einem Körper K heißt *K-Vektorraum*.

Beispiele.

1. $\{0\}$ mit der offensichtlichen (und einzig möglichen) Operation ist ein K-Vektorraum.
2. $(K, +_K, 0_K)$ mit der Operation $K \times K \to K$, $(a, v) \mapsto av$ ist ein K-Vektorraum.
3. $K^n = \underbrace{K \times \cdots \times K}_{n-\text{mal}}$ wird zum K-Vektorraum durch

   $$(v_1, \ldots, v_n) + (w_1, \ldots, w_n) = (v_1 + w_1, \ldots, v_n + w_n)$$

   und $a(v_1, \ldots, v_n) = (av_1, \ldots, av_n)$.
4. $\mathbb{C}$ ist ein $\mathbb{R}$-Vektorraum.
   Allgemeiner: Ist L ein Körper und $K \subset L$ ein Teilkörper, so ist L ein K-Vektorraum.

5. Die Menge $C^n(\mathbb{R},\mathbb{R})$ der n-mal stetig differenzierbaren reellwertigen Funktionen auf $\mathbb{R}$ $(0 \leq n \leq \infty)$ ist ein $\mathbb{R}$-Vektorraum durch

$$\begin{aligned} &\text{Addition:} && (f_1+f_2)(x) = f_1(x)+f_2(x).\\ &\text{Skalarmultiplikation:} && (af)(x) = af(x) \end{aligned}$$

Vereinbarung: Von jetzt an sei K ein fixierter Körper, den wir manchmal von der Notation ausschließen.

Lemma 3.3. *Sei V ein K-Vektorraum. Dann gilt für alle $v \in V$, $a \in K$*

(i) $0_K \cdot v = 0_V$,

(ii) $(-1)v = -v$,

(iii) $a \cdot 0_V = 0_V$.

Beweis.
(i) $0_V + 0_K \cdot v = 0_K \cdot v = (0_K + 0_K)v = 0_K \cdot v + 0_K \cdot v$. Jetzt kürzen.
(ii) $0_V = 0_K \cdot v = (1_K + (-1)_K) \cdot v = v + (-1_K)v$.
(iii) $a \cdot 0_V = a \cdot (0_V + 0_V) = a \cdot 0_V + a \cdot 0_V$. □

Definition 3.4. Es seien V, W K-Vektorräume. Ein Gruppenhomomorphismus

$$f\colon V \to W$$

heißt *(K-)lineare Abbildung* oder *(K-)Vektorraumhomomorphismus*, wenn

$$f(ax) = a \cdot f(x)$$

für alle $x \in V$ gilt. Eine lineare Abbildung heißt Vektorraum-Monomorphismus, Vektorraum-Epimorphismus bzw. Vektorraum-Isomorphismus, wenn sie injektiv, surjektiv bzw. bijektiv ist. Die Menge der linearen Abbildungen von V nach W wird mit $\mathrm{Hom}_K(V,W)$ bezeichnet.

Bemerkung. Sei K fixiert. Die K-Vektorräume (Objekte) bilden zusammen mit den K-linearen Abbildungen (Morphismen) eine Kategorie.

Weitere Notationen:

$$\begin{aligned} \mathrm{End}_K(V) &= \mathrm{Hom}_K(V,V)\\ \mathrm{Gl}(V) &= \mathrm{Aut}_K(V) = \{\varphi \in \mathrm{End}_K(V),\ |\ \varphi \text{ ist Iso.}\} \end{aligned}$$

Beispiele linearer Abbildungen.

1. $K^n \to K^1 = K,\ (a_1,\ldots,a_n) \mapsto a_1$.

2. $\mathbb{C} \to \mathbb{C}$, $\lambda \mapsto \bar{\lambda}$ ist $\mathbb{R}$-linear, aber nicht $\mathbb{C}$-linear.

3. $K^n \to K^{2n}$, $(a_1, \ldots, a_n) \mapsto (a_1, \ldots, a_n, a_1, \ldots, a_n)$.

4. $C^0(\mathbb{R}, \mathbb{R}) \to \mathbb{R}$, $f \mapsto \int_0^1 f(x)dx$.

5. $n \geq 1$:
$$\begin{array}{rcl} C^n(\mathbb{R}, \mathbb{R}) & \longrightarrow & C^{n-1}(\mathbb{R}, \mathbb{R}) \\ f & \longmapsto & f' = \frac{df}{dx} \quad \text{(Ableitung)}. \end{array}$$

Definition 3.5. Eine Teilmenge V eines Vektorraumes W heißt *Untervektorraum*, wenn sie mit den von W ererbten Strukturen ein Vektorraum ist, d.h.

(i) V ist Untergruppe von W,

(ii) $v \in V \Rightarrow a \cdot v \in V$ für alle $a \in K$.

Lemma 3.6. *Seien $f : V \to W$ eine (K-)lineare Abbildung. Dann gilt:*

(i) $\ker(f) \subset V$ *ist Untervektorraum,*

(ii) $\mathrm{im}(f) \subset W$ *ist Untervektorraum.*

Beweis. $\ker(f)$ und $\mathrm{im}(f)$ sind Untergruppen nach 2.22. Für $v \in \ker(f)$ gilt $f(av) = af(v) = a \cdot 0 = 0$, also $av \in \ker(f)$ für alle $a \in K$. Für $w = f(v) \in \mathrm{im}(f)$ gilt $a \cdot w = a \cdot f(v) = f(a \cdot v) \in \mathrm{im}(f)$. □

3.2 Konstruktionen auf Vektorräumen

Seien U, V K-Vektorräume und M eine Menge.

Abbildungen in Vektorräume

$\mathrm{Abb}(M, V)$ wird zum Vektorraum durch

$$(f + g)(m) = f(m) + f(m), \quad (af)(m) = a(f(m))$$

für $m \in M$, $f, g \in \mathrm{Abb}(M, V)$, $a \in K$. Neutrales Element: $e(m) = 0_V$ für alle $m \in M$. ("Nullabbildung")
Bez.: $0 \in \mathrm{Abb}(M, V)$

$\mathrm{Hom}_K(U, V) \subset \mathrm{Abb}(U, V)$ ist ein Untervektorraum, weil:

- 0 ist eine lineare Abbildung,

- f, g linear $\Rightarrow f + g$ linear,

- $a \in K$, f linear $\Rightarrow af$ linear.

Spezialfall: $V = K$
$\mathrm{Hom}_K(U, K) =: U^*$ heißt der *Dualraum* zu U, seine Elemente heißen *Linearformen* auf U.
Ist $f\colon U \to V$ eine lineare Abbildung, so ist die *duale Abbildung*

$$f^*\colon V^* \longrightarrow U^*,\ \varphi \longmapsto \varphi \circ f$$

linear. Die Abbildung

$$\begin{array}{rccc} *\colon & \mathrm{Hom}_K(U,V) & \longrightarrow & \mathrm{Hom}_K(V^*,U^*) \\ & f & \longmapsto & f^* \end{array}$$

ist linear.
Wir haben eine kanonische lineare Abbildung

$$U \longrightarrow (U^*)^*, \qquad u \longmapsto \phi_u.$$

Dabei bezeichnet für $u \in U$

$$\phi_u\colon U^* \longrightarrow K, \qquad f \longmapsto f(u)$$

die *Auswertungsabbildung.*

Direktes Produkt, direkte Summe.

Das kartesische (direkte) Produkt $U \times V$ wird durch

$$(u, v) + (u', v') = (u + u', v + v'), \quad a(u, v) = (au, av)$$

($a \in K$, $u, u' \in U$, $v, v' \in V$) zu einem K-Vektorraum.
Alternative Bezeichnung: $U \oplus V$ (die *d*irekte Summe).

Schnitt und Summe von Unterräumen.

Seien $U_1, U_2 \subset V$ Untervektorräume. Dann sind

- der Schnitt $U_1 \cap U_2$
- die Summe $U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1,\ u_2 \in U_2\}$

Untervektorräume von V.

Lemma 3.7. *Die natürliche Abbildung*

$$\varphi\colon U_1 \oplus U_2 \to U_1 + U_2, \quad (u_1, u_2) \longmapsto u_1 + u_2$$

ist surjektiv. Gilt $U_1 \cap U_2 = \{0\}$, *so ist* φ *ein Isomorphismus.*

Beweis.
Beh: φ ist linear.
Bew: Seien $u_1, v_1 \in U_1,\ u_2, v_2 \in U_2$ und $a \in K$:

$$\begin{aligned}\varphi((u_1,u_2)+(v_1,v_2)) &= \varphi((u_1+v_1,u_2+v_2))\\ &= u_1+v_1+u_2+v_2 = (u_1+u_2)+(v_1+v_2)\\ &= \varphi((u_1,u_2))+\varphi((v_1,v_2)).\\ \varphi(a(u_1,u_2)) &= \varphi((au_1,au_2))\\ &= au_1+au_2 = a(u_1+u_2) = a\varphi((u_1,u_2)).\end{aligned}$$

Beh: φ ist surjektiv.
Bew: folgt aus der Definition von U_1+U_2.

Sei $U_1 \cap U_2 = \{0\}$. Beh: φ ist injektiv.
Bew: Sei $(u_1,u_2) \in \ker(\varphi)$. Dann gilt $u_1+u_2=0 \Rightarrow u_1 = -u_2$. Folglich $u_1 \in U_2$, $u_2 \in U_1$, also $u_1, u_2 \in U_1 \cap U_2 = \{0\} \Rightarrow (u_1,u_2) = (0,0)$. Daher gilt $\ker(\phi) = \{0\}$ und die Injektivität von φ folgt aus 2.22 (iii). □

Faktorvektorraum

Sei $U \subset V$ ein Untervektorraum. Die Faktorgruppe

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

der Restklassen

$$v+U = \{v+u \mid u \in U\}$$

von V modulo U wird ein K-Vektorraum durch

$$a \cdot (v+U) = a \cdot v + U.$$

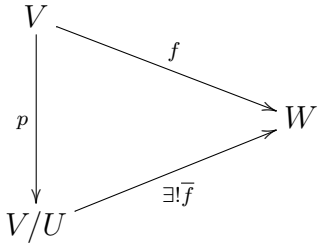
für $a \in K$, $v+U \in V/U$.

Unabhängigkeit von der Auswahl des Repräsentanten $v \in v+U$:
Ist $v+U = v'+U$, so gilt $v-v' \in U$.
Folglich: $av - av' = a(v-v') \in U$ und daher $av+U = av'+U$

V/U heißt der *Faktorvektorraum*. Die kanonische Projektion $p\colon V \to V/U$ ist linear.

Satz 3.8. *(Universelle Eigenschaft des Faktorraums). Sei $U \subset V$ ein Untervektorraum und $p\colon V \to V/U$ die kanonische Projektion. Zu jeder linearen Abbildung $f\colon V \to W$ mit $U \subset \ker(f)$ gibt es eine eindeutig bestimmte lineare Abbildung*

$\overline{f}\colon V/U \to W$ mit $f = \overline{f} \circ p$. Bild:



Beweis.
Existenz: Definiere $\overline{f}(v+U) = f(v)$.
Wohldefiniertheit: Gilt $v + U = v' + U$, so gilt $v - v' \in U \subset \ker(f)$. Daher gilt

$$\begin{aligned} f(v) &= f(v) - f(v - v') \\ &= f(v - v + v') = f(v'). \end{aligned}$$

Eindeutigkeit: Seien $\overline{f_1}$ und $\overline{f_2}$ zwei solche Abbildungen und $v+U \in V/U$ beliebig.
Beh: $\overline{f}_1(v+U) = \overline{f}_2(v+U)$.
Bew: Wegen $f(v) = \overline{f}_1(p(v)) = \overline{f}_1(v+U)$ und $f(v) = \overline{f}_2(p(v)) = \overline{f}_2(v+U)$ gilt $\overline{f}_1(v+U) = \overline{f}_2(v+U)$. □

Korollar 3.9. *Seien $U \subset V$ Vektorräume und W ein weiterer Vektorraum. Dann gibt es einen natürlichen Isomorphismus von Vektorräumen*

$$F\colon \{\varphi \in \mathrm{Hom}_K(V,W) \mid U \subset \ker(\varphi)\} \xrightarrow{\sim} \mathrm{Hom}_K(V/U, W)$$

Beweis. Die Abbildung F ist durch die Universaleigenschaft des Faktorraums gegeben (3.8), d.h. $F(\varphi) =: \psi$ ist die eindeutig bestimmte Abbildung mit $\psi(v + U) = \varphi(v) \in W$. Dass F linear ist, folgt direkt aus der Definition.
Um zu zeigen, dass F ein Isomorphismus ist, genügt es, eine Umkehrabbildung anzugeben. Sei $p\colon V \to V/U$ die kanonische Projektion und für $\psi\colon V/U \to W$ setze $G(\psi) = \psi \circ p : V \to W$. Dann gilt $F \circ G(\psi) = \psi$ und $G \circ F(\varphi) = \varphi$ für alle φ und ψ. □

Korollar 3.10. *Sei $f\colon V \to W$ eine K-lineare Abbildung. Dann gibt es einen natürlichen Isomorphismus*

$$\Big(W/\operatorname{im}(f)\Big)^* \cong \ker(f^*).$$

Beweis.

$$\Big(W/\operatorname{im}(f)\Big)^* = \mathrm{Hom}_K(W/\operatorname{im}(f), K)$$

und

$$\begin{aligned} \ker f^* &= \{\varphi\colon W \longrightarrow K,\ f^*(\varphi) = 0\} \\ &= \{\varphi\colon W \longrightarrow K,\ \varphi \circ f = 0\} \\ &= \{\varphi\colon W \longrightarrow K \mid \operatorname{im}(f) \subset \ker \varphi\} \end{aligned}$$

Die Aussage folgt aus 3.9 (mit $U = \operatorname{im}(f)$ und $W = K$). □

Satz 3.11. *(Homomorphiesatz für lineare Abbildungen). Seien V, W Vektorräume und $f\colon V \to W$ eine lineare Abbildung. Dann gibt es einen natürlichen Vektorraumisomorphismus*

$$F\colon V/\ker(f) \xrightarrow{\sim} \operatorname{im}(f)$$

mit der Eigenschaft $f = i \circ F \circ p$. Hier bezeichnet $p\colon V \to V/\ker(f)$ die kanonische Projektion und $i\colon \operatorname{im}(f) \to W$ die Inklusion.

Beweis. Nach 3.8 erhalten wir eine Abbildung

$$\overline{f}\colon V/\ker(f) \to W \text{ mit } \overline{f}(v + \ker(f)) = f(v),$$

d. h. $f = \overline{f} \circ p$.

$\overline{f}$ ist injektiv, da

$$\overline{f}(v + \ker(f)) = 0 \Rightarrow f(v) = 0 \Rightarrow v \in \ker(f) \Rightarrow v + \ker(f) = 0 + \ker(f)$$

Das Bild von $\overline{f}$ ist gleich $\operatorname{im}(f)$ (klar).

Damit können wir $\overline{f}$ in der Form $\overline{f} = i \circ F$ mit $F\colon V/\ker(f) \to \operatorname{im}(f)$ schreiben. Die Abbildungen i und $\overline{f}$ sind injektiv, also auch F. Außerdem ist F nach Konstruktion surjektiv. Daher ist F ein Isomorphismus und es gilt

$$f = \overline{f} \circ p = i \circ F \circ p.$$

□

Konstruktionen für unendliche Familien

Sei $(U_i)_{i \in I}$ eine Familie von Vektorräumen. Das kartesische Produkt

$$\prod_{i \in I} U_i$$

wird (analog zum Produkt zweier Vektorräume) durch komponentenweise Addition und Skalarmultiplikation zu einem Vektorraum.

Notation: Ist I eine Indexmenge und sind $(a_i)_{i \in I}$ Objekte die durch I induziert sind, so sagt man, dass eine Eigenschaft "für fast alle $i \in I$" erfüllt ist, wenn es eine endliche Teilmenge $J \subset I$ gibt, so dass a_i die Eigenschaft für alle $i \in I \setminus J$ hat.

Der Untervektorraum

$$\bigoplus_{i \in I} U_i := \{(u_i)_{i \in I} \mid u_i = 0 \quad \text{für fast alle} \quad i\}$$

des kartesischen Produkts heißt die *direkte Summe* der Vektorräume U_i. Ist I selbst eine endliche Menge so gilt

$$\bigoplus_{i\in I} U_i = \prod_{i\in I} U_i.$$

Sei nun $(U_i)_{i\in I}$ eine Familie von Untervektorräumen eines Vektorraums V. Dann haben wir die Untervektorräume

$$\bigcap_{i\in I} U_i = \{u \mid u \in U_i \quad \text{für alle} \quad i\}$$

$$\sum_{i\in I} U_i = \left\{\sum_{i\in I} u_i \mid u_i \in U_i \text{ für alle } i, \text{ und } u_i = 0 \text{ f. f. a. } i.\right\}$$

Wegen der Bedingung $u_i = 0$ f. f. a. i ist die scheinbar unendliche Summe nur eine endliche Summe und darum überhaupt erst definiert.
Ist die Indexmenge I endlich und nicht leer, nimmt man sich typischerweise eine bijektive Abbildung $I \xrightarrow{\sim} \{1, \ldots, n\}$ und schreibt

$$\bigoplus_{i\in I} U_i = \bigoplus_{i=1}^{n} U_i$$

und analog für die anderen Konstruktionen.

Konvention:

$$\bigoplus_{i\in\emptyset} U_i = \prod_{i\in\emptyset} U_i = \{0\}$$

Das einzige Element $0 = \emptyset$ wird auch als *leere Familie* bezeichnet. Für Untervektorräume in V:

$$\sum_{i\in\emptyset} U_i = \{0 = \sum_{i\in\emptyset} v_i\}, \quad \bigcap_{i\in\emptyset} U_i = V.$$

Endomorphismenring

Wir haben gesehen, dass $\mathrm{Hom}_K(V, W)$ wieder eine Vektorraumstruktur trägt. Insbesondere ist es eine abelsche Gruppe bzgl. $+$. Ist $V = W$, so definieren wir auf $\mathrm{Hom}_K(V, V) = \mathrm{End}_K(V)$ eine Multiplikation durch $\circ$ (Komposition).

Lemma 3.12. *Mit diesen Operationen ist* $(\mathrm{End}_K(V), +, \circ, 0, \mathrm{id}_V)$ *ein unitärer Ring. Die Abbildung*

$$K \to \mathrm{End}_K(V),\ a \longmapsto a \cdot \mathrm{id}_V$$

ist ein Ringhomomorphismus. Durch die Operation

$$\begin{aligned} \mathrm{End}_K(V) \times V &\longrightarrow V \\ (f, v) &\longmapsto f(v) \end{aligned}$$

wird V zu einem (unitären, Links-) $\mathrm{End}_K(V)$*-Modul .*

Beweis. Wir verifizieren die Ringaxiome für $(\mathrm{End}_K(V), +, \circ, 0, \mathrm{id}_V)$. Zunächst ist $\circ$ assoziativ. Weiter gilt $g \circ (f_1 + f_2) = g \circ f_1 + g \circ f_2$ für $g, f_1, f_2 \in \mathrm{End}_K(V)$ wegen

$$\begin{aligned}(g \circ (f_1 + f_2))(v) &= g((f_1 + f_2)(v)) = g(f_1(v) + f_2(v)) = g(f_1(v)) + g(f_2(v)) \\ &= g \circ f_1(v) + g \circ f_2(v) = (g \circ f_1 + g \circ f_2)(v)\end{aligned}$$

für $v \in V$.

Analog $(g_1 + g_2) \circ f = g_1 \circ f + g_2 \circ f$.

Also ist $\mathrm{End}_K(V)$ ein Ring, in dem id_V offenbar ein 1-Element ist. Dass die Abbildung $K \to \mathrm{End}_K(V),\ a \mapsto a \cdot \mathrm{id}_V$ ein Ringhomomorphismus ist, liest man leicht an den Definitionen ab. Die gegebene Operation macht V zu einem $\mathrm{End}_K(V)$-Modul weil:

(M1) $g \cdot (f \cdot v) = g(f(v)) = (g \circ f)(v) = (g \cdot f)(v)$.

(M2) $(f + g) \cdot v = (f + g)(v) = f(v) + g(v) = f \cdot v + g \cdot v$

(M3)

$$\begin{aligned}f \cdot (v + w) &= f(v + w) = f(v) + f(w) \\ &= f \cdot v + f \cdot w.\end{aligned}$$

(M4) $1_{\mathrm{End}_K(V)} \cdot v = \mathrm{id}_V(v) = v$.

□

3.3 Basen

Ziel: Klassifikation aller Vektorräume bis auf Isomorphie.
Zentrales Hilfsmittel: Begriff der Basis.

Erinnerung: Eine über eine Indexmenge I indizierte Familie $(m_i)_{i \in I}$ von Elementen einer Menge M ist nichts weiter als eine Abbildung $m\colon I \to M$ und wir schreiben $m(i) = m_i \in M$ und $m = (m_i)_{i \in I} \in M^I$.

Sprechweise: $(m_i)_{i \in I}$ ist ein *System von Elementen in M*.

Definition 3.13. Ein System von Skalaren $(\alpha_i)_{i \in I} \in K^I$ heißt *endlich*, wenn $\alpha_i = 0$ für fast alle i gilt. Der Untervektorraum aller endlichen Systeme von Skalaren wird mit

$$K^{(I)} = \bigoplus_{i \in I} K$$

bezeichnet.

Bemerkung. Sei V ein K-Vektorraum, $(v_i)_{i\in I} \in V^I$ ein System von Vektoren in V und $(\alpha_i)_{i\in I} \in K^{(I)}$ ein endliches System von Skalaren. Dann gilt $\alpha_i v_i = 0$ für fast alle i, so dass man der Summe

$$\sum_{i\in I} \alpha_i v_i$$

einen Sinn geben kann.

Definition 3.14. Sei V ein K-Vektorraum, I eine Indexmenge und $v = (v_i)_{i\in I}$ ein System von Vektoren in V. Der Untervektorraum

$$\mathrm{Lin}((v_i)_{i\in I}) = \left\{ \sum_{i\in I} \alpha_i v_i \mid (\alpha_i)_{i\in I} \in K^{(I)} \subset V \right\}$$

heißt die *lineare Hülle* des Systems $(v_i)_{i\in I}$. Jeder Vektor $v \in \mathrm{Lin}((v_i)_{i\in I})$ heißt *Linearkombination* der v_i. Man nennt $\mathrm{Lin}((v_i)_{i\in I})$ auch den von den Vektoren $(v_i)_{i\in I}$ *aufgespannten* Untervektorraum.

Lemma 3.15. *Die lineare Hülle* $\mathrm{Lin}((v_i)_{i\in I})$ *ist der kleinste Untervektorraum von* V*, der für alle* $i \in I$ *den Vektor* v_i *enthält.*

Beweis. Sei $U \subset V$ ein Untervektorraum mit $v_i \in U$ für alle $i \in I$. Dann gilt für alle $(\alpha_i)_{i\in I} \in K^{(I)}$:

$$\sum_{i\in I} \alpha_i v_i \in U$$

und somit $\mathrm{Lin}((v_i)_{i\in I}) \subset U$. □

Definition 3.16. Sei $(v_i)_{i\in I}$ ein System von Vektoren eines Vektorraums V.

(i) $(v_i)_{i\in I}$ heißt *Erzeugendensystem* von V, wenn $\mathrm{Lin}((v_i)_{i\in I}) = V$ gilt

(ii) $(v_i)_{i\in I}$ heißt *linear unabhängig*, wenn für jedes endliche System von Skalaren $(\alpha_i)_{i\in I} \in K^{(I)}$ die Implikation

$$\sum_{i\in I} \alpha_i v_i = 0 \implies \alpha_i = 0 \text{ für alle } i$$

gilt.

(iii) $(v_i)_{i\in I}$ heißt *Basis* von V, wenn es zu jedem Vektor $v \in V$ ein eindeutig bestimmtes endliches System von Skalaren $(\alpha_i)_{i\in I} \in K^{(I)}$ mit $v = \sum_{i\in I} \alpha_i v_i$ gibt.

Beispiel. Für jeden Vektorraum V ist $(v)_{v\in V} \in V^V$ ein Erzeugendensystem von V.

Beispiel. Im K^n bilden die Vektoren

$$e_1 = (1, 0, \ldots, 0),\ e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 1)$$

eine Basis. Diese heißt die *kanonische Basis* des K^n. Der Vektor e_i, $(i = 1, \ldots, n)$, heißt der i-te *Einheitsvektor*.

Lemma 3.17. *Der K-Vektorraum V habe die endliche Basis $(v_1, \ldots, v_n)$. Dann ist die Abbildung*

$$\begin{array}{rcl} \phi : K^n & \longrightarrow & V \\ (\alpha_1, \ldots, \alpha_n) & \longmapsto & \sum\limits_{i=1}^{n} \alpha_i v_i \end{array}$$

ein Vektorraumisomorphismus.

Beweis. Zunächst ist ϕ linear.
Definition von Basis: Zu jedem $v \in V$ gibt es genau ein $\alpha = (\alpha_1, \ldots, \alpha_n) \in K^n$ mit

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

d. h. ϕ ist bijektiv. □

Bemerkung. Im Moment wissen wir noch nicht, ob für $n \neq m$ eventuell ein Isomorphismus $K^n \cong K^m$ existieren könnte.

Definition 3.18. Ein Vektorraum V heißt *endlich erzeugt*, wenn es ein endliches Erzeugendensystem $(v_1, \ldots, v_n)$ von V gibt.

Beispiel. K^n ist endlich erzeugt.

Bemerkung. Im Moment wissen wir noch nicht, ob ein Untervektorraum eines endlich erzeugten Vektorraums wieder endlich erzeugt ist.

Wir wollen Charakterisierungen der Eigenschaft, Basis zu sein, herleiten.

Lemma 3.19. *Ein System von Vektoren $(v_i)_{i \in I}$ ist genau dann eine Basis von V, wenn es ein Erzeugendensystem und linear unabhängig ist.*

Beweis.
($\Rightarrow$) Sei $(v_i)_{i \in I}$ eine Basis. Da jeder Vektor als Linearkombination darstellbar ist, gilt $\mathrm{Lin}((v_i)_{i \in I}) = V$, d.h. $(v_i)_{i \in I}$ ist ein Erzeugendensystem. Ist nun $(\alpha_i)_{i \in I} \in K^{(I)}$ ein endliches System von Skalaren mit $\sum\limits_{i \in I} \alpha_i v_i = 0$, so gilt wegen $\sum\limits_{i \in I} 0 \cdot v_i = 0$ und der Eindeutigkeit der Darstellung: $\alpha_i = 0$ für alle $i \in I$.

($\Leftarrow$) Sei nun $(v_i)_{i \in I}$ ein Erzeugendensystem. Dann ist jeder Vektor Linearkombination der $(v_i)_{i \in I}$.

Z. z.: Ist das System $(v_i)_{i \in I}$ linear unabhängig, so ist die Darstellung jedes Vektors $v \in V$ als Linearkombination der v_i eindeutig.

Bew.: Seien $(\alpha_i), (\beta_i) \in K^{(I)}$ endliche Familien und $\sum \alpha_i v_i = v = \sum \beta_i v_i$. Dann ist die Familie $(\alpha_i - \beta_i)_{i \in I}$ auch endlich und es gilt:

$$\sum_I (\alpha_i - \beta_i) v_i = 0 \quad \Rightarrow \quad \alpha_i - \beta_i = 0 \text{ für alle } i.$$

Hieraus folgt $\alpha_i = \beta_i$ für alle i. □

Definition 3.20. Ein Erzeugendensystem $(v_i)_{i \in I}$ eines Vektorraums V heißt *minimal*, wenn für jede echte Teilmenge $J \subsetneq I$ das System $(v_i)_{i \in J}$ kein Erzeugendensystem ist.

Beispiel. Das Erzeugendensystem $(e_1, \ldots, e_n)$ des K^n ist minimal. Läßt man den i-ten Einheitsvektor weg, so kann man nur noch Elemente $(\alpha_1, \ldots, \alpha_n) \in K^n$ mit $\alpha_i = 0$ als Linearkombination erhalten. Wegen $1 \neq 0$ in K fehlt also z.B. der Vektor $(0, \ldots, 1, \ldots, 0)$ (die 1 steht an i-ter Stelle).

Lemma 3.21. *Ein Erzeugendensystem ist genau dann minimal, wenn es eine Basis ist.*

Beweis.
($\Leftarrow$) Sei $(v_i)_{i \in I}$ eine Basis und $J \subsetneq I$ eine echte Teilmenge. Wähle ein $i_0 \in I \setminus J$. Trivialerweise gilt $v_{i_0} = 1 \cdot v_{i_0}$. Wegen der Eindeutigkeit der Darstellung, läßt sich also v_{i_0} nicht als Linearkombination der v_j, $j \in J$, schreiben, und deshalb ist $(v_i)_{i \in J}$ kein Erzeugendensystem. Folglich ist $(v_i)_{i \in I}$ ein minimales Erzeugendensystem.
($\Rightarrow$) Sei nun $(v_i)_{i \in I} \in V^I$ ein minimales Erzeugendensystem. Nach 3.19 müssen wir zeigen, dass $(v_i)_{i \in I}$ linear unabhängig ist. Angenommen es gäbe ein von 0 verschiedenes endliches System $(\alpha_i) \in K^{(I)}$ mit

$$\sum_{i \in I} \alpha_i v_i = 0.$$

Sei $i_0 \in I$ mit $\alpha_{i_0} \neq 0$. Dann gilt

$$-\alpha_{i_0} v_{i_0} = \sum_{i \in I \setminus \{i_0\}} \alpha_i v_i,$$

also

$$v_{i_0} = \sum_{i \in I \setminus \{i_0\}} -\frac{\alpha_i}{\alpha_{i_0}} v_i.$$

Beh: $(v_i)_{i \in I \setminus \{i_0\}}$ ist auch ein Erzeugendensystem.

Bew.: Sei $v \in V$. Dann existiert eine endliche Familie $\beta_i \in K^{(I)}$ mit $v = \sum_{i \in I} \beta_i v_i$. Nun gilt

$$v = \beta_{i_0} v_{i_0} + \sum_{i \in I \setminus \{i_0\}} \beta_i v_i = \sum_{i \in I \setminus \{i_0\}} \left(-\frac{\beta_{i_0} \cdot \alpha_i}{\alpha_{i_0}} + \beta_i \right) v_i.$$

Dies zeigt die Behauptung und wir erhalten einen Widerspruch zur Minimalität des Systems $(v_i)_{i \in I}$. □

Notation: Sind $(v_i)_{i\in I} \in V^I$ und $(w_i)_{i\in J} \in V^J$ zwei Systeme von Vektoren eines Vektorraums V, so bezeichnet man das System

$$(u_i)_{i\in I\dot\cup J} \in V^{I\dot\cup J} \quad \text{mit } u_i = \begin{cases} v_i & i \in I \\ w_i & i \in J \end{cases}.$$

als die *Vereinigung* der Systeme $(v_i)_{i\in I}$ und $(w_i)_{i\in J}$.

Definition 3.22. Ein linear unabhängiges System von Vektoren $(v_i)_{i\in I} \in V^I$ heißt *maximal*, wenn für jeden Vektor $v \in V$ das System $(v_i)_{i\in I\dot\cup\{*\}}$ mit $v_i = v_i$ für $i \in I$ und $v_* = v$ nicht linear unabhängig ist.

Lemma 3.23. *Ein linear unabhängiges System ist genau dann maximal, wenn es eine Basis ist.*

Beweis.
($\Leftarrow$) Sei $(v_i)_{i\in I}$ eine Basis und $v \in V$ beliebig. Dann existiert $(\alpha_i)_{i\in I} \in K^{(I)}$ mit $v = \sum \alpha_i v_i$. Dies formuliert man zu

$$\sum \alpha_i v_i + (-1)v = 0$$

um und sieht, dass die Vereinigung von $(v_i)_{i\in I}$ mit dem 1-elementigen System (v) nicht linear unabhängig ist. Also ist (v_i) ein maximales linear unabhängiges System.

($\Rightarrow$) Sei (v_i) ein maximales linear unabhängiges System. Nach 3.19 ist zu zeigen, dass (v_i) ein Erzeugendensystem ist. Angenommen nicht. Dann gäbe es ein

$$v \in V \setminus \operatorname{Lin}((v_i)_{i\in I}).$$

Beh.: Das System $(v_i)_{i\in I\dot\cup\{*\}}$ mit $v_i = v_i$ für $i \in I$, $v_* = v$ ist linear unabhängig.

Bew.: Sei $(\alpha_i)_{i\in I\dot\cup\{*\}}$ eine endliche Familie mit $\sum\limits_{i\in I\dot\cup\{*\}} \alpha_i v_i = 0$. Dann gilt $\alpha_* v = -\sum\limits_{i\in I} \alpha_i v_i$. Wegen $v \notin \operatorname{Lin}((v_i)_{i\in I})$ folgt $\alpha_* = 0$. Da das System $(v_i)_{i\in I}$ linear unabhängig ist, folgt $\alpha_i = 0$ für alle $i \in I$. Dies zeigt die Behauptung.

Wegen der Maximalität von $(v_i)_{i\in I}$ erhalten wir einen Widerspruch, also gilt $\operatorname{Lin}((v_i)_{i\in I}) = V$. □

Zusammenfassend:

Theorem 3.24 (Charakterisierung von Basen)**.** *Sei $(v_i)_{i\in I} \in V^I$ ein System von Vektoren in einen Vektorraum V. Dann sind äquivalent:*

(i) $(v_i)_{i\in I}$ ist eine Basis von V,

(ii) $(v_i)_{i\in I}$ ist ein linear unabhängiges Erzeugendensystem von V,

(iii) $(v_i)_{i\in I}$ *ist ein minimales Erzeugendensystem von* V*,*

(iv) $(v_i)_{i\in I}$ *ist ein maximales linear unabhängiges System von* V*.*

Beweis. 3.19, 3.21, 3.23. □

Frage: Hat jeder Vektorraum eine Basis?

Theorem 3.25 (Basisergänzungssatz)**.** *Sei* $(v_i)_{i\in I}$ *ein Erzeugendensystem des Vektorraums* V *und* $I' \subset I$ *eine Teilmenge, so dass das System* $(v_i)_{i\in I'}$ *linear unabhängig ist. Dann gibt es eine Teilmenge* $J \subset I$*, mit* $I' \subset J$ *so dass* $(v_i)_{i\in J}$ *eine Basis ist. Insbesondere besitzt jeder Vektorraum eine Basis und jeder endlich erzeugte Vektorraum besitzt eine endliche Basis.*

Beweis. Wir begründen zunächst das „insbesondere“. Sei $(v_i)_{i\in I}$ ein Erzeugendensystem von V (jeder Vektorraum besitzt ein solches). Setzt man $I' = \varnothing$, so erhält man $J \subset I$ so dass $(v_i)_{i\in J}$ eine Basis von V ist. War I endlich, so ist auch J endlich.
Wir beweisen die Aussage des Satzes hier nur im Fall, dass I endlich ist. Betrachte

$$S = \{M \subset I | I' \subset M,\ (v_i)_{i\in M} \text{ ist linear unabhängig}\} \subset \mathcal{P}(I).$$

Dann besitzt S ein maximales Element bezüglich der Inklusion von Mengen, d. h. ein $J \in S$, für das gilt:

$$\forall M \in S : J \subset M \Rightarrow J = M.$$

Es reicht dazu, $J \in S$ mit $\#J$ maximal zu wählen. Das geht, weil $\#J$ durch $\#I$ beschränkt ist.
Setze $V' = \mathrm{Lin}((v_i)_{i\in J})$.
Beh.: Es gilt $V = V'$. Insbesondere ist $(v_i)_{i\in J}$ ein Erzeugendensystem von V und damit eine Basis von V.
Bew.: Sei $k \in I \setminus J$. Setze $M = J \cup \{k\}$. Dann ist $(v_i)_{i\in M}$ linear abhängig. Sonst wäre $J \subset M \in S$, aber $J \neq M$. Also gibt es $0 \neq (\alpha_i)_{i\in M} \in K^{(}M)$ mit

$$\sum_{i\in M} \alpha_i v_i = 0.$$

Wäre $\alpha_k = 0$, so folgt $(\alpha_i)_{i\in J} = 0$ wegen der linearen Unabhängigkeit von $(v_i)_{i\in J}$ und damit $(\alpha_i)_{i\in M} = 0$. Widerspruch. Also gilt $\alpha_k \neq 0$ und somit

$$v_k = \sum_{i\in J} \frac{\alpha_i}{\alpha_k} v_i \in V'.$$

Also gilt $v_i \in V'$ für alle $i \in I$. Weil $(v_i)_{i\in I}$ ein Erzeugendensystem von V ist, gilt mit Lemma 3.15

$$V = \mathrm{Lin}((v_i)_{i\in I}) \subset V' \subset V,$$

also $V = V'$. □

Bemerkung. Die Endlichkeit von I haben wir nur gebraucht, um ein maximales Element J in S zu finden. Falls I nicht endlich ist, folgt die Existenz eines maximalen Elementes aus dem Lemma von Zorn, siehe z. B. in W. Greub *Lineare Algebra* oder in F. Brieskorn *Lineare Algebra und analytische Geometrie*.

Mit dem Basisergänzungssatz 3.25 und Lemma 3.17 sehen wir, dass jeder endlich erzeugte Vektorraum isomorph zu einem K^n ist. Aber ist das n eindeutig bestimmt? Wichtig ist nun der

Satz 3.26. *Sei $(v_1, \dots, v_n)$ linear unabhänging und $(w_1, \dots, w_m)$ eine Basis von V. Dann gilt $n \leq m$.*

Beweis. Setze

$$A = \{v_1, \dots, v_n\}, \quad B_0 = \{w_1, \dots, w_m\} \subset V$$

In linear unabhängigen Systemen kommt kein Vektor doppelt vor, also gilt:

$$\#A = n, \#B_0 = m$$

Die Systeme $(v_1, \dots, v_n)$, $(w_1, \dots, w_m)$ können wir mit den Systemen $(v)_{v\in A}$ bzw. $(w)_{w\in B_0}$ identifizieren (Das ist nur ein notationeller Trick).

Angenommen, $A \subset B_0$. Dann gilt $n \leq m$ und wir sind fertig.

Beweisidee: Ersetze sukzessive die Basis $(w)_{w\in B_0}$ durch eine Basis $(w)_{w\in B'}$ mit $B' \subset V$, $\#B' \leq \#B_0$ und $\#A \cap B' > \#A \cap B_0$; solange, bis $A \subset B'$.

Angenommen $x \in A \setminus B_0$. Setze $I' = (A \cap B_0) \cup \{x\} \subset B_0 \cup \{x\} = I$. Dann ist $(v)_{v\in I'}$ als Teilsystem von $(v)_{v\in A}$ linear unabhängig. Das System $(w)_{w\in I}$ ist Erzeugendensystem, weil es die Basis $(w)_{w\in B_0}$ als Teilsystem hat. Mit dem Basisergänzungssatz 3.25 erhalten wir ein $I' \subset B_1 \subset I$, so dass $(w)_{w\in B_1}$ Basis ist. Wäre $B_1 = I$, so $B_0 \subsetneq B_1$ und $(w)_{w\in B_0}$ ist kein maximales linear unabhängiges System. Widerspruch. Also $B_1 \subsetneq I$ und $\#B_1 < \#I = \#B_0 + 1$, d. h. $\#B_1 \leq \#B_0$. Andererseits $I' = (A \cap B_0) \cup \{x\} \subset A \cap B_1$, also $\#A \cap B_1 \geq \#I' = \#A \cap B_0 + 1$.

Führe nun dieselbe Betrachtung für B_1 anstelle von B_0 durch, usw. Wir erhalten so eine streng monoton wachsende Folge

$$A \cap B_0 \subsetneq A \cap B_1 \subsetneq A \cap B_2 \subsetneq \dots$$

mit $B_i \subset V$, $\#B_i \leq m$, $(w)_{w\in B_i}$ Basis von V. Weil $\#A \cap B_i$ durch $\#A$ beschränkt ist, muss dieser Prozess terminieren, d. h. es gibt ein $k \in \mathbb{N}_0$ mit $A \setminus B_k = \emptyset$, d. h. $A \subset B_k$. Also gilt $n = \#A \leq B_k \leq \#B_0 = m$. □

Korollar 3.27. *Sei V ein Vektorraum, der eine Basis aus n Vektoren hat. Dann gilt:*

(i) Mehr als n Vektoren sind stets linear abhängig.

(ii) Jede Basis von V besteht aus genau n Vektoren.

(iii) Jedes Erzeugendensystem besteht aus mindestens n Vektoren.

Beweis.
(i): folgt aus 3.26.
(ii): ist $(v_1, \ldots, v_n)$ eine Basis und $(w_1, \ldots, w_m)$ eine andere Basis, so folgt aus 3.26 dass $n \leq m$ und $m \leq n$.
(iii); Wäre $(w_1, \ldots, w_m)$, $m < n$, ein Erzeugendensystem, so gäbe es nach 3.25 eine Basis aus weniger als n Vektoren, was (ii) widerspräche. □

Definition 3.28. Ist V ein endlich erzeugter Vektorraum, so nennt man die Kardinalität einer (jeder) Basis die *Dimension* von V. (*Bezeichnung:* $\dim_K V$ oder einfach $\dim V$). Ist V nicht endlich erzeugt, so setzen wir $\dim V = \infty$.

Bemerkung. Besser: Der Begriff der Kardinalität existiert für jede Menge (i. A. ist das keine natürliche Zahl). Für jeden Vektorraum V hat jede Basis von V dieselbe Kardinalität, so dass man auch allgemein $\dim V$ als die Kardinalität einer Basis definieren kann. Mit dieser Definition gilt dann

$$\dim_K K^{(\mathbb{N})} = \#\mathbb{N} < \#\mathbb{R} = \dim_K K^{(\mathbb{R})}.$$

Für uns haben beide Vektorräume die Dimension ∞.

Beispiele.

- $V = \{0\} \iff \dim V = 0$.
- $\dim K^n = n$

Satz 3.29. *Ist V ein endlich erzeugter Vektorraum und $W \subset V$ ein Untervektorraum, so ist W endlich erzeugt und es gilt $\dim W \leq \dim V$. Die Gleichheit ist äquivalent zu $W = V$.*

Beweis. Sei $n = \dim V$. Wir erhalten eine endliche Basis von W wie folgt.

1. Falls $W = \{0\}$: fertig.
2. Ansonsten: Wähle $w_1 \in W \setminus \{0\}$.
3. (w_1) Basis von W: fertig.
4. Ansonsten ist (w_1) kein maximales linear unabhängiges System und wir finden $w_2 \in W$ mit (w_1, w_2) linear unabhängig.
5. (w_1, w_2) Basis: fertig.
6. Ansonsten: Suche w_3 usw.

Dieser Prozess bricht ab, weil mehr als n Vektoren in V (und damit in W) stets linear abhängig sind. Wir erhalten eine Basis $(w_1, \ldots, w_m)$ von W mit $m \leq n$. Im Fall $m = n$ ist $(w_1, \ldots, w_m)$ ein maximales linear unabhängiges System von Vektoren in V, also eine Basis, daher $W = \mathrm{Lin}(w_1, \ldots, w_n) = V$. □

Konvention: Von jetzt an benutzen wir das Wort endlich-dimensionaler Vektorraum (e. d. VR) für endlich erzeugte Vektorräume.

Lemma 3.30. *Sei $f\colon V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen Vektorräumen. Dann gilt*

(i) ist f injektiv und $(v_1, \ldots, v_n)$ linear unabhängig in V, so ist $(f(v_1), \ldots, f(v_n))$ linear unabhängig in W. Insbesondere gilt $\dim V \leq \dim W$ und Gleichheit gilt dann und nur dann, wenn f ein Isomorphismus ist.

(ii) Ist f surjektiv und $(v_1, \ldots, v_n)$ ein Erzeugendensystem von V, so ist $(f(v_1), \ldots, f(v_n))$ ein Erzeugendensystem von W. Es gilt $\dim V \geq \dim W$ und Gleichheit gilt genau dann, wenn f ein Isomorphismus ist.

Beweis.
(i) Sei f injektiv und $(v_1, \ldots, v_n)$ ein linear unabhängiges System in V. Dann gelten die Implikationen

$$\sum_{i=1}^{n} \alpha_i f(v_i) = 0 \Rightarrow f\left(\sum_{i=1}^{n} \alpha_i v_i\right) = 0 \Rightarrow \sum_{i=1}^{n} \alpha_i v_i = 0 \Rightarrow \alpha_i = 0 \text{ für } i = 1, \ldots, n.$$

Daher ist das System $(f(v_1), \ldots, f(v_n))$ von Vektoren in W linear unabhängig. Ist nun $(v_1, \ldots, v_n)$ eine Basis von V, folgt aus der linearen Unabhängigkeit von $(f(v_1), \ldots, f(v_n))$ und aus 3.29 die Ungleichung $\dim V = n \leq \dim W$.
Da f injektiv ist, induziert f einen Isomorphismus

$$F\colon V \xrightarrow{\sim} \mathrm{im}(f),\ v \mapsto f(v) \in \mathrm{im}(f) \subset W.$$

Wir erhalten $\dim \mathrm{im}(f) = \dim V$. Nun folgt

$$f \text{ Isom.} \Leftrightarrow \mathrm{im}(f) = W \overset{3.29}{\Leftrightarrow} \dim \mathrm{im}(f) = \dim W \Leftrightarrow \dim V = \dim W.$$

(ii) Sei f surjektiv und $(v_1, \ldots, v_n)$ ein Erzeugendensystem von V. Es sei $w = f(v) \in W$ beliebig. Dann existieren $\alpha_1, \ldots, \alpha_n \in K$ mit $v = \sum_{i=1}^{n} \alpha_i v_i$. Es folgt $w = \sum_{i=1}^{n} \alpha_i f(v_i)$. Daher ist $(f(v_1), \ldots, f(v_n))$ ein Erzeugendensystem von W. Wählen wir nun eine Basis $(v_1, \ldots, v_{\dim(V)})$, so ist $(f(v_1), \ldots, f(v_{\dim(V)}))$ ein Erzeugendensystem, also $\dim V \geq \dim W$.

Ist f ein Isomorphismus, so gilt offenbar $\dim V = \dim W$. Umgekehrt gelte $\dim V = \dim W =: n$. Sei $(w_1, \ldots, w_n)$ eine Basis von W und $w_i = f(v_i)$, $i = 1, \ldots, n$. Dann ist $(v_1, \ldots, v_n)$ linear unabhängig (Grund: $\sum \alpha_i v_i = 0 \Rightarrow \sum \alpha_i f(v_i) = 0 \Rightarrow \alpha_i = 0$, $i = 1, \ldots, n$), wegen $n = \dim V$ maximal linear unabhängig, also eine Basis von V. Ist nun $\sum_{i=1}^{n} \alpha_i v_i \in \ker(f)$, so folgt $\sum_{i=1}^{n} \alpha_i w_i = 0$, also $\alpha_i = 0$, $i = 1, \ldots, n$. Daher gilt $\ker(f) = \{0\}$ und f ist ein Isomorphismus. □

Korollar 3.31. *Sei V ein endlich-dimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus. Dann sind äquivalent:*

(i) f ist injektiv

(ii) f ist surjektiv

(iii) f ist ein Isomorphismus.

Beweis. Ist f injektiv, so ist f Isomorphismus nach 3.30 (i). Ist f surjektiv, so ist f Isomorphismus nach 3.30 (ii). Die verbleibenden Implikationen sind trivial. □

Theorem 3.32 (Klassifikation endlichdimensionaler Vektorräume)**.** *Endlichdimensionale K-Vektorräume werden bis auf Isomorphie vollständig durch ihre Dimension klassifiziert:*

(i) Für jeden endlichdimensionalen K-Vektorraum V gibt es ein $n \in \mathbb{N}_0$ mit $V \cong K^n$, nämlich $n = \dim V$.

(ii) Zwei endlichdimensionale K-Vektorräume W und V sind genau dann isomorph, wenn $\dim_K W = \dim_K V$.

(iii) Für jedes $n \in \mathbb{N}_0$ existiert ein endlichdimensionaler Vektorraum V mit $\dim_K V = n$, z. B. $V = K^n$.

Mit anderen Worten: Die Zuordnung $V \mapsto \dim_K V$ induziert eine Bijektion von der Menge der Isomorphieklassen endlichdimensionaler K-Vektorräume nach $\mathbb{N}_0$.

Beweis.
(i) Mit dem Basisergänzungssatz 3.25 und Lemma 3.17 sehen wir, dass V isomorph zu $K^{\dim V}$ ist.
(ii) siehe Lemma 3.30.
(iii) klar. □

3.4 Dimensionsformeln

Wir untersuchen nun, wie sich die Dimension unter den verschiedenen Konstruktionen auf Vektorräumen verhält.

Satz 3.33 (Dimensionsformel für direkte Summen)**.** *Sind U, V Vektorräume, so gilt*

$$\dim U \oplus V = \dim U + \dim V$$

Bemerkung. Per Konvention gilt $\infty + n = \infty = \infty + n$ für $n \in \mathbb{N}_0$, sowie $\infty + \infty = \infty$.

Beweis. Gilt $\dim U = \infty$ oder $\dim V = \infty$, so gilt nach 3.29 auch $\dim U \oplus V = \infty$. Sind $n = \dim U$ und $m = \dim V$ endlich, $(u_1, \ldots, u_n)$ eine Basis von U und $(v_1, \ldots, v_m)$ eine Basis von V, so sieht man unmittelbar, dass

$$((u_1, 0), (u_2, 0), \ldots, (u_n, 0), (0, v_1), \ldots, (0, v_m))$$

eine Basis von $U \oplus V$ ist. Dies zeigt die Dimensionsformel. □

Anwendung:

Definition 3.34. Es sei V ein Vektorraum und $U \subset V$ ein Untervektorraum. Ein Untervektorraum U' von V heißt *Komplement* zu U, wenn gilt

$$U \cap U' = \{0\} \quad \text{und} \quad U + U' = V.$$

Bemerkung. Ist U' ein Komplement zu U, so ist U ein Komplement zu U'.

Satz 3.35. *Sei V ein Vektorraum und $U \subset V$ ein Untervektorraum. Dann existiert ein Komplement U' zu U.*

Beweis. Sei $(u_i)_{i \in I}$ eine Basis von U. Wir ergänzen diese zu einer Basis $(u_i)_{i \in I \dot\cup J}$ von V. Setze $U' = \mathrm{Lin}((u_i)_{i \in J})$. Dann gilt $U \cap U' = \{0\}$, $U + U' = V$. □

Satz 3.36 (Dimensionsformel für Komplemente)**.** *Sei V ein Vektorraum, $U \subset V$ ein Untervektorraum und $U' \subset V$ ein Komplement zu U. Dann ist die natürliche Abbildung*

$$U \oplus U' \longrightarrow V, \quad (u, u') \mapsto u + u'$$

ein Isomorphismus. Insbesondere gilt $\dim V = \dim U + \dim U'$.

Beweis. Der Isomorphismus folgt aus 3.7, die Dimensionsformel aus 3.33. □

Satz 3.37 (Dimension des Faktorraums)**.** *Sei V ein endlich-dimensionaler Vektorraum und $U \subset V$ ein Untervektorraum. Dann ist V/U endlich-dimensional und es gilt*

$$\dim V/U = \dim V - \dim U.$$

Beweis. Sei U' ein Komplement zu U in V. Wir betrachten die zusammengesetzte Abbildung

$$\varphi \colon U' \xrightarrow{i} V \xrightarrow{p} V/U.$$

Behauptung: φ ist ein Isomorphismus.

Beweis der Behauptung: $\ker(\varphi) = \ker(p) \cap U' = U \cap U' = \{0\}$. Also ist φ injektiv. Sei nun $v + U \in V/U$ beliebig. Wegen $U + U' = V$ existiert $u \in U$ und $u' \in U'$ mit $u + u' = v$, also $u' + U = v + U$. Somit ist u' ein Urbild von $v + U$ unter φ. □

Satz 3.38 (Dimensionsformel für lineare Abbildungen)**.** *Ist $f\colon V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen Vektorräumen, so gilt*

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

Beweis. Nach Satz 3.11 gilt $V/\ker(f) \cong \mathrm{im}(f)$. Die Aussage folgt aus 3.37. □

Satz 3.39 (Dimensionsformel für Schnitt und Summe)**.** *Seien U_1, U_2 Untervektorräume des endlich-dimensionalen Vektorraums U Dann gilt*

$$\dim U_1 + \dim U_2 = \dim(U_1 + U_2) + \dim(U_1 \cap U_2)$$

Beweis. Wir betrachten wie im Beweis von Lemma 3.7 die surjektive lineare Abbildung

$$\varphi\colon U_1 \oplus U_2 \longrightarrow U_1 + U_2,\ (u_1, u_2) \longmapsto u_1 + u_2$$

Behauptung: Die Abbildung

$$i\colon U_1 \cap U_2 \longrightarrow \ker(\varphi),\ u \longmapsto (u, -u)$$

ist ein Isomorphismus.

Beweis der Behauptung: Injektivität ist klar. Sei $(u_1, u_2) \in \ker(\varphi)$. Dann gilt $u_1 = -u_2$, also $u_1, u_2 \in U_1 \cap U_2$. Setze $u = u_1$, dann gilt $i(u) = (u_1, u_2)$. Dies zeigt die Surjektivität.

Nach 3.38 gilt nun $\dim(U_1 \oplus U_2) = \dim(U_1 \cap U_2) + \dim(U_1 + U_2)$ und nach 3.33 gilt $\dim(U_1 \oplus U_2) = \dim U_1 + \dim U_2$. □

3.5 Basen und lineare Abbildungen

Sei $f\colon V \to W$ eine lineare Abbildung und $(v_i)_{i\in I}$ eine Basis von V. Wir erhalten das System $(f(v_i))_{i\in I}$ von Vektoren in W. Umgekehrt gilt:

Satz 3.40. *Sei $(v_i)_{i\in I}$ eine Basis von V. Zu jedem System $(w_i)_{i\in I}$ von Vektoren in W gibt es eine eindeutig bestimmte lineare Abbildung $f\colon V \to W$ mit $w_i = f(v_i)$ für alle $i \in I$.*

Beweis. Jedes $v \in V$ hat eine endliche Darstellung $v = \sum\limits_{i\in I} \alpha_i v_i$, $\alpha_i = 0$ f.f.a. i. Wegen der Eindeutigkeit der Darstellung ist die Abbildung

$$f\colon V \longrightarrow W,\ f(v) = \sum_{i\in I} \alpha_i w_i$$

wohldefiniert, linear und hat die gewünschte Eigenschaft. Sind nun f_1, f_2 zwei lineare Abbildungen und $f_1(v_i) = w_i = f_2(v_i)$ für alle $i \in I$, so gilt für $v = \sum \alpha_i v_i$:

$$f_1(v) = \sum_{i\in I} \alpha_i f_1(v_i) = \sum \alpha_i w_i = \sum \alpha_i f_2(v_i) = f_2(\sum \alpha_i v_i) = f_2(v),$$

also $f_1 = f_2$. □

Korollar 3.41. *Für jeden Vektorraum V ist die kanonische Abbildung*

$$\Phi\colon V \to V^{**}, \quad u \mapsto \phi_u, \qquad (\phi_u\colon V^* \to K, \quad f \mapsto f(u))$$

injektiv.

Beweis. Zu zeigen: $\ker(\Phi) = 0$, d.h.: aus $f(u) = 0$ für alle $f \in V^*$ folgt $u = 0$. Sei $u \neq 0$. Nach dem Basisergänzungssatz 3.25 existiert eine Basis $(u_i)_{i\in I}$, die u enthält, d.h. $u_{i_0} = u$ für ein $i_0 \in I$. Wir definieren eine Linearform $f\colon V \to K$ durch $f(u_{i_0}) = 1$ und $f(u_i) = 0$ für $i \neq i_0$. Dann gilt $f(u) \neq 0$. □

Satz 3.42. *Ist V ein endlich-dimensionaler Vektorraum, so gilt* $\dim V = \dim V^*$.

Beweis. Sei $(v_1, \dots, v_n)$ eine Basis von V. Seien $v_1^*, \dots, v_n^* \in V^*$ definiert durch $v_i^*(v_j) = \delta_{ij}$, wobei

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \qquad \text{(Kroneckersymbol)}.$$

Behauptung: $(v_1^*, \dots, v_n^*)$ ist eine Basis von V^*.
Beweis: Sei $\alpha_1 v_1^* + \cdots + \alpha_n v_n^*$ der Nullhomomorphismus $V \to K$. Zu zeigen: $\alpha_1 = \cdots = \alpha_n = 0$. Es gilt

$$\begin{aligned} \alpha_1 &= \alpha_1 + 0 + \cdots + 0 \\ &= \alpha_1 v_1^*(v_1) + \alpha_2 v_2^*(v_1) + \cdots + \alpha_n v_n^*(v_1) \\ &= (\alpha_1 v_1^* + \cdots + \alpha_n v_n^*)(v_1) = 0. \end{aligned}$$

Analog zeigt man $\alpha_2 = \cdots = \alpha_n = 0$, weshalb $(v_1^*, \dots, v_n^*)$ ein linear unabhängiges System in V^* ist. Sei nun $f\colon V \to K$ beliebig und sei $f(v_i) = \alpha_i$, $i = 1, \dots, n$. Dann gilt $f = \alpha_1 v_1^* + \cdots + \alpha_n v_n^*$, weil beide Seiten Linearformen auf V sind, die die gleichen Bilder auf den Basisvektoren $v_1, \dots, v_n$ haben. □

Definition 3.43. Die eben definierte Basis $(v_1^*, \dots, v_n^*)$ von V^* heißt die zur Basis $(v_1, \dots, v_n)$ von V *duale Basis*.

Korollar 3.44. *Ist V ein endlich-dimensionaler Vektorraum, so ist die kanonische Abbildung $\Phi\colon V \to V^{**}$ ein Isomorphismus.*

Beweis. Nach 3.41 ist Φ injektiv. Nach 3.42 gilt

$$\dim V = \dim V^* = \dim V^{**}$$

Wegen 3.30 (i) ist Φ ein Isomorphismus. □

Bemerkung. Für unendlich-dimensionale Vektorräume ist 3.44 falsch!

3.6 Der Rangsatz

Definition 3.45. Seien V, W endlich-dimensionale Vektorräume und $f\colon V \to W$ eine lineare Abbildung. Der *Rang von* f ist definiert durch

$$\mathrm{Rg}(f) = \dim(\mathrm{im}(f)).$$

Bemerkungen.

(i) $f = 0 \iff \mathrm{Rg}(f) = 0$

(ii) f surjektiv $\iff \mathrm{Rg}(f) = \dim W$

(iii) Allgemein gilt die Ungleichung

$$0 \leq \mathrm{Rg}(f) \leq \min(\dim V, \dim W).$$

Nun definiert f die duale Abbildung $f^* : W^* \to V^*$, $\varphi \mapsto \varphi \circ f$. Es gilt der

Satz 3.46 (Rangsatz)**.** *Sei $f\colon V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen Vektorräumen. Dann gilt*

$$\mathrm{Rg}(f) = \mathrm{Rg}(f^*)$$

Beweis. Es gilt für $f : V \to W$:

$$\begin{aligned}
\mathrm{Rg}(f^*) &= \dim(\mathrm{im}(f^*)) \\
&\overset{3.38}{=} \dim(W^*) - \dim(\ker(f^*)) \\
&\overset{3.10}{=} \dim(W^*) - \dim(W/\mathrm{im}(f))^* \\
&\overset{3.42}{=} \dim(W) - \dim(W/\mathrm{im}(f)) \\
&\overset{3.38}{=} \dim(W) - \dim(W) + \dim(\mathrm{im}\, f) = \mathrm{Rg}(f).
\end{aligned}$$

□

4 Matrizen

Im ganzen Kapitel sei K ein fester Körper.
Ziel: Beschreibung aller linearen Abbildungen zwischen zwei endlichdimensionalen Vektorräumen V und W.
Erinnerung: Jeder endlich-dimensionale Vektorraum V über K ist isomorph zum K^n, $n = \dim V$. Der Isomorphismus ist nicht kanonisch, sondern hängt von der Auswahl einer Basis ab (Lemma 3.17).

4.1 Matrizen

Definition 4.1. Eine $m \times n$-*Matrix mit Einträgen in* K ist ein Schema

$$(a_{ij})_{\substack{1\le i\le m\\ 1\le j\le k}} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \text{ mit } a_{ij} \in K \text{ für } 1 \le i \le m,\ 1 \le j \le n.$$

Die Menge der $m \times n$-Matrizen über K wird mit $M_{m,n}(K)$ bezeichnet. $M_{m,n}(K)$ wird zum K-Vektorraum durch

$$\begin{aligned} (a_{ij})_{\substack{1\le i\le m\\ 1\le j\le k}} + (b_{ij})_{\substack{1\le i\le m\\ 1\le j\le k}} &= (a_{ij} + b_{ij})_{\substack{1\le i\le m\\ 1\le j\le k}}, \\ \alpha \cdot (a_{ij})_{\substack{1\le i\le m\\ 1\le j\le k}} &= (\alpha a_{ij})_{\substack{1\le i\le m\\ 1\le j\le k}} \end{aligned}$$

für $(a_{ij})_{\substack{1\le i\le m\\ 1\le j\le k}}, (b_{ij})_{\substack{1\le i\le m\\ 1\le j\le k}} \in M_{m,n}(K)$, $\alpha \in K$.

Notation: Die $m \times n$ Matrix mit 0 an allen Stellen wird mit $0 \in M_{m,n}(K)$ bezeichnet.

Konvention: Wir werden ab jetzt durch

$$(a_1, \dots, a_n) \longmapsto \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

K^n mit $M_{n,1}(K)$ (Spaltenvektoren) identifizieren.

Definition 4.2. Sind $A = (a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}} \in M_{m,n}(K)$ und $B = (b_{ij})_{\substack{1\le i\le n\\ 1\le j\le k}} \in M_{n,k}(K)$ Matrizen, so heißt die Matrix

$$C = (c_{ij})_{\substack{1\le i\le m\\ 1\le j\le k}} \in M_{m,k}(K) \quad \text{mit} \quad c_{ij} = \sum_{s=1}^{n} a_{is} \cdot b_{sj}.$$

das *Produkt* der Matrizen A und B. Schreibweise: $C = A \cdot B$.

Warnung: i.A. ist $B \cdot A$ nicht definiert und selbst im Fall $m = n = k$ gilt i.A. $A \cdot B \neq B \cdot A$!

Beispiel.

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ 5 \end{pmatrix} = (15)$$

Die Multiplikation mit einer festen Matrix $A \in M_{m,n}(K)$ definiert eine Abbildung

$$F_{m,n}(A)\colon K^n \to K^m.$$

Wir erhalten so eine Zuordnung $A \longmapsto F_{m,n}(A)$, die jeder $m \times n$-Matrix eine Abbildung $K^n \to K^m$ zuordnet.

Satz 4.3. *Die eben definierte Zuordnung definiert einen Vektorraum-Isomorphismus*

$$F_{m,n}\colon M_{m,n}(K) \xrightarrow{\sim} \mathrm{Hom}_K(K^n, K^m)$$

Für $A \in M_{m,n}(K)$ und $B \in M_{n,k}(K)$, gilt

$$F_{m,k}(A \cdot B) = F_{m,n}(A) \circ F_{n,k}(B)$$

in $\mathrm{Hom}_K(K^k, K^m)$.

Beweis.
(i) $F_{m,n}(A)$ ist eine lineare Abbildung: Sei $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in K^n$, $\alpha \in K$. Schreibe $A = (a_{ij})$.

$$\begin{aligned} A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + A\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 & +\cdots+ & a_{1n}x_n \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 & +\cdots+ & a_{mn}x_n \end{pmatrix} + \begin{pmatrix} a_{11}y_1 + \cdots + a_{1n}y_n \\ \vdots \\ a_{m1}y_1 + \cdots + a_{mn}y_n \end{pmatrix} \\ &= \begin{pmatrix} a_{11}(x_1+y_1) & +\cdots+ & a_{1n}(x_n+y_n) \\ \vdots & & \vdots \\ a_{m1}(x_1+y_1) & +\cdots+ & a_{mn}(x_n+y_n) \end{pmatrix} \\ &= A\begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} = A\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) \end{aligned}$$

$$\begin{aligned} A\left(\alpha\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) &= A\begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix} = \begin{pmatrix} a_{11}\alpha x_1 & +\cdots+ & a_{1n}\alpha x_1 \\ & \vdots & \\ a_{m1}\alpha x_1 & +\cdots+ & a_{mn}\alpha x_n \end{pmatrix} \\ &= \alpha\begin{pmatrix} a_{11}x_1 & +\cdots+ & a_{1n}x_n \\ \vdots & & \vdots \\ a_{m1}x_1 & +\cdots+ & a_{mn}x_n \end{pmatrix} = \alpha A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{aligned}$$

(ii) $F_{m,n} : M_{m,n}(K) \to \mathrm{Hom}_K(K^n, K^m)$ ist linear: Man muß nachrechnen:

$$\begin{aligned} F_{m,n}(A + B) &= F_{m,n}(A) + F_{m,n}(B), \\ F_{m,n}(\alpha \cdot A) &= \alpha F_{m,n}(A). \end{aligned}$$

Beide Gleichheiten von linearen Abbildungen zeigt man durch Einsetzen eines beliebigen Vektors $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ und Anwendung der Definitionen. Dies sind einfache Rechnungen, ähnlich denen wie unter (i).

(iii) $F_{m,n}$ ist ein Isomorphismus: Beobachtung: Für $1 \leq i \leq n$ gilt

$$F_{m,n}(A)(e_i) = A \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \text{ } i\text{-te Stelle} = \begin{pmatrix} a_{1,i} \\ \vdots \\ a_{m,i} \end{pmatrix} = i\text{-te Spalte von } A.$$

Nun ist $(e_1, \ldots, e_n)$ eine Basis des K^n und nach 3.40 ist eine lineare Abbildung $\varphi\colon K^n \to K^m$ eindeutig durch das System $(\varphi(e_1), \ldots, \varphi(e_n))$ im K^m gegeben und umgekehrt. Daher ist $F_{m,n}$ ein Isomorphismus.

(iv) Es verbleibt, die Formel für die Komposition zu zeigen. Es genügt zu zeigen: für $i = 1, \ldots, k$ gilt

$$F_{m,k}(A \cdot B)(e_i) = F_{m,n}(A)(F_{n,k}(B)(e_i))$$

Linke Seite:

$$F_{m,k}(A \cdot B)(e_i) = i\text{-te Spalte von } (A \cdot B) \quad = \begin{pmatrix} a_{11}b_{1i} + a_{12}b_{2i} + \cdots + a_{1n}b_{ni} \\ \vdots \\ a_{m1}b_{1i} + a_{m2}b_{2i} + \cdots + a_{mn}b_{ni} \end{pmatrix}$$

Rechte Seite:

$$A \cdot (i\text{-te Spalte von } B) = \text{dasselbe.}$$

□

Definition 4.4. Die Matrix $E_n = (\delta_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ (also Einsen auf der Hauptdiagonale und sonst Nullen) heißt die *Einheitsmatrix vom Rang n*.

Bemerkung. Es gilt $F_{n,n}(E_n) = \mathrm{id}_{K^n}$.

Korollar 4.5. *Das Matrizenprodukt ist assoziativ* [1]. *Für $m = n = k$ erhalten wir einen Isomorphismus unitärer Ringe*

$$F_{n,n}\colon M_{n,n}(K) \xrightarrow{\sim} \mathrm{End}_K(K^n).$$

[1] das kann man natürlich auch zu Fuß ausrechnen

Definition 4.6. Eine $n \times n$ Matrix A heißt *invertierbar*, wenn $F_{n,n}(A) \colon K^n \to K^n$ ein Automorphismus ist. Die Menge der invertierbaren $n \times n$-Matrizen wird mit $\mathrm{Gl}_n(K)$ bezeichnet und bildet eine Gruppe bezüglich Matrizenmultiplikation, die durch $F_{n,n}$ isomorph auf $\mathrm{Gl}(K^n) = \mathrm{Aut}_K(K^n)$ abgebildet wird. Das Inverse zu einer invertierbaren Matrix $A \in \mathrm{Gl}_n(K)$ heißt die zu A inverse Matrix und wird mit A^{-1} bezeichnet.

Lemma 4.7. *Eine Matrix $A \in M_{n,n}(K)$ ist genau dann invertierbar, wenn eine Matrix $B \in M_{n,n}(A)$ mit $B \cdot A = E_n$ existiert. Es gilt dann automatisch auch $A \cdot B = E_n$ und somit $B = A^{-1}$.*

Beweis.
($\Leftarrow$): Es gilt $F_{n,n}(B) \circ F_{n,n}(A) = \mathrm{id}$, insbesondere ist $F_{n,n}(A)$ ein injektiver Endomorphismus des endlichdimensionalen K-Vektorraums K^n, also ein Automorphismus nach 3.31.
($\Rightarrow$): trivial.
$A \cdot B = E_n$, $B = A^{-1}$: Eindeutigkeit des Inversen in einer Gruppe. □

Definition 4.8. Wir nennen ein Diagramm von Vektorräumen und linearen Abbildungen *kommutativ*, wenn jede Verbindung zwischen zwei Vektorräumen im Diagramm dieselbe Abbildung repräsentiert.

Beispiele.

$$\begin{array}{ccc} U & \xrightarrow{f} & V \\ & \searrow^{h} & \downarrow{\scriptstyle g} \\ & & W \end{array}$$

ist kommutativ, falls $h = g \circ f$.

$$\begin{array}{ccc} U & \xrightarrow{f} & V \\ {\scriptstyle g'}\downarrow & & \downarrow{\scriptstyle g} \\ W & \xrightarrow[f']{} & S \end{array}$$

ist kommutativ, falls $g \circ f = f' \circ g'$.

Seien nun V, W endlich-dimensionale Vektorräume, $n = \dim V$, $m = \dim W$ und $(v_1, \ldots, v_n)$, $(w_1, \ldots, w_m)$ Basen von V und W. Nach 3.17 erhalten wir Isomorphismen

$$\phi_{v_1,\ldots,v_n} \colon K^n \xrightarrow{\sim} V, \qquad \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \longmapsto \sum_{i=1}^{n} \alpha_i v_i$$

$$\psi_{w_1,\ldots,w_m} \colon K^m \xrightarrow{\sim} W, \qquad \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \longmapsto \sum_{j=1}^{m} \beta_j w_j.$$

Für $A \in M_{m,n}(K)$ erhalten wir ein kommutatives Diagramm

$$\begin{array}{ccc} K^n & \xrightarrow{F_{m,n}(A)} & K^m \\ {\scriptstyle \phi_{v_1,\dots,v_n}}\downarrow{\scriptstyle \sim} & & {\scriptstyle \sim}\downarrow{\scriptstyle \psi_{w_1,\dots,w_m}} \\ V & \xrightarrow[\psi_{w_1,\dots,w_m}\circ F_{m,n}(A)\circ\phi^{-1}_{v_1,\dots,v_n}]{} & W \end{array}$$

Korollar 4.9. *Wir erhalten einen Isomorphismus von Vektorräumen*

$$\begin{aligned} F^{v_1,\dots,v_n}_{w_1,\dots,w_m} \colon M_{m,n}(K) &\longrightarrow \mathrm{Hom}_K(V,W) \\ A &\longmapsto \psi_{w_1,\dots,w_m} \circ F_{m,n}(A) \circ \phi^{-1}_{v_1,\dots,v_n} \end{aligned}$$

Definition 4.10. Den inversen Isomorphismus bezeichnen wir mit

$$M^{v_1,\dots,v_n}_{w_1,\dots,w_m} \colon \mathrm{Hom}_K(V,W) \xrightarrow{\sim} M_{m,n}(K).$$

Für $f \in \mathrm{Hom}_K(V,W)$ heißt $M^{v_1,\dots,v_n}_{w_1,\dots,w_m}(f)$ die *darstellende Matrix* der linearen Abbildung f bzgl. der Basen $(v_1,\dots,v_n)$ und $(w_1,\dots,w_n)$. Alternative Bezeichnungen: *Darstellungsmatrix*, *Koordinatenmatrix*.

Aus den Definitionen folgt für eine lineare Abbildung $f : V \to W$ das kommutative Diagramm

$$\begin{array}{ccc} K^n & \xrightarrow{F_{m,n}\left(M^{v_1,\dots,v_n}_{w_1,\dots,w_m}(f)\right)} & K^m \\ {\scriptstyle \phi_{v_1,\dots,v_n}}\downarrow{\scriptstyle \sim} & & {\scriptstyle \sim}\downarrow{\scriptstyle \psi_{w_1,\dots,w_m}} \\ V & \xrightarrow[f]{} & W \end{array}$$

4.2 Wechsel der Basen

Seien nun $(v'_1,\dots,v'_n)$ und $(w'_1,\dots,w'_n)$ andere Basen von V bzw. W. Wie ändert sich für eine lineare Abbildung $f\colon V \to W$ die repräsentierende Matrix?

Definition 4.11. Sind $(v_1,\dots,v_n)$ und $(v'_1,\dots,v'_n)$ zwei Basen desselben Vektorraums V, so heißt die Matrix

$$T = M^{v_1,\dots,v_n}_{v'_1,\dots,v'_n}(\mathrm{id}_V) \in M_{n,n}(K)$$

die *Transformationsmatrix* von $(v_1,\dots,v_n)$ nach $(v'_1,\dots,v'_n)$.

Lemma 4.12. *T ist invertierbar. T^{-1} ist die Transformationsmatrix von $(v'_1,\dots,v'_n)$ zu $(v_1,\dots,v_n)$.*

Beweis. Das Diagramm

$$\begin{array}{ccccc} K^n & \xrightarrow{F_{n,n}\left(M^{v_1,\ldots,v_n}_{v'_1,\ldots,v'_n}(\mathrm{id}_V)\right)} & K^n & \xrightarrow{F_{n,n}\left(M^{v'_1,\ldots,v'_n}_{v_1,\ldots,v_n}(\mathrm{id}_V)\right)} & K^n \\ \sim\downarrow{\scriptstyle \phi_{v_1,\ldots,v_n}} & & \sim\downarrow{\scriptstyle \phi_{v'_1,\ldots,v'_n}} & & \sim\downarrow{\scriptstyle \phi_{v_1,\ldots,v_n}} \\ V & \xrightarrow[\mathrm{id}_V]{} & V & \xrightarrow[\mathrm{id}_V]{} & V \end{array}$$

setzt sich aus zwei kommutativen Diagrammen zusammen und ist deshalb kommutativ. Daher gilt

$$\mathrm{id}_V \circ \mathrm{id}_V \circ \phi_{v_1,\ldots,v_n} = \phi_{v_1,\ldots,v_n} \circ F_{n,n}(M^{v'_1,\ldots,v'_n}_{v_1,\ldots,v_n}(\mathrm{id}_V)) \circ F_{n,n}(M^{v_1,\ldots,v_n}_{v'_1,\ldots,v'_n}(\mathrm{id}_V)).$$

Eine Anwendung von $\phi^{-1}_{v_1,\ldots,v_n}$ liefert

$$\mathrm{id}_{K^n} = F_{n,n}(M^{v'_1,\ldots,v'_n}_{v_1,\ldots,v_n}(\mathrm{id}_V)) \circ F_{n,n}(M^{v_1,\ldots,v_n}_{v'_1,\ldots,v'_n}(\mathrm{id}_V)).$$

Wegen $\mathrm{id}_{K^n} = F_{n,n}(E_n)$ und weil $F_{n,n}$ nach 4.5 ein Isomorphismus ist, folgt

$$E_n = M^{v'_1,\ldots,v'_n}_{v_1,\ldots,v_n}(\mathrm{id}_V) \cdot M^{v_1,\ldots,v_n}_{v'_1,\ldots,v'_n}(\mathrm{id}_V).$$

□

Bemerkung. Im Moment haben wir noch kein Verfahren zur Berechnung der inversen Matrix T^{-1} zur Hand. Diesem Problem widmen wir uns später.

Wir erinnern uns an das kommutative Diagramm

$$\begin{array}{ccc} K^n & \xrightarrow{F_{m,n}\left(M^{v_1,\ldots,v_n}_{w_1,\ldots,w_m}(f)\right)} & K^m \\ {\scriptstyle \phi_{v_1,\ldots,v_n}}\downarrow\sim & & \sim\downarrow{\scriptstyle \psi_{w_1,\ldots,w_m}} \\ V & \xrightarrow[f]{} & W \end{array}$$

Hieraus erkennen wir:

Lemma 4.13. *Ist $M^{v_1,\ldots,v_n}_{w_1,\ldots,w_m}(f) = (a_{ij}) \in M_{m,n}(K)$, so gilt für $j = 1, \ldots, n$:*

$$f(v_j) = a_{1j}w_1 + \ldots + a_{mj}w_m.$$

Beweis. In der j-ten Spalte der Matrix $M^{v_1,\ldots,v_n}_{w_1,\ldots,w_m}(f) = (a_{ij})$ steht das Bild des j-ten Basisvektors, also der Vektor $F_{m,n}(M^{v_1,\ldots,v_n}_{w_1,\ldots,w_m}(f))(e_j) \in K^m$. Nun gilt $\phi_{v_1,\ldots,v_n}(e_j) = v_j$ und die Kommutativität des Diagramms zeigt

$$\begin{aligned} f(v_j) = \psi_{w_1,\ldots,w_m}(F_{m,n}(M^{v_1,\ldots,v_n}_{w_1,\ldots,w_m}(f))(e_j)) &= \psi_{w_1,\ldots,w_m}\left(\begin{pmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{pmatrix}\right) \\ &= a_{1j}w_1 + \ldots + a_{mj}w_m. \end{aligned}$$

□

Lemma 4.14. *Seien U, V, W endlich-dimensionale Vektorräume $(u_1, \dots, u_n)$, $(v_1, \dots, v_m)$ und $(w_1, \dots, w_k)$ Basen von U, V und W und $f : U \to V$ und $g : V \to W$ lineare Abbildungen. Dann gilt:*

$$M^{v_1,\dots,v_m}_{w_1,\dots,w_k}(g) \cdot M^{u_1,\dots,u_n}_{v_1,\dots,v_m}(f) = M^{u_1,\dots,u_n}_{w_1,\dots,w_k}(g \circ f).$$

Beweis. Es sei $M^{u_1,\dots,u_n}_{w_1,\dots,w_k}(g \circ f) = (c_{ij})$. Dann gilt:

$$g(f(u_j)) = c_{1j}w_1 + \cdots + c_{kj}w_k.$$

Setzt man $M^{u_1,\dots,u_n}_{v_1,\dots,v_m}(f) = (b_{ij})$ und $M^{v_1,\dots,v_m}_{w_1,\dots,w_k}(g) = (a_{ij})$, so gilt

$$f(u_j) = b_{1j}v_1 + \cdots b_{mj}v_m$$

und

$$g(v_i) = a_{1i}w_1 + \ldots + a_{ki}w_k.$$

Zusammen ergibt sich

$$g(f(u_j)) = g(b_{1j}v_1 + \cdots b_{mj}v_m) = b_{1,j}g(v_1) + \cdots + b_{mj}g(v_m) =$$

$$b_{1j}a_{1\,1}w_1 + \cdots + b_{1j}a_{k1}w_k + b_{2j}a_{1\,2}w_1 + \cdots .$$

Wir erhalten durch Koeffizientenvergleich von $g(f(u_j))$ vor w_i für $i = 1, \dots, k$, $j = 1, \dots, n$

$$c_{ij} = a_{i1}b_{1j} + \cdots + a_{im}b_{mj}.$$

□

Aus diesem Lemma erhalten wir einen neuen Beweis für 4.12, denn es gilt nach 4.14

$$M^{v'_1,\dots,v'_n}_{v_1,\dots,v_n}(\mathrm{id}_V) \cdot M^{v_1,\dots,v_n}_{v'_1,\dots,v'_n}(\mathrm{id}_V) = M^{v_1,\dots,v_n}_{v_1,\dots,v_n}(\mathrm{id}_V) = E_n.$$

Satz 4.15. *(Basiswechselsatz). Seien V, W endlich-dimensionale Vektorräume und $f\colon V \to W$ eine lineare Abbildung.*
Seien $(v_1, \dots, v_n)$ und $(v'_1, \dots, v'_n)$ zwei Basen von V und $T_1 \in M_{n,n}(K)$ die Transformationsmatrix.
Seien weiterhin $(w_1, \dots, w_m)$ und $(w'_1, \dots, w'_m)$ zwei Basen von W und T_2 die Transformationsmatrix. Dann gilt

$$M^{v'_1,\dots,v'_n}_{w'_1,\dots,w'_m}(f) = T_2 \cdot M^{v_1,\dots,v_n}_{w_1,\dots,w_m}(f) \cdot T_1^{-1}$$

Beweis. Wir wenden 4.14 auf die Abbildungen

$$V \xrightarrow{\mathrm{id}_V} V \xrightarrow{f} W \xrightarrow{\mathrm{id}_W} W$$

und die Basen $(v'_1, \ldots, v'_n)$, $(v_1, \ldots, v_n)$, $(w_1, \ldots, w_m)$, $(w'_1, \ldots, w'_m)$ an und erhalten unter Verwendung von $T_1^{-1} = M_{v_1,\ldots,v_n}^{v'_1,\ldots,v'_n}(\mathrm{id}_V)$ (siehe 4.12)

$$\begin{aligned} T_2 \cdot M_{w_1,\ldots,w_m}^{v_1,\ldots,v_n}(f) \cdot T_1^{-1} &= M_{w'_1,\ldots,w'_m}^{w_1,\ldots,w_m}(\mathrm{id}_W) \cdot M_{w_1,\ldots,w_m}^{v_1,\ldots,v_n}(f) \cdot M_{v_1,\ldots,v_n}^{v'_1,\ldots,v'_n}(\mathrm{id}_V) \\ &= M_{w'_1,\ldots,w'_m}^{v_1,\ldots,v_n}(f) \cdot M_{v_1,\ldots,v_n}^{v'_1,\ldots,v'_n}(\mathrm{id}_V) \\ &= M_{w'_1,\ldots,w'_m}^{v'_1,\ldots,v'_n}(f) \end{aligned}$$

□

4.3 Ränge von Matrizen

Definition 4.16. Sei $A \in M_{m,n}(K)$ eine Matrix. Der *Zeilen-(bzw. Spalten-)Rang* von A ist die Dimension des durch die Zeilen (bzw. Spalten) von A im K^n (bzw. im K^m) aufgespannten Untervektorraums.
Bezeichnung: $Z\,\mathrm{Rg}(A)$, $S\,\mathrm{Rg}(A)$.

Es spannen n Vektoren im K^m höchstens einen Vektorraum der Dimension $\min(m, n)$ auf. Daher gilt:

$$0 \leq Z\,\mathrm{Rg}(A), S\,\mathrm{Rg}(A) \leq \min(m, n).$$

Ziel dieses Abschnitts ist der Beweis von

Satz 4.17. *(Rangsatz für Matrizen). Für jede Matrix $A \in M_{m,n}(K)$ gilt*

$$Z\,\mathrm{Rg}(A) = S\,\mathrm{Rg}(A).$$

Danach werden wir die Notation $\mathrm{Rg}(A)$ für diese Zahl benutzen.

Um den Rangsatz zu zeigen, beginnen wir mit

Lemma 4.18.

(i) Für $A \in M_{m,n}(K)$ gilt

$$S\,\mathrm{Rg}(A) = \mathrm{Rg}(F_{m,n}(A)).$$

(ii) Seien V, W endlich-dimensionale Vektorräume, $n = \dim V$, $m = \dim W$, und $(v_1, \ldots, v_n)$, $(w_1, \ldots, w_m)$ Basen von V und W. Sei $f : V \to W$ ein Homomorphismus. Dann gilt

$$\mathrm{Rg}(f) = S\,\mathrm{Rg}(M_{w_1,\ldots,w_m}^{v_1,\ldots,v_n}(f)).$$

Beweis.
(i): Die Spalten von A sind die Bilder der Basisvektoren $e_1, \dots, e_n \in K^n$ unter $F_{m,n}(A)$. Diese Bilder spannen $F_{m,n}(A)(K^n)$ auf. Dies zeigt (i).
(ii): Aus dem kommutativen Diagramm

$$\begin{array}{ccc} K^n & \xrightarrow{F_{m,n}\left(M^{v_1,\dots,v_n}_{w_1,\dots,w_m}(f)\right)} & K^m \\ {\scriptstyle \phi_{v_1,\dots,v_n}}\downarrow{\scriptstyle \sim} & & {\scriptstyle \sim}\downarrow{\scriptstyle \psi_{w_1,\dots,w_m}} \\ V & \xrightarrow[f]{} & W \end{array}$$

folgt, dass die Einschränkung von $\psi_{w_1,\dots,w_m}$ auf den Untervektorraum

$$\operatorname{im}(F_{m,n}(M^{v_1,\dots,v_n}_{w_1,\dots,w_m}(f))) \subset K^m$$

einen Isomorphismus

$$\operatorname{im}(F_{m,n}(M^{v_1,\dots,v_n}_{w_1,\dots,w_m}(f))) \xrightarrow{\sim} \operatorname{im}(f)$$

induziert. Insbesondere gilt

$$\operatorname{Rg}(f) = \operatorname{Rg}(F_{m,n}(M^{v_1,\dots,v_n}_{w_1,\dots,w_m}(f))) \overset{(i)}{=} S\operatorname{Rg}(M^{v_1,\dots,v_n}_{w_1,\dots,w_m}(f)).$$

□

Definition 4.19. Zu $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in M_{m,n}(K)$ definiert man die *transponierte Matrix* $A^t \in M_{n,m}(K)$ (oft auch ${}^tA$) durch

$$A^t = (a_{ji})_{\substack{j=1,\dots,n \\ i=1,\dots,m}}.$$

Beispiele.

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

$$\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^t = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Bemerkung. Offenbar gilt

$$Z\operatorname{Rg}(A) = S\operatorname{Rg}(A^t)$$

Für $A \in M_{m,n}(K)$, $B \in M_{n,k}(K)$ gilt

$$(A \cdot B)^t = B^t \cdot A^t \in M_{k,m}(K).$$

Erinnerung: Ist V ein endlich-dimensionaler Vektorraum mit Basis $(v_1, \dots, v_n)$, so hat der duale Vektorraum $V^* = \mathrm{Hom}_K(V, K)$ die duale Basis $v_1^*, \dots, v_n^*$ die durch $v_i^*(v_j) = \delta_{ij}$ gegeben ist.
Eine lineare Abbildung $f\colon V \to W$ induziert die duale Abbildung

$$f^*\colon W^* \to V^*, \varphi \mapsto \varphi \circ f.$$

Für $V \xrightarrow{f} W \xrightarrow{g} U$ gilt $(g \circ f)^* = f^* \circ g^*\colon U^* \to V^*$.

Lemma 4.20. *Seien V, W endlich-dimensionale Vektorräume und $(v_1, \dots, v_n)$, $(w_1, \dots, w_m)$ Basen. Dann gilt für jede lineare Abbildung $f\colon V \to W$:*

$$M^{w_1^*,\dots,w_m^*}_{v_1^*,\dots,v_n^*}(f^*) = (M^{v_1,\dots,v_n}_{w_1,\dots,w_m}(f))^t \in M_{n,m}(K).$$

Beweis. Sei $M^{v_1,\dots,v_n}_{w_1,\dots,w_m}(f) = (a_{ij}) \in M_{m,n}(K)$, d.h. für $i = 1, \dots, n$ gilt

$$f(v_i) = a_{1i}w_1 + \cdots + a_{mi}w_m.$$

Die Spalten der Matrix auf der linken Seite sind die Koordniaten (in $(v_1^*, \dots, v_n^*)$) der Bilder der Basisvektoren $w_1^*, \dots, w_m^*$ unter $f*$. Um Gleichheit mit $(a_{ij})^t$ zu zeigen, ist also für jedes $1 \le j \le m$ zu zeigen, dass

$$f^*(w_j^*) = a_{j1}v_1^* + \cdots + a_{jn}v_n^*.$$

Beide Seiten sind Linearformen auf V, d. h. Elemente in $V^* = \mathrm{Hom}_K(V, K)$. Es genügt daher zu zeigen, dass für alle i, $1 \le i \le n$, gilt:

$$f^*(w_j^*)(v_i) = (a_{j1}v_1^* + \cdots + a_{jn}v_n^*)(v_i) = a_{ji}.$$

Nun ist $f^*(w_j^*)$ die Komposition $V \xrightarrow{f} W \xrightarrow{w_j^*} K$. Daher gilt

$$\begin{aligned} f^*(w_j^*)(v_i) &= w_j^*(f(v_i)) \\ &= w_j^*(a_{1i}w_1 + \cdots + a_{mi}w_m) \\ &= a_{ji} \end{aligned}$$

□

Beweis des Rangsatzes für Matrizen 4.17. Wir betrachten die Abbildung

$$f = F_{m,n}(A)\colon K^n \longrightarrow K^m.$$

Nach 4.20 wird $f^* : (K^m)^* \to (K^n)^*$ bezüglich der zu den kanonischen Basen dualen Basen $(e_1^*, \dots, e_m^*)$ von $(K^m)^*$ und $(e_1^*, \dots, e_n^*)$ von $(K^n)^*$ durch die transponierte Matrix A^t dargestellt.
Der Rangsatz für lineare Abbildungen 3.46 und 4.18 zeigen uns daher

$$S\,\mathrm{Rg}(A) = \mathrm{Rg}\, f = \mathrm{Rg}\, f^* = S\,\mathrm{Rg}(A^t) = Z\,\mathrm{Rg}(A).$$

□

Korollar 4.21. *Für eine Matrix $A \in M_{n,n}(K)$ sind die folgenden Aussagen äquivalent:*

(i) A ist invertierbar, d.h. $A \in Gl_n(K)$

(ii) die Zeilen von A bilden eine Basis des K^n

(iii) die Spalten von A bilden eine Basis des K^n

(iv) $\mathrm{Rg}(A) = n$.

Beweis. Die Äquivalenz (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv) folgt aus dem Rangsatz. Schließlich gilt

$$\begin{aligned} A \text{ invertierbar} &\iff F_{n,n}(A) \text{ ist bijektiv} \\ &\overset{3.31}{\iff} F_{n,n}(A) \text{ ist surjektiv} \\ &\iff \mathrm{Rg}(A) = n. \end{aligned}$$

□

Korollar 4.22. *Ist $A \in M_{m,n}(K)$ und $T \in \mathrm{Gl}_n(K)$, $S \in \mathrm{Gl}_m(K)$, so gilt*

$$\mathrm{Rg}(A) = \mathrm{Rg}(SAT).$$

Beweis. Sei

$$f = F_{m,n}(A) \colon K^n \to K^m.$$

Sei $w_1, \dots, w_m$ die durch die Spalten von S^{-1} gegebene Basis des K^m und $v_1, \dots, v_n$ die durch die Spalten von T gegebene Basis des K^n. Dann ist S die Transformationsmatrix von $(e_1, \dots, e_m)$ zu $(w_1, \dots, w_m)$ im K^n und T^{-1} die Transformationsmatrix von $(e_1, \dots, e_n)$ zu $(v_1, \dots, v_n)$ im K^n. Nach 4.15 gilt

$$\begin{aligned} M^{v_1,\dots,v_n}_{w_1,\dots,w_m}(f) &= SA(T^{-1})^{-1} \\ &= SAT \end{aligned}$$

und deshalb nach 4.18

$$\mathrm{Rg}(A) = \mathrm{Rg}(f) = \mathrm{Rg}(SAT).$$

□

Korollar 4.23. *Sei $A \in M_{m,n}(K)$ und $r = \mathrm{Rg}(A)$. Dann gibt es eine invertierbare $r \times r$-Untermatrix von A, d.h. Indizes $1 \le i_1 < \cdots < i_r \le m$, $1 \le j_1 < \cdots < j_r \le n$, so dass die Matrix*

$$\tilde{A} = (a_{ij})_{\substack{i \in \{i_1,\dots,i_r\} \\ j \in \{j_1,\dots,j_r\}}}$$

invertierbar ist. Umgekehrt folgt aus der Existenz einer invertierbaren $s \times s$ Untermatrix, dass $s \le r$ gilt.

Beweis. Wähle r Zeilen so aus, dass diese eine Basis des (r-dimensionalen) von den Zeilen von A im K^n aufgespannten Untervektorraums bilden. Dann streiche die anderen Zeilen. Die so erhaltene Matrix hat Rang r. Dann wähle r linear unabhängige Spalten aus und erhalte eine $r \times r$-Matrix mit Rang $= r$.
Umgekehrt: Ist A' eine invertierbare $s \times s$-Untermatrix von A, so sind die Spalten von A' linear unabhängig, also auch die s-vielen Spalten von A, die A' treffen $\Rightarrow r \geq s$. □

Satz 4.24. *Sei $f\colon V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen Vektorräumen, $n = \dim V$, $m = \dim W$, $r = \mathrm{Rg}(f)$.*

Dann gibt es Basen $(v_1, \ldots, v_n)$, $(w_1, \ldots, w_m)$ von V und W, so dass

$$M^{v_1,\ldots,v_n}_{w_1,\ldots,w_m}(f) = \underbrace{\left(\begin{array}{cccc|c} \overbrace{1}^{r} & & & 0 & \\ & 1 & & & \\ & & \ddots & & 0 \\ 0 & & & 1 & \\ \hline & & 0 & & 0 \end{array}\right)}_{n} \Bigg\} m$$

Beweis. $\mathrm{Rg}(f) = \mathrm{Rg}(M^{\text{irgendeine Basis}}_{\text{irgendeine Basis}}(f)) \leq \min(n, m)$.
Es gilt $r = \dim f(V) = \dim(V) - \dim \ker(f)$ also $\dim \ker(f) = n - r$.

Sei $U \subset V$ ein Komplement zu $\ker(f)$ (siehe 3.35). Wähle eine Basis $(v_1, \ldots, v_n)$ von V, so dass $\mathrm{Lin}(v_1, \ldots, v_r) = U$ und $\mathrm{Lin}(v_{r+1}, \ldots, v_n) = \ker(f)$. Die eingeschränkte Abbildung $f\colon U \to W$ ist wegen $U \cap \ker(f) = \{0\}$ injektiv. Sei nun $w_i = f(v_i)$ für $i = 1, \ldots, r$ und wähle $w_{r+1}, \ldots, w_m$ so dass $w_1, \ldots, w_m$ eine Basis von W ist. □

Korollar 4.25. *Sei $A \in M_{m,n}(K)$ und $r = \mathrm{Rg}(A)$. Dann existieren Matrizen $S \in Gl_m(K)$, $T \in \mathrm{Gl}_n(K)$ mit*

$$SAT = \underbrace{\left(\begin{array}{cccc|c} \overbrace{1}^{r} & & & 0 & \\ & 1 & & & \\ & & \ddots & & 0 \\ 0 & & & 1 & \\ \hline & & 0 & & 0 \end{array}\right)}_{n} \Bigg\} m$$

Beweis. Wende 3.10 auf $F_{m,n}(A)\colon K^n \to K^m$ und den Basiswechselsatz 4.15 an. □

5 Lineare Gleichungssysteme

5.1 Gauß-Elimination

Aufgabe: $v_1, \dots, v_m \in K^n$ seien gegeben. Berechne eine Basis von $\mathrm{Lin}(v_1, \dots, v_m)$!

Definition 5.1. Zwei Systeme $(v_1, \dots, v_m)$ und $(w_1, \dots, w_k)$ von Vektoren im K^n heißen *linear äquivalent*, wenn $\mathrm{Lin}(v_1, \dots, v_m) = \mathrm{Lin}(w_1, \dots, w_k)$ gilt. Man setzt:

$$\mathrm{Rg}(v_1, \dots, v_m) = \dim \mathrm{Lin}(v_1, \dots, v_m).$$

Die folgenden Operationen heißen *Zeilen-Umformungen*. Mit ihrer Hilfe erhält man aus einem System von Vektoren ein linear äquivalentes System:

(i) Multiplikation von v_i mit $\lambda \neq 0$ für ein i

(ii) Ersetzen von v_i durch $v_i + \lambda v_j$, $i \neq j$

(iii) Vertauschen der v_i

(iv) Streichen von v_i, wenn $v_i = 0$.

(Nullvektoren sind entbehrlich, die anderen Operationen sind umkehrbar.)

Bemerkung. Schreibt man $v_1, \dots, v_m$ als Zeilenvektoren

$$\begin{aligned} v_1^t &= \begin{pmatrix} a_{11} & \dots & a_{1n} \end{pmatrix} \\ &\vdots \quad\quad \vdots \\ v_m^t &= \begin{pmatrix} a_{m1} & \dots & a_{mn} \end{pmatrix}, \end{aligned}$$

so bewirken die Operationen (i), (ii), (iii) auf der $m \times n$-Matrix $A = (a_{ij})$ das folgende:
(i) Multiplikation von links mit der $m \times m$-Matrix $E_i(\lambda)$, die definiert ist durch

$$E_i(\lambda)_{k,\ell} = \begin{cases} 0 & k \neq \ell \\ 1 & k = \ell \neq i \\ \lambda & k = \ell = i \end{cases},$$

d.h.

$$E_i(\lambda) = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & \lambda & & \\ & & & \ddots & \\ 0 & & & & 1 \end{pmatrix} \quad (\lambda \neq 0 \text{ steht an der Stelle } (i,i)).$$

(ii) Multiplikation von links mit der $m \times m$ Matrix

$$E_{i,j}(\lambda)_{k,\ell} = \begin{cases} 1 & k = \ell \\ \lambda & k = i \text{ und } \ell = j \\ 0 & \text{sonst.} \end{cases}$$

d.h.

$$E_{i,j}(\lambda) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ \lambda & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \quad \lambda \text{ steht an der Stelle } (i,j),\ (i \neq j).$$

(iii) Multiplikation von links mit der $m \times m$-Matrix P_{ij}, die gegeben ist durch

$$(P_{i,j})_{k,\ell} = \begin{cases} 1 & (k,\ell) = (i,j) \text{ oder } (j,i) \\ 1 & i \neq k = \ell \neq j \\ 0 & \text{sonst.} \end{cases}$$

(vertausche i-te und j-te Zeile in E_m), d.h.

$$P_{ij} = \left(\begin{array}{ccc|c|ccc|c} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ \hline & & & 0 & & & 1 & \\ \hline & & & & 1 & & & \\ & & & & & \ddots & & \\ & & & & & & 1 & \\ \hline & i & & 1 & & & 0 & \\ \hline & & & & & & & 1 \\ & & & & & & & & \ddots \\ & & & & & & & & & 1 \end{array}\right) \begin{matrix} \\ \\ \\ j \\ \\ \\ \\ i \\ \\ \\ \\ \end{matrix}$$

$$\begin{matrix} & & & j & & & i & & \end{matrix}$$

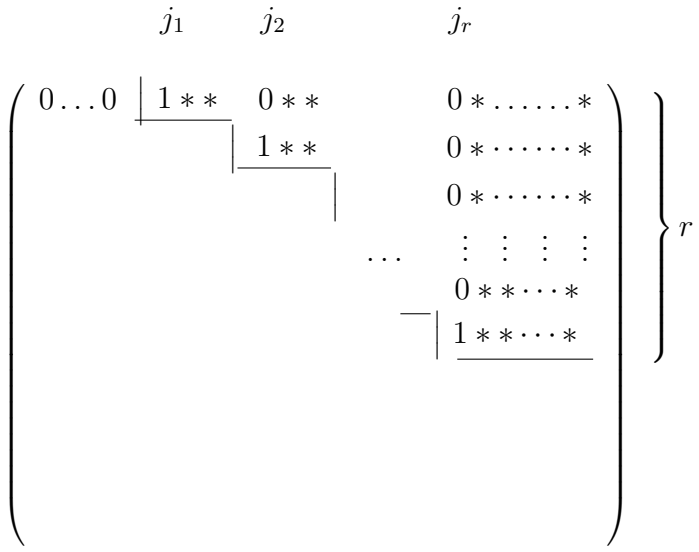
Definition 5.2. Eine $m \times n$-Matrix A hat *Zeilenstufenform*, wenn es ganze Zahlen $0 \leq r \leq m$, $1 \leq j_1 < \cdots < j_r \leq n$ gibt, so dass gilt

(1) $a_{ij} = 0$ falls $i > r$ oder falls $i \leq r$ und $j < j_i$.
(2) $a_{ij_i} = 1$ für alle $1 \leq i \leq r$.

Man sagt dass A *strenge Zeilenstufenform* hat, wenn dazu noch gilt

(3) $a_{ij_k} = 0$ für alle $1 \leq k \leq r$, $k \neq i$.

Bild:



Satz 5.3 (Gauß-Elimination)**.** *Durch Zeilenumformungen (i), (ii), (iii) kann man eine $(m \times n)$-Matrix in strenge Zeilenstufenform bringen.*

Beweis. Die j-te Spalte sei die erste, die nicht nur aus Nullen besteht. $j_1 = j$
(1) Heraufbringen: Ein Zeilentausch (iii) bringt $a_{1j} \neq 0$
(2) Normieren: Anwenden von (i) bewirkt $a_{1j} = 1$
(3) Ausräumen: Durch Ersetzen von v_i durch $v_i - a_{ij}v_1$ erhält die Matrix bis zur j-ten Spalte die gewünschte Form
Jetzt betrachte die Matrix ohne die 1. Zeile und wähle j_2. Mache Schritt 1, 2, 3 und beachte, dass sich auch in Schritt 3 der wesentliche Teil der ursprünglich ersten Zeile nicht ändert usw. □

Beispiel. Sei $\mathrm{char}(K) \neq 2$

$$\begin{pmatrix} 2 & 4 & 2 & 6 \\ 3 & 6 & 3 & 9 \\ 4 & 8 & 5 & 9 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 1 & 3 \\ 3 & 6 & 3 & 9 \\ 4 & 8 & 5 & 9 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -3 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 0 & 6 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Satz 5.4. *Der von Vektoren $v_1, \ldots, v_m$ im K^n erzeugte Untervektorraum ist durch die strenge Stufe eindeutig bestimmt und umgekehrt.*

D.h. gilt $\mathrm{Lin}(v_1, \ldots, v_m) = \mathrm{Lin}(w_1, \ldots, w_k)$ *so unterscheiden sich die strengen Zeilenstufenmatrizen nur um 0-Zeilen unterhalb der Stufe.*

Beweis. Sei $U = \mathrm{Lin}(v_1, \ldots, v_m)$ und $r := \dim U$. Dann ist $U = \mathrm{Lin}(v'_1, \ldots, v'_r)$ wobei $v'_1, \ldots, v'_r$ die ersten r-Zeilen der assoziierten Matrix in strenger Zeilenstufenform sind.
Seien nun $(v_1, \ldots, v_m)$ und $(w_1, \ldots, w_k)$ mit $\mathrm{Lin}(v_1, \ldots, v_m) = \mathrm{Lin}(w_1, \ldots, w_k)$ gegeben und seien $A \in M_{m,n}(K)$ und $B \in M_{k,n}(K)$ die assoziierten Matrizen in strenger Zeilenstufenform. Wegen $\mathrm{Rg}(A) = \dim U = \mathrm{Rg}(B)$ sind bei A und B die i-ten Zeilen mit $i > r = \dim U$ alle Null.
Betrachten wir für $1 \leq j \leq r$ die Projektion $p_j \colon K^n \to K^j$, $(a_1, \ldots, a_n) \mapsto (a_1, \ldots, a_j)$, so sind die Stufen bei A wie bei B gerade die j, mit $\dim p_j(U) > \dim p_{j-1}(U)$. Also haben die Stufen bei A und B die gleiche geometrische Form.

Seien nun $a_1 = (a_{11}, \ldots, a_{1n}), \ldots, a_r = (a_{r1}, \ldots, a_{rn})$ die ersten r-Zeilen von A und $b_1 = (b_{11}, \ldots, b_{1n}), \ldots, b_r = (b_{r1}, \ldots, b_{rn})$ die ersten r-Zeilen von B. Dann sind $(a_1, \ldots, a_r)$ und $(b_1, \ldots, b_r)$ beides Basen von U.

Schreiben wir nun $a_j = \sum_{i=1}^{n} \lambda_{ij} b_i$, so ist (weil die j_i-te Spalte von B genau eine 1 bei (i, j_i) hat) $\lambda_{ij} = a_{jj_i}$. Aber A ist in strenger Zeilenstufenform, also $a_{jj_i} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$. Hieraus folgt $\lambda_{ij} = 0$, falls $i \neq j$, und $= 1$, falls $i = j$, und damit $a_j = b_j$ für $j = 1, \ldots, r$. □

Berechnung der inversen Matrix

Ist $A \in \mathrm{Gl}_n(K)$, d.h. $A \in M_{n,n}(K)$ und $n = \mathrm{Rg}(A)$, so ist die strenge Zeilenstufenform von A die Einheitsmatrix E_n. Dies gibt die folgende Methode zur Berechnung von A^{-1}:
Führe die gleichen Operationen die A auf strenge Zeilenstufenform (d.h. auf E_n) bringen mit E_n aus. Das Ergebnis ist A^{-1}.
Begründung: Jede der Operationen (i), (ii) und (iii) entspricht der Linksmultiplikation mit einer Matrix. Ist M das (von rechts nach links gebildete) Produkt dieser Matrizen, so gilt $M \cdot A = E_n$ und $M \cdot E_n = M$ und also $M = A^{-1}$.

Beispiel. Suche

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

wenn $\mathrm{char}(K) \neq 2$, d.h. $2 \neq 0$, $\frac{1}{2} \in K$:

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

Also gilt $\left(\begin{array}{ccc} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 2 & 2 \end{array}\right) \left(\begin{array}{ccc} 2 & 0 & -1 \\ -1 & -1 & 2 \\ 0 & 1 & -1 \end{array}\right) = E_3$

Berechnen der dualen Basis.

Elemente des Dualraums $(K^n)^*$ sind Linearformen auf K^n. Die kanonische Basis des $(K^n)^*$ ist durch die zur kanonischen Basis $(e_1, \ldots, e_n)$ des K^n duale Basis $(e_1^*, \ldots, e_n^*)$ gegeben. Wie vorher schreiben wir Elemente des K^n als Spaltenvektoren (d.h. $(n\times 1)$-Matrizen). Jeder Zeilenvektor (d.h. $(1\times n)$-Matrix) $(a_1, \ldots, a_n)$ definiert mit Hilfe der Matrixmultiplikation $M_{1,n}(K) \times M_{n,1}(K) \to M_{1,1}(K) = K$ durch

$$(x_1, \ldots, x_n)^t \longmapsto (a_1, \ldots, a_n) \cdot (x_1, \ldots, x_n)^t = a_1x_1 + \cdots + a_nx_n \in K$$

ein Element von $(K^n)^*$. Es gilt

$$(1, 0, \ldots, 0) = e_1^*, \ \ldots \ , (0, \ldots, 0, 1) = e_n^*.$$

Daher läßt sich die Linearform $\varphi = a_1 e_1^* + \cdots + a_n e_n^*$ durch den Zeilenvektor $(a_1, \ldots, a_n)$ darstellen, mit anderen Worten: wir können den $(K^n)^*$ mit dem Vektorraum der Zeilenvektoren der Länge n identifizieren.

Sei nun $(v_1, \ldots, v_n)$ eine Basis des K^n und $(v_1^*, \ldots, v_n^*)$ die duale Basis, d. h. die Basis des $(K^n)^*$, die durch $v_i^*(v_j) = \delta_{ij}$ charakterisiert ist. Die Zeilenvektorform der dualen Basis berechnet man wie folgt:

1. Bilde die Matrix A deren i-te Spalte gleich v_i ist.
2. die i-te Zeile von A^{-1} ist v_i^*.

Begründung: v_i^* ist durch $v_i^*(v_j) = \delta_{ij}$ charakterisiert, d. h. für den als Zeile geschriebenen Vektor v_i^* gilt: $v_i^* \cdot v_j = \delta_{ij}$. Bildet man die Matrix A mit den v_i als Spalten und die Matrix B mit den v_i^* als Zeilen, so gilt $B \cdot A = E_n$, also $B = A^{-1}$.

Beispiel. Betrachte die Basis

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

des $\mathbb{R}^2$. Dann gilt

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \qquad A^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix},$$

also

$$v_1^* = (2, -1) = 2e_1^* - e_2^*, \quad v_2^* = (-1, 1) = -e_1^* + e_2^*.$$

Basisergänzung

Gegeben: Linear unabhängiges System $(v_1, \ldots, v_k)$
und Erzeugendensystem $(w_1, \ldots, w_m)$ im K^n.
Gesucht: Indizes $1 \leq j(1) < \cdots < j(s) \leq m$, $s = n - k$, so dass

$$(v_1, \ldots, v_k, w_{j(1)}, \ldots, w_{j(s)})$$

eine Basis ist (diese Indizes existieren nach 3.25).
Methode: Wir bringen die Matrix A mit Zeilenvektoren $v_1, \ldots, v_k, w_1, \ldots, w_m$ auf strenge Zeilenstufenform. Hierbei mußte (evtl.) mehrere Male der Schritt (i) (Zeilentausch) durchgeführt werden. Wir gehen dabei so vor, dass der notwendige Zeilentausch stets mit der Zeile ausgeführt wird, die am weitesten oben steht und einen nicht verschwindenden Eintrag in der untersuchten Spalte hat. Dann gilt: Unter den ersten n Zeilen sind $s = n - k$ Stück, die ursprünglich zu einem Zeilenvektor $w_{j(i)}$ gehörten ($i = 1, \ldots, s$). Dies liefert die gesuchten $j(i)$.

Beispiel. Im $\mathbb{R}^3$ sei $v_1 = e_1$ und das Erzeugendensystem

$$w_1 = \begin{pmatrix} 2 & 0 & 0 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 1 & 2 & 0 \end{pmatrix}, \quad w_3 = \begin{pmatrix} 0 & 2 & 0 \end{pmatrix}, \quad w_4 = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$$

gegeben, d. h.

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 2 & 0 \\ 1 & 2 & 3 \end{pmatrix} \longrightarrow \begin{matrix} \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 2 & 0 \\ 0 & 2 & 3 \end{pmatrix} \longrightarrow \begin{matrix} \\ 2 \\ 1 \\ 3 \\ 4 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 2 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix} \longrightarrow \begin{matrix} \\ 2 \\ 4 \\ 3 \\ 1 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Also ist (v_1, w_2, w_4) eine Basis von $\mathbb{R}^3$.

Lineares Komplement

Gegeben: $U = \mathrm{Lin}(v_1, \ldots, v_m) \subset K^n$.
Gesucht: $U' \subset K^n$ mit $U \cap U' = \{0\}$ und $U + U' = K^n$ und eine Basis von U'.
Methode: Forme die Matrix mit Zeilen $v_1, \ldots, v_m$ in strenge Zeilenstufenform um. Die von 0 verschiedenen Zeilen sind eine Basis von U, die Vektoren e_j, wobei j kein Stufenindex ist, sind Basis eines Teilraums U' mit $U \cap U' = \{0\}$, $U + U' = K^n$.

Beispiel. Sei $U = \mathrm{Lin}(v_1, v_2) \subset \mathbb{R}^4$ mit

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

Die Matrix

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

ist bereits in strenger Zeilenstufenform. Ein Komplement von U ist $U' = \mathrm{Lin}(e_2, e_4)$.

5.2 Lineare Gleichungssysteme

Definition 5.5. Ein lineares Gleichungssystem ist ein System von Gleichungen

$$\begin{matrix} a_{11}x_1 + a_{12}x_2 & + \cdots + a_{1n}x_n = & b_1 \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 & + \cdots + a_{mn}x_n = & b_m \end{matrix} \qquad (*)$$

$(*)$ heißt *homogen* wenn alle $b_i = 0$ sind, ansonsten *inhomogen*. Ist $(*)$ inhomogen, so nenne man das System $(**)$ mit den gleichen a_{ij} und rechts überall Nullen das *zugehörige homogene System*. Wir schreiben $(*)$ in der Form

$$A \cdot x = b$$

mit $A = (a_{ij})$, $x = (x_1, \ldots, x_n)^t$, $b = (b_1, \ldots, b_m)^t$.

Bemerkung. Wir schreiben ab jetzt A für die zu A assoziierte lineare Abbildung $K^n \xrightarrow{F_{m,n}(A)} K^m$.

Satz 5.6. *Die Menge der Lösungen $x \in K^n$ des homogenen Gleichungssystems*

$$Ax = 0 \qquad (**)$$

ist ein Untervektorraum des K^n. Seine Dimension ist gleich $n - \mathrm{Rg}(A)$.

Beweis. Offensichtlich ist die Lösungsmenge gleich dem Kern von A. Die Aussage über die Dimension folgt aus 3.38 sowie aus $\mathrm{Rg}(A) = \dim(\mathrm{im}(A))$. □

Korollar 5.7. *Die folgenden Aussagen sind äquivalent:*

(i) $\mathrm{Rg}(A) = n$.

(ii) Das homogene System (**) hat genau die triviale Lösung $x = (0, \ldots, 0) \in K^n$.

Ist $n = m$ so sind (i) und (ii) zusätzlich äquivalent zu

(iii) A ist invertierbar.

Beweis. (i) $\Longleftrightarrow$ (iii) haben wir schon in 4.21 gesehen. (i) $\Longleftrightarrow$ (ii) folgt aus 5.6. □

Definition 5.8. Eine Teilmenge A eines Vektorraums V heißt *affiner Teilraum*, wenn es ein $v \in V$ und einen Untervektorraum $U \subset V$ gibt so dass

$$A = v + U = \{v + u \mid u \in U\}$$

gilt.

Bemerkung: Anders ausgedrückt: ein affiner Teilraum ist ein Element in V/U, wobei U ein Untervektorraum von V ist.

Lemma 5.9. *Ist $A \subset V$ ein affiner Teilraum, so gibt es genau einen Untervektorraum U mit der Eigenschaft $A = v + U$ für ein $v \in V$.*

Beweis. Ein solcher Untervektorraum U existiert nach Definition. Sei $A = v_0 + U_0 = v_1 + U_1$. Dann gilt

$$U_0 = \{a - a' \mid a, a' \in A\} = U_1$$

□

Definition 5.10. Ist $A \subset V$ ein affiner Teilraum, so setzt man

$$\dim A := \dim U,$$

wobei U der nach 5.9 eindeutig bestimmte Untervektorraum mit $A = v + U$ ist.

Satz 5.11. *Für das inhomogene Gleichungssystem $Ax = b$ $(*)$ gibt es genau die beiden folgenden Möglichkeiten*

(i) $b \notin \mathrm{im}(A)$ *und die Lösungsmenge L von $(*)$ ist leer.*

(ii) $b \in \mathrm{im}(A)$. *Dann ist die Lösungsmenge L ein affiner Teilraum des K^n der Dimension $n - \mathrm{Rg}(A)$. Ist $v_0 \in L$ eine Lösung von $(*)$, so gilt*

$$L = v_0 + U,$$

wobei U der Lösungsraum des zugehörigen homogenen Systems $Ax = 0$ $(**)$ ist.

Beweis. Es gilt offensichtlich $b \notin \mathrm{im}(A) \iff L = \emptyset$. Ist $b \in \mathrm{im}(A)$ so gibt es eine Lösung $v_0 \in K^n$ mit $Av_0 = b$. Sei $U = \ker(A)$ die Lösungsmenge von $(**)$. Für $u \in U$ gilt $A(v_0 + u) = Av_0 + Au = b$, also $v_0 + U \subset L$. Andererseits sei $v \in L$, d. h. $Av = b$. Dann gilt für $u = v - v_0$: $Au = Av - Av_0 = b - b = 0$, also $u \in U$ und $v = v_0 + u$. Daher gilt auch $L \subset v_0 + U$. Schließlich gilt

$$\dim L = \dim U = n - \mathrm{Rg}(A).$$

□

Korollar 5.12. *Das inhomogene System $(*)$ hat genau dann eine Lösung wenn*

$$\mathrm{Rg}(A) = \mathrm{Rg}(A|b)$$

gilt. Hier ist $(A|b)$ die $m \times (n+1)$ Matrix, die durch Anfügen von b an A als $(n+1)$-te Spalte entsteht.

Beweis. $\mathrm{Rg}(A) = \mathrm{Rg}(A|b) \iff b \in \mathrm{Lin}$ (Spalten von A) $\iff b \in \mathrm{im}(A)$. □

5.3 Explizite Lösung linearer Gleichungssysteme

Wir betrachten das homogene System

$$Ax = 0 \qquad (*)$$

Das Ausführen von Zeilenumformungen (i) – (iii) ändert den Lösungsraum $\ker(A)$ nicht (z.B. weil diese Umformungen der Multiplikation von links und invertierbaren Matrizen entspricht).
Daher erhalten wir folgendes Verfahren:

1. Schritt: bringe A auf strenge Zeilenstufenform
2. Schritt: Sei nun $S = (s_{ij})$ die strenge Zeilenstufenform von A

$$\overset{j_1}{\qquad} \overset{j_2}{\qquad\qquad} \overset{j_r}{\qquad\qquad}$$

$$S = \left(\begin{array}{cccccc} 0\dots0 & 1** & 0** & & 0*\dots\dots* \\ & & 1** & & 0*\cdots\cdots* \\ & & & & 0*\cdots\cdots* \\ & & & \dots & \vdots\ \vdots\ \vdots\ \vdots \\ & & & & 0**\cdots* \\ & & & & 1**\cdots* \end{array}\right) \Bigg\} r$$

Die von $j_1, \dots, j_r$ verschiedenen Indizes in $\{1, \dots, n\}$ seien $k_1 < \cdots < k_{n-r}$, d.h.

$$\{1, \dots, n\} = \{j_1, \dots, j_r, k_1, \dots, k_{n-r}\}$$

Dann gilt

$$Sx = 0 \iff \begin{pmatrix} x_{j_1} \\ \vdots \\ x_{j_r} \end{pmatrix} = - \begin{pmatrix} \sum_{c=1}^{n-r} s_{1k_c} x_{k_c} \\ \vdots \\ \sum_{c=1}^{n-r} s_{rk_c} x_{k_c} \end{pmatrix}$$

Folglich kann man $x_{k_1}, \dots, x_{k_{n-r}}$ beliebig wählen und $x_{j_1}, \dots, x_{j_r}$ ergeben sich dann eindeutig. Setzt man nun für $(x_{k_1}, \dots, x_{k_{n-r}})$ den i-ten Standardbasisvektor im K^{n-r} ein, erhält man eine Basis $(v_1, \dots, v_r)$ der Lösungsmenge, wobei

$$v_c = e_{k_c} - \left(\sum_{i=1}^{r} s_{ik_c} e_{j_i} \right)$$

und $e_j \in K^n$ den j-te Standardbasisvektor bezeichnet. Tatsächlich gilt sogar $s_{ik_c} = 0$ für $j_i > k_c$.

Beispiel. Sei $K = \mathbb{R}$ und

$$\begin{aligned} 2x_1 + 4x_2 + 2x_3 + 6x_4 &= 0 \\ 3x_1 + 6x_2 + 3x_3 + 9x_4 &= 0 \\ 4x_1 + 8x_2 + 5x_3 + 9x_4 &= 0 \end{aligned}$$

Dann gilt

$$A = \begin{pmatrix} 2 & 4 & 2 & 6 \\ 3 & 6 & 3 & 9 \\ 4 & 8 & 5 & 9 \end{pmatrix} \overset{\text{Bsp. nach 5.3}}{\longrightarrow} S = \begin{pmatrix} 1 & 2 & 0 & 6 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Wir lesen ab: $j_1 = 1$, $j_2 = 3$, $k_1 = 2$, $k_2 = 4$.
Basis des 2-dimensionalen Lösungsraums: $((-2, 1, 0, 0), (-6, 0, 3, 1))$

Nun betrachten wir das inhomogene System

$$Ax = b \qquad (*)$$

Die Zeilenumformungen (i), (ii), (iii) auf der Matrix $(A|b)$ verändern nicht den Lösungsraum. Wir kommen auf $(S|s)$ in strenger Zeilenstufenform. Sei $r = \mathrm{Rg}(A) = \mathrm{Rg}(S)$. Wegen $\mathrm{Rg}(A|b) = \mathrm{Rg}(S|s)$ ist die Existenz von Lösungen nach 5.12 äquivalent zu $s_{r+1} = \cdots = s_n = 0$. Ist dies erfüllt, setze $x_j = 0$ für $j \notin \{j_1, \ldots, j_r\}$ und $(x_{j_1}, \ldots, x_{j_r}) = (s_1, \ldots, s_r)$, um eine spezielle Lösung zu erhalten. Alle anderen Lösungen erhält man durch Addition von Lösungen des zugeordneten homogenen Systems.

Beispiel. Sei $K = \mathbb{R}$ und

$$\begin{aligned} 2x_1 + 4x_2 + 2x_3 + 6x_4 &= 4 \\ 3x_1 + 6x_2 + 3x_3 + 9x_4 &= 6 \\ 4x_1 + 8x_2 + 5x_3 + 9x_4 &= 9 \end{aligned}$$

Wir erhalten

$$(A|b) = \begin{pmatrix} 2 & 4 & 2 & 6 & 4 \\ 3 & 6 & 3 & 9 & 6 \\ 4 & 8 & 5 & 9 & 9 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 2 \\ 3 & 6 & 3 & 9 & 6 \\ 4 & 8 & 5 & 9 & 9 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -3 & 1 \end{pmatrix}$$
$$\rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 2 \\ 0 & 0 & 1 & -3 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 0 & 6 & 1 \\ 0 & 0 & 1 & -3 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Es gilt $\mathrm{Rg}(A|b) = 2 = \mathrm{Rg}\, A$, also existiert eine Lösung. Außerdem haben wir die Stufenindizes $j_1 = 1$, $j_2 = 3$: Spezielle Lösung $(1, 0, 1, 0)$.
Die allgemeine Lösung ist $(1, 0, 1, 0) + \mathrm{Lin}((-2, 1, 0, 0), (-6, 0, 3, 1))$
d. h.

$$L = \{(1 - 2x_2 - 6x_4, x_2, 1 + 3x_4, x_4) \in K^4 \mid x_2, x_4 \in K\}.$$

6 Determinanten und Eigenwerte

6.1 Permutationen

Erinnerung:

$$S_n = \mathrm{Aut}(\{1, \ldots, n\})$$

Ein Element $\pi \in S_n$ können wir in der Form

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix}$$

schreiben.

Definition 6.1. Sei $1 \leq k \leq n$. Ein Element $\sigma \in S_n$ heißt *Zykel* der Länge k, wenn es k Zahlen $i_1, \ldots, i_k$ zwischen 1 und n zyklisch vertauscht und alle anderen festhält, d. h.

$$\sigma(j) = \begin{cases} j & \text{falls } j \notin \{i_1, \ldots, i_k\}, \\ i_{r+1} & \text{falls } j = i_r,\ r < k, \\ i_1 & \text{falls } j = i_k. \end{cases} .$$

Schreibweise: $\sigma = \begin{pmatrix} i_1 & i_2 & \ldots & i_k \end{pmatrix}$. Ein Zykel der Länge 2 heißt auch *Transposition*.

Beispiel. Sei $n = 5$.

$$\begin{pmatrix} 1 & 3 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 4 & 1 \end{pmatrix}$$

Lemma 6.2. *Die Transpositionen erzeugen S_n, d. h. jedes Element in S_n kann (auf nicht notwendig eindeutige Weise) als Produkt von Transpositionen geschrieben werden.*

Beweis. Induktion über n.
$n = 1, 2$: trivial.
Sei $n > 2$. Betrachte die Einbettung

$$S_{n-1} \hookrightarrow S_n, \qquad \pi \mapsto \begin{pmatrix} 1 & 2 & \ldots & n-1 & n \\ \pi(1) & \pi(2) & \ldots & \pi(n-1) & n \end{pmatrix},$$

bezüglich derer wir S_{n-1} als Untergruppe in S_n auffassen.
Sei nun $\sigma \in S_n$ beliebig. Gilt $\sigma(n) = n$, so ist $\sigma \in S_{n-1}$ und nach Induktionsvoraussetzung ist σ Produkt von Transpositionen.
Ist $\sigma(n) = m$, $1 \leq m \leq n-1$, so ist $(mn) \cdot \sigma \in S_{n-1}$, also Produkt von Transpositionen $(mn)\sigma = t_1 \ldots t_r$, also $\sigma = (mn) \cdot t_1 \ldots t_r$. □

Wie eindeutig ist die Darstellung als Produkt von Transpositionen? Wir werden sehen, dass die Anzahl der Transpositionen modulo 2 wohlbestimmt ist.

Lemma 6.3. *Die Abbildung*

$$\operatorname{sgn} : S_n \longrightarrow \{\pm 1\}, \quad \sigma \longmapsto \prod_{1 \leq i < j \leq n} \frac{i-j}{\sigma(i) - \sigma(j)}$$

ist ein Gruppenhomomorphismus.

Beweis. Zunächst bemerken wir, dass in $\prod_{1 \leq i < j \leq n} \frac{i-j}{\sigma(i)-\sigma(j)}$ im Zähler wie Nenner die gleichen Faktoren vorkommen, aber eventuell mit verschiedenem Vorzeichen. Also $\operatorname{sgn}(\sigma) \in \{\pm 1\}$. Für $\sigma, \tau \in S_n$ und $1 \leq i < j \leq n$ mit $\tau(i) > \tau(j)$ gilt

$$\frac{\tau(i) - \tau(j)}{\sigma\tau(i) - \sigma\tau(j)} = \frac{\tau(j) - \tau(i)}{\sigma\tau(j) - \sigma\tau(i)},$$

also gilt

$$\prod_{1\le i<j\le n} \frac{\tau(i)-\tau(j)}{\sigma\tau(i)-\sigma\tau(j)} \overset{i'=\tau(i),j'=\tau(j)}{=} \prod_{1\le i'<j'\le n} \frac{i'-j'}{\sigma(i')-\sigma(j')}.$$

Wir erhalten

$$\begin{aligned}
\operatorname{sgn}(\sigma\tau) &= \prod_{1\le i<j\le n} \frac{i-j}{\sigma\tau(i)-\sigma\tau(j)} \\
&= \prod_{1\le i<j\le n} \frac{\tau(i)-\tau(j)}{\sigma\tau(i)-\sigma\tau(j)} \cdot \prod_{1\le i<j\le n} \frac{i-j}{\tau(i)-\tau(j)} \\
&= \prod_{1\le i<j\le n} \frac{i-j}{\sigma(i)-\sigma(j)} \cdot \prod_{1\le i<j\le n} \frac{i-j}{\tau(i)-\tau(j)} \\
&= \operatorname{sgn}(\sigma)\cdot\operatorname{sgn}(\tau).
\end{aligned}$$

□

Definition 6.4. Die Zahl $\operatorname{sgn}(\sigma) \in \{\pm 1\}$ heißt das *Vorzeichen* oder *Signum* der Permutation σ.

Bemerkung. Ein Paar (i,j) mit $1 \le i < j \le n$ und $\sigma(i) > \sigma(j)$ nennt man auch *Fehlstellung* von σ. Damit gilt

$$\operatorname{sgn}(\sigma) = (-1)^{\text{Anzahl der Fehlstellungen von } \sigma}$$

Lemma 6.5. *Für eine Transposition $t = (ij) \in S_n$ gilt* $\operatorname{sgn}(t) = -1$.

Beweis. Sei $t = (ij) \in S_n$ und ohne Einschränkung $i < j$. Wir betrachten für $1 \le \alpha < \beta \le n$ das Vorzeichen von

$$\frac{\alpha-\beta}{t(\alpha)-t(\beta)}.$$

Der Zähler ist stets negativ und wir erhalten die folgende Fallunterscheidung

$\{\alpha,\beta\}\cap\{i,j\} = \varnothing$	$+$	
$\alpha = i < \beta \le j$	$-$	$(j-i)$ mal.
$\alpha = i < j < \beta$	$+$	
$\alpha < i < j = \beta$	$+$	
$i < \alpha < j = \beta$	$-$	$(j-i-1)$ mal
$i < j = \alpha < \beta$	$+$	
$\alpha < \beta = i < j$	$+$	

Also folgt

$$\operatorname{sgn}(t) = (-1)^{2j-2i-1} = -1.$$

□

Satz 6.6. *Ist*

$$\sigma = t_1 \cdots t_r$$

eine Darstellung von σ als Produkt von Transpositionen, so gilt $\operatorname{sgn}(\sigma) = (-1)^r$. *Insbesondere ist die Restklasse von r modulo 2 von der Wahl der Darstellung unabhängig.*

Beweis. Es ist sgn ein Homomorphismus, also

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(t_1) \cdots \operatorname{sgn}(t_r) = (-1)^r.$$

□

Definition 6.7. Die Untergruppe

$$A_n := \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\}$$

heißt die *alternierende Gruppe* (über n Elementen)

Bemerkung. Es gilt $A_n = \ker(\operatorname{sgn}\colon S_n \to \{\pm 1\})$ und für $n \geq 2$ ist sgn surjektiv. Also folgt für $n \geq 2$

$$\#A_n = \frac{\#S_n}{2} = \frac{n!}{2}.$$

6.2 Determinanten

Sei R ein kommutativer, unitärer Ring und $M_{n,n}(R)$ der Ring der $n \times n$-Matrizen mit Einträgen aus R.

Definition 6.8 (Leibniz-Formel der Determinante)**.** Die Determinante von $A = (a_{ij}) \in M_{n,n}(R)$ ist das Element

$$\det A = |A| = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \in R.$$

Hier fassen wir $\operatorname{sgn}(\sigma) \in \{\pm 1\}$ als Element von R auf.

Beispiele.

$n = 1$ $\det(a) = a$.

$n = 2$ $\det \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = ac - bd$.

$n = 3$ $\det \left(\begin{smallmatrix} a & b & c \\ d & e & f \\ g & h & i \end{smallmatrix}\right) = aei - afh + bfg - bdi + cdh - ceg$

Beispiel.

$$\det E_n = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \delta_{1\sigma(1)} \cdots \delta_{n\sigma(n)} = 1.$$

Bemerkung. Sei $f\colon R \to S$ ein unitärer Ringhomomorphismus. Indem wir f auf jeden Eintrag einer Matrix $A \in M_{n,n}(R)$ anwenden, erhalten wir einen unitären Ringhomomorphismus $f\colon M_{n,n}(R) \to M_{n,n}(S)$ und es gilt

$$\det f(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) f(a_{1\sigma(1)}) \cdots f(a_{n\sigma(n)}) = f(\det A).$$

Bemerkung. Die Leibniz-Formel ist für die praktische Berechnung ungeeignet (zu viele Summanden).

Lemma 6.9. *Sei $A \in M_{n,n}(R)$. Dann gilt $\det A = \det A^t$.*

Beweis.

$$\begin{aligned} \det A &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\ &\overset{\tau = \sigma^{-1}}{=} \sum_{\tau \in S_n} \operatorname{sgn}(\tau) a_{\tau(1)1} a_{\tau(2)2} \cdots a_{\tau(n)n} \\ &= \det A^t \end{aligned}$$

□

Lemma 6.10. *Die Abbildung $\det\colon M_{n,n}(R) \to R$ ist in jeder Zeile (jeder Spalte) linear: Sei $1 \le i \le n$ und $v_1, \ldots, v_n, v_i' \in R^n$, $\lambda, \mu \in R$. Dann gilt*

$$\det \begin{pmatrix} v_1 \\ \vdots \\ \lambda v_i + \mu v_i' \\ \vdots \\ v_n \end{pmatrix} = \lambda \det \begin{pmatrix} v_1 \\ \vdots \\ v_i \\ \vdots \\ v_n \end{pmatrix} + \mu \det \begin{pmatrix} v_1 \\ \vdots \\ v_i' \\ \vdots \\ v_n \end{pmatrix} \qquad \textit{(i-te Zeile).}$$

Beweis. Schreibe $v_k = (a_{k1}, \ldots, a_{kn})$,

$$c_\sigma = a_{1,\sigma(1)} \cdots a_{i-1,\sigma(i-1)} a_{i+1,\sigma(i+1)} \cdots a_n.$$

Dann ist die Abbildung

$$R^n \to R, \qquad w = (w_1, \ldots, w_n) \mapsto \det \begin{pmatrix} v_1 \\ \vdots \\ w \\ \vdots \\ v_n \end{pmatrix}_{(i\text{-te Zeile})} = \sum_{\sigma \in S_n} c_\sigma w_{\sigma(i)}$$

offensichtlich R-linear. Die Aussage für Spalten folgt wegen $\det A = \det A^t$. □

Lemma 6.11. *Es gilt* $\det A = 0$ *falls* A *zwei gleiche Zeilen (Spalten) hat.*

Beweis. Schreibe $A = (a_{ij})$ und nehme an, die k-te und ℓ-te Zeile sind gleich, d. h. $a_{kj} = a_{\ell j}$ für alle j. Setze $\tau = (k\ell) \in S_n$. Dann gilt $S_n = A_n \sqcup \tau A_n$ und für alle $\sigma \in A_n$

$$a_{1\sigma(1)} \cdots a_{n\sigma(n)} = a_{1\tau\sigma(1)} \cdots a_{n\tau\sigma(n)},$$

also

$$\det A = \sum_{\sigma \in A_n} (a_{1\sigma(1)} \cdots a_{n\sigma(n)} - a_{1\tau\sigma(1)} \cdots a_{n\tau\sigma(n)}) = 0.$$

□

Definition 6.12. Wir nennen eine Abbildung $\alpha\colon M_{n,n}(R) \to R$ eine *alternierende n-Form*, falls

(i) α ist in jeder Zeile linear,

(ii) $\alpha(A) = 0$ falls A zwei gleiche Zeilen hat.

Erinnerung:
Für $1 \leq i \neq j \leq n$, $\lambda \in R$ haben wir die Elementarmatrizen $E_i(\lambda), E_{ij}(\lambda), P_{ij} \in M_{n,n}(R)$. Für $A \in M_{n,n}(R)$ gilt

$$\begin{aligned} E_i(\lambda)A &= \text{„Multipliziere } i\text{-te Zeile von } A \text{ mit } \lambda\text{”} \\ E_{ij}(\lambda)A &= \text{„Addiere das } \lambda\text{-fache der } j\text{-ten Zeile zur } i\text{-ten Zeile von } A\text{”} \\ P_{ij}A &= \text{„Vertausche } i\text{-te und } j\text{-te Zeile in } A\text{”} \end{aligned}$$

Definition 6.13. Eine Abbildung $\alpha\colon M_{n,n}(R) \to R$ heißt

(i) *homogen*, wenn $\alpha(E_j(\lambda)A) = \lambda\alpha(A)$ für alle $A \in M_{n,n}(R)$, $\lambda \in R$, $1 \leq j \leq n$;

(ii) *scherungsinvariant*, wenn $\alpha(E_{ij}(\lambda)A) = \alpha(A)$ für alle $\lambda \in R$, $1 \leq i \neq j \leq n$.

Lemma 6.14. *Sei* $\alpha\colon M_{n,n}(R) \to R$ *homogen und scherungsinvariant. Dann gilt für* $A \in M_{n,n}(R)$, $1 \leq i \neq j \leq n$

$$\alpha(P_{ij}A) = -\alpha(A)$$

Beweis. Es gilt

$$P_{ij} = E_i(-1)E_{ij}(1)E_{ji}(1)E_{ij}(-1)$$

□

Lemma 6.15. *Sei* $\alpha\colon M_{n,n}(R) \to R$ *eine alternierende n-Form. Dann ist α homogen und scherungsinvariant.*

Beweis. Die Homogenität folgt sofort aus der Linearität in jeder Zeile. Scherungsinvarianz: Seien $v_1, \dots v_n \in R^n$ die Zeilen der Matrix $A \in M_{n,n}(R)$, $1 \leq i \neq j \leq n$, $\lambda \in R$. Dann gilt

$$\begin{aligned}\alpha(E_{ij}(\lambda)A) &= \alpha\begin{pmatrix} v_1 \\ \vdots \\ v_i + \lambda v_j \\ \vdots \\ v_n \end{pmatrix} = \alpha\begin{pmatrix} v_1 \\ \vdots \\ v_i \\ \vdots \\ v_n \end{pmatrix} + \lambda\begin{pmatrix} v_1 \\ \vdots \\ v_j \\ \vdots \\ v_n \end{pmatrix} \qquad (v_j \text{ in } i\text{-ter und } j\text{-ter Zeile}) \\ &= \alpha(A)\end{aligned}$$

□

Bemerkung. Ist K ein Körper, gilt auch die Umkehrung: Jede homogene, scherungsinvariante Abbildung $\alpha\colon M_{n,n}(K) \to K$ ist eine alternierende n-Form.

Notation: Sei $1 \leq i, j \leq n$. Für $A \in M_{n,n}(R)$ sei $A_{ij} \in M_{n-1,n-1}(R)$ die Matrix, die aus A durch Streichen der i-ten Zeilen und j-ten Spalte entsteht. Für $B = (b_{k\ell}) \in M_{n-1,n-1}(R)$ sei

$$\epsilon_{ij}(B) = \begin{pmatrix} b_{1,1} & \dots & b_{1,j-1} & 0 & b_{1,j} & \dots & b_{1,n} \\ \vdots & & & \vdots & & & \vdots \\ b_{i-1,1} & & & 0 & & & b_{i-1,n} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ b_{i,1} & & & 0 & & & b_{i,n} \\ \vdots & & & \vdots & & & \vdots \\ b_{n,1} & \dots & b_{n,j-1} & 0 & b_{n,j} & \dots & b_{n,n} \end{pmatrix}$$

d. h. füge in B vor der i-ten Zeile / j-ten Spalte eine Nullzeile/-spalte ein und setze den Eintrag auf der Position (i, j) gleich 1.

Lemma 6.16. *Sei $\alpha\colon M_{n,n}(R) \to R$ eine alternierende n-Form und $1 \leq i, j \leq n$. Dann ist*

$$\tilde{\alpha}_j\colon M_{n-1,n-1}(R) \to R, \qquad B \mapsto (-1)^{i+j}\alpha(\epsilon_{ij}(B))$$

eine alternierende $(n-1)$-Form und es gilt

(i) $\tilde{\alpha}_j$ *ist unabhängig von* i.

(ii) $\alpha(E_n) = \tilde{\alpha}_j(E_{n-1})$,

(iii) $\alpha(A) = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\tilde{\alpha}_j(A_{ij})$ *für* $A = (a_{k\ell}) \in M_{n,n}(R)$.

Beweis. $\tilde{\alpha}_j$ ist alternierende $(n-1)$-Form: folgt sofort aus den entsprechenden Eigenschaften von α.
Zu (i): Es gilt $P_{i,i+1}\epsilon_{ij}(B) = \epsilon_{i+1,j}(B)$, nach 6.14 also

$$(-1)^{i+1+j}\alpha(\epsilon_{i+1,j}(B)) = (-1)^{i+j}\alpha(\epsilon_{ij}(B)).$$

Zu (ii): Nach (i) dürfen wir annehmen, dass $i = j$. Nun gilt $\epsilon_{jj}(E_{n-1}) = E_n$ und $(-1)^{2j} = 1$.
Zu (iii): Sei A'_{ij} die Matrix, die aus A entsteht, indem die i-te Zeile durch e_j ersetzt wird. Aus der Linearität von α in der i-ten Zeile folgt

$$\alpha(A) = \sum_{j=1}^{n} a_{ij}\alpha(A'_{ij}).$$

Aus der Scherungsinvarianz von α folgt (räume j-te Spalte aus):

$$\alpha(A'_{ij}) = \alpha(\epsilon_{ij}(A_{ij})) = (-1)^{i+j}\tilde{\alpha}_j(A_{ij}).$$

□

Bemerkung. Gleich werden wir feststellen, dass $\tilde{\alpha}_j$ auch von j unabhängig ist.

Satz 6.17 (Charakterisierung der Determinante)**.** *Sei $\alpha\colon M_{n,n}(R) \to R$ eine alternierende n-Form. Dann gilt $\alpha = c\det$ mit $c = \alpha(E_n)$. Insbesondere gilt $\alpha = \det$ falls $\alpha(E_n) = 1$.*

Beweis. Induktion über n. Für $n = 1$ und $r \in R$ gilt $\alpha(r) = r\alpha(1) = cr = c\det(r)$
Sei $n > 1$ und die Behauptung für $n-1$ bewiesen. Nach Lemma 6.16 und Induktionsannahme gilt für $B \in M_{n-1,n-1}(R)$

$$\tilde{\alpha}_j(B) = c\det{}_{n-1}(B) = \widetilde{c(\det{}_n)}_j(B) \qquad (\text{wobei } \det{}_i\colon M_{i,i}(R) \to R)$$

und somit für $A \in M_{n,n}(R)$

$$\alpha(A) = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\tilde{\alpha}_j(A_{ij}) = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}c\det(A_{ij}) = c\det(A).$$

□

Korollar 6.18 (Entwicklungssatz von Laplace)**.** *Für $A = (a_{ij}) \in M_{n,n}(R)$ gilt*

$$\det A = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij}) \quad \textit{für } 1 \le j \le n \textit{ (Entwicklung nach der } j\textit{-ten Spalte)}$$

$$\det A = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij}) \quad \textit{für } 1 \le i \le n \textit{ (Entwicklung nach der } i\textit{-ten Zeile)}$$

Beweis. Die Entwicklung nach Spalten folgt wie eben, nach Zeilen wegen $\det A = \det A^t$ und $(A^t)_{ji} = (A_{ij})^t$. □

Satz 6.19 (Produktsatz)**.** *Sind $A, B \in M_{n,n}(R)$, so gilt*

$$\det(A \cdot B) = \det(A) \cdot \det(B).$$

Beweis. Für festes B betrachten wir die Abbildung

$$\alpha \colon M_{n,n}(R) \longrightarrow R,\ A \longmapsto \det(AB)$$

Seien $v_1, \ldots, v_n$ die Zeilen von A und $w_1, \ldots, w_n$ die Spalten von B. Dann gilt $AB = (v_i w_j)_{\substack{1 \le i \le n \\ 1 \le j \le n}}$. Gilt $v_k = v_\ell$, so ist auch in AB die k-te und ℓ-te Zeile gleich, also $\alpha(A) = \det(AB) = 0$. Ferner ist die Abbildung

$$R^n \xrightarrow{v \mapsto (vw_1, \ldots, vw_n)} R^n \xrightarrow{d_i} R$$

mit

$$d_i(v) = \det(\text{Ersetze } i\text{-te Zeile von } AB \text{ durch } v)$$

für $1 \le i \le n$ linear. Somit ist α eine alternierende n-Form und es gilt

$$\det(AB) = \alpha(A) = \alpha(E_n) \det(A) = \det(B) \det(A)$$

nach Satz 6.17. □

Definition 6.20. Die Matrix $\tilde{A} = (\tilde{a}_{ij})$ mit $\tilde{a}_{ij} = (-1)^{i+j} \det A_{ji}$ heißt die *Adjunkte* zu A.

Beispiel.

$$A = \begin{pmatrix} 3 & 5 \\ 1 & 3 \end{pmatrix} \quad \Rightarrow \quad \tilde{A} = \begin{pmatrix} 3 & -5 \\ -1 & 3 \end{pmatrix}$$

Es gilt

$$\tilde{A} \cdot A = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} = A \cdot \tilde{A} = \det(A) E_2$$

Das ist kein Zufall:

Satz 6.21 (Erste Cramersche Regel)**.** *Sei $A \in M_{n,n}(R)$. Dann gilt:*

$$\tilde{A} \cdot A = A \cdot \tilde{A} = \det(A) E_n.$$

Beweis. Sei $c_{ij}(B)$ der Koeffizient der Matrix B an der Stelle (i, j). Dann gilt

$$\begin{aligned} c_{ii}(A \cdot \tilde{A}) &= \sum_{j=1}^{n} a_{ij} \tilde{a}_{ji} = \sum_{j=1}^{n} a_{ij} (-1)^{i+j} \det A_{ij} \\ &= \det(A) \end{aligned}$$

Für $i \neq j$ ist

$$c_{ij}(A \cdot \tilde{A}) = \sum_{k=1}^{n} (-1)^{j+k} a_{ik} \det(A_{jk})$$

Rechts steht die Entwicklung nach der j-ten Zeile der Matrix, die man erhält, wenn man in A die j-te durch die i-te Zeile ersetzt, also die Determinante einer Matrix mit zwei gleichen Zeilen, d. h. 0.
Der Beweis von $\tilde{A} \cdot A = \det(A) E_n$ geht analog mit Spaltenentwicklung. □

Korollar 6.22. *$A \in M_{n,n}(R)$ ist genau dann invertierbar wenn $\det A \in R^\times$.*

Beweis. Ist A invertierbar, so gilt $1 = \det(E_n) = \det(A \cdot A^{-1}) = \det(A) \det(A^{-1})$, also $\det(A) \in R^\times$. Ist $\det(A) \in R^\times$, so gilt $\det(A)^{-1} \cdot \tilde{A} \cdot A = E_n$, also ist A invertierbar. □

Korollar 6.23. *Die Determinante induziert einen surjektiven Gruppenhomomorphismus*

$$\det\colon \mathrm{Gl}_n(R) \longrightarrow R^\times.$$

Beweis. $A \in \mathrm{Gl}_n(R) \Rightarrow \det(A) \in R^\times$ nach 6.22. Die Homomorphismuseigenschaft folgt aus 6.19. Surjektivität folgt aus

$$\det E_j(\lambda) = \lambda \det E_n = \lambda.$$

□

Korollar 6.24. *Sei $A \in M_{n,n}(\mathbb{Z})$. Dann existiert $A^{-1} \in M_{n,n}(\mathbb{Z})$ genau dann, wenn $|A| \in \{\pm 1\}$ gilt.*

Beweis. Es gilt $\mathbb{Z}^\times = \{\pm 1\}$. □

Satz 6.25 (Zweite Cramersche Regel). *Das lineare Gleichungssystem*

$$A \cdot x = b\ ,\ A \in \mathrm{Gl}_n(R),\ b \in R^n$$

hat die Lösung

$$x = \det(A)^{-1} \tilde{A} \cdot b.$$

Für die i-te Komponente x_i von x gilt daher

$$x_i = \det(A)^{-1} \sum_{j=1}^{n} (-1)^{i+j} \det(A_{ji}) b_j = \frac{\det(A, i, b)}{\det(A)}.$$

Hier ist (A, i, b) die Matrix, die aus A entsteht, wenn man die i-te Spalte durch b ersetzt.

Beweis. $A^{-1} = \det(A)^{-1} \tilde{A}$. □

Bemerkung. Diese Formel ist zur praktischen Berechnung ungeeignet, ist aber manchmal zu theoretischen Zwecken nützlich.

Definition 6.26. Die Gruppe

$$\mathrm{Sl}_n(R) \colon = \ker(\det : \mathrm{Gl}_n(R) \longrightarrow R^\times),$$

also die multiplikative Gruppe der (invertierbaren) Matrizen mit Determinante 1, heißt die *spezielle lineare Gruppe*.

Praktische Berechnungsregeln für det:

Ist $A \in M_{s,s}(R)$ und $k \in \mathbb{N}$

$$\left|\begin{array}{c|c} E_k & 0 \\ \hline B & A \end{array}\right| = |A| = \left|\begin{array}{c|c} E_k & B \\ \hline 0 & A \end{array}\right|$$

(Entwicklung nach der ersten Zeile und Induktion nach k).

Analog:

$$\left|\begin{array}{c|c} A & 0 \\ \hline B & E_k \end{array}\right| = |A| = \left|\begin{array}{c|c} E_k & B \\ \hline 0 & A \end{array}\right|$$

$$\left|\begin{array}{c|c} 0 & A \\ \hline E_k & B \end{array}\right| = (-1)^{sk}|A| = \left|\begin{array}{c|c} B & A \\ \hline E_k & 0 \end{array}\right|$$

Sind A und C quadratisch, so erhalten wir

$$\left|\begin{array}{c|c} A & 0 \\ \hline B & C \end{array}\right| = \left|\begin{array}{c|c} A & 0 \\ \hline B & E \end{array}\right| \cdot \left|\begin{array}{c|c} E & 0 \\ \hline 0 & C \end{array}\right| = |A| \cdot |C|$$

und durch Transponieren

$$\left|\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right| = |A| \cdot |C|.$$

Induktiv erhält man so für $\lambda_1 \dots \lambda_n \in R$

$$\left|\begin{array}{ccc} \lambda_1 & & 0 \\ & \ddots & \\ * & & \lambda_n \end{array}\right| = \lambda_1 \cdots \lambda_n = \left|\begin{array}{ccc} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{array}\right|$$

Also sollte man zur praktischen Berechnung von det die Matrix in Dreiecksgestalt bringen.

Beispiel.

$$\left|\begin{array}{ccc} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{array}\right| = \left|\begin{array}{ccc} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{array}\right| = \left|\begin{array}{ccc} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{array}\right| = 0$$

$$\left|\begin{array}{ccc} 1 & 2 & 3 \\ 4 & 8 & 6 \\ 1 & 3 & 3 \end{array}\right| = \left|\begin{array}{ccc} 1 & 2 & 3 \\ 0 & 0 & -6 \\ 0 & 1 & 0 \end{array}\right| = -\left|\begin{array}{ccc} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & -6 \end{array}\right| = 6$$

6.3 Ähnliche Matrizen

Sei K wieder ein Körper.

Definition 6.27. Zwei Matrizen $A, B \in M_{n,n}(K)$ heißen ähnlich, wenn es $T \in \mathrm{Gl}_n(K)$ gibt mit

$$B = TAT^{-1}$$

Bemerkungen.

- Änlichkeit ist eine Äquivalenzrelation auf $M_{n,n}(K)$.
- Sei $A = M_v^v(f)$ die Darstellungsmatrix eines Endomorphismus $f \in \mathrm{End}_K(V)$ eines n-dimensionalen Vektorraums bezüglich der Basis $v = (v_1, \ldots, v_n)$. Dann ist B genau dann ähnlich zu A, wenn es eine Basis $w = (w_1, \ldots, w_n)$ von V mit $B = M_w^w(f)$ gibt (siehe Basiswechselsatz 4.15).

Ziel:
Klassifikation aller Matrizen bis auf Ähnlichkeit, d. h. Beschreibung aller Ähnlichkeitsklassen von Matrizen.
Äquivalent dazu: Betrachte die Kategorie

Objekte: Paare (V, f) mit V endlichdimensionaler Vektorraum, $f \in \mathrm{End}_K(V)$

Morphismen: $\alpha\colon (V, f) \to (W, g)$ mit $\alpha\colon V \to W$ linear, $\alpha \circ f = g \circ \alpha$

Ziel: Klassifikation aller Objekte (V, f) bis auf Isomorphie.

Methode:
Finde ausreichend viele Invarianten, d. h. Abbildungen $c\colon M_{n,n}(K) \to K$ mit $c(TAT^{-1}) = c(A)$.

Sei $A \in M_{n,n}(K)$.

Definition 6.28. Die *Spur* von $A = (a_{ij})$ ist definiert durch

$$\mathrm{sp}(A) = \sum_{i=1}^{n} a_{ii}$$

Bemerkung. $\mathrm{sp}\colon M_{n,n}(K) \to K$ ist K-linear.

Lemma 6.29. $\mathrm{sp}(A \cdot B) = \mathrm{sp}(B \cdot A)$.

Beweis. Sei $c_{ij}(M)$ der Eintrag der Matrix M an der Position (i, j). Dann gilt

$$\begin{aligned} c_{ii}(A \cdot B) &= \sum_j c_{ij}(A) c_{ji}(B) \\ &= \sum_i c_{ji}(B) c_{ij}(A) = c_{jj}(B \cdot A). \end{aligned}$$

Also

$$\mathrm{sp}(A \cdot B) = \sum_i c_{ii}(AB) = \sum_j c_{jj}(BA) = \mathrm{sp}(B \cdot A)$$

□

Korollar 6.30. *Für $T \in \mathrm{Gl}_n(K)$ gilt*

$$\begin{aligned} \det(TAT^{-1}) &= \det(A) \\ \mathrm{sp}(TAT^{-1}) &= \mathrm{sp}(A) \end{aligned}$$

Beweis. Es gilt $\det(AB) = \det(BA)$ nach 6.19, sowie $\mathrm{sp}(AB) = \mathrm{sp}(BA)$ nach 6.29. Wir erhalten

$$\det(TAT^{-1}) = \det(T^{-1}TA) = \det(EA) = \det(A).$$

Das Argument für die Spur ist das gleiche. □

Im folgenden wollen wir noch weitere Invarianten definieren. Wir brauchen dazu ein paar Hilfsmittel.

6.4 Polynome

Sei R ein unitärer Ring.

Definition 6.31. Ein *Polynom* mit Koeffizienten in R ist ein Ausdruck

$$f = f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n, a_j \in R.$$

Das Zeichen x heißt die Unbestimmte oder Variable des Polynoms. Die Menge der Polynome mit Koeffizienten in R wird mit $R[x]$ bezeichnet.

Bemerkungen.

(i) Ein Polynom ist ein formaler Ausdruck, d. h. nichts weiter als die Familie seiner Koeffizienten $a_0, a_1, \ldots$, bzw. eine Abbildung $f\colon \mathbb{N}_0 \to R$ mit $f(i) = 0$ f.f.a. i. Insbesondere sind zwei Polynome genau dann gleich, wenn die Familien ihrer Koeffizienten gleich sind.

(ii) Jedes Polynom $f(x) = a_0 + \cdots + a_n x^n \in R[x]$ induziert eine Abbildung

$$R \longrightarrow R, \quad \alpha \mapsto f(\alpha) = a_0 + a_1\alpha + \cdots + a_n\alpha^n.$$

I. A. ist f durch diese Abbildung nicht eindeutig bestimmt. Z. B. induzieren die Polynome x^2 und $x \in \mathbb{Z}/2\mathbb{Z}[x]$ die gleiche Abbildung $\mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$, aber $x \neq x^2$.

(iii) $R[x]$ wird folgendermaßen zu einem unitären Ring:

Sei $f = \sum\limits_{i\in\mathbb{N}_0} a_i x^i$, $g = \sum\limits_{i\in\mathbb{N}_0} b_i x^i$. Dann setzt man

$$f + g = \sum_{i\in\mathbb{N}_0} (a_i + b_i) x^i$$

$$f \cdot g = \sum_{s\in\mathbb{N}_0} \left(\sum_{i=0}^{s} a_i \cdot b_{s-i} \right) x^s.$$

Ist R kommutativ, so auch $R[x]$.

(iv) Wir haben eine natürliche Inklusion von unitären Ringen $R \hookrightarrow R[x]$, die einem $\alpha \in R$ das „konstante" Polynom α (d.h. $a_0 = \alpha$, $a_i = 0$ für alle $i > 0$) zuordnet.

(v) Die Abbildung

$$M_{n,n}(R[x]) \xrightarrow{\sim} M_{n,n}(R)[x],$$
$$(f_{ij}) = \left(\sum_k a_{ijk} x^k\right)_{i,j\in\{1\dots n\}} \mapsto \sum_k (a_{ijk})_{i,j\in\{1\dots n\}}\, x^k.$$

ist ein Isomorphismus unitärer Ringe. Wir werden zwischen beiden Ringen nicht unterscheiden.

Definition 6.32. Ist $f = a_0 + a_1 x + \cdots + a_n x^n \in R[x]$ und $a_n \neq 0$, so heißt n der *Grad von* f (Notation: $\deg(f)$) und a_n der *Leitkoeffizient* (Notation: $a_n = e(f)$). Gilt $a_n = 1$, so nennt man f *normiert*.

Konvention: $\deg(0) = -\infty$, $e(0) = 0$.

Lemma 6.33. *Sei R kommutativ und nullteilerfrei, $f, g \in R[x]$.*

(i) $\deg(f \cdot g) = \deg(f) + \deg(g)$

(ii) *Es gilt* $\deg(f+g) \leq \max(\deg(f), \deg(g))$ *und wenn* $\deg(f) \neq \deg(g)$, *so gilt* $\deg(f+g) = \max(\deg(f), \deg(g))$

(iii) *Es gilt* $e(f \cdot g) = e(f) \cdot e(g)$.

Beweis. Für $f = 0$ oder $g = 0$: klar. Sei $f, g \neq 0$.
(i)+(iii): Es gilt

$$f = e(f)x^{\deg f} + \text{Terme niederer Ordnung.}$$

Analog für g. Also

$$fg = e(f)e(g)x^{\deg f + \deg g} + \text{T. n. O. .}$$

Wegen der Nullteilerfreiheit gilt $e(fg) = e(f)e(g) \neq 0$, also $\deg fg = \deg f + \deg g$.
(ii): Ähnlich. □

Korollar 6.34. *Ist R kommutativ und nullteilerfrei, so auch $R[x]$, d. h. aus $f \cdot g = 0$ folgt $f = 0$ oder $g = 0$.*

Beweis. Dies folgt aus $e(f \cdot g) = e(f) \cdot e(g) \neq 0$. □

Definition 6.35. Ein $\alpha \in R$ heißt *Nullstelle* oder *Wurzel* von $f \in R[x]$, wenn $f(\alpha) = 0$ gilt.

Lemma 6.36. *Ist $\alpha \in R$ eine Wurzel von $0 \neq f \in R[x]$, so gibt es ein $g \in R[x]$ mit $f = g \cdot (x - \alpha)$ und $\deg g = \deg f - 1$.*

Beweis. Sei $f = \sum_{i=0}^{n} a_i x^i \in R[x]$.
Fall $\alpha = 0$, d. h. $a_0 = f(0) = 0$: Setze

$$g = \sum_{i=0}^{n-1} a_{i+1} x^i.$$

Dann gilt $f(x) = g(x)x$ wie erwünscht.
Allgemeiner Fall: Für $\alpha \in R$ und $h = \sum_{i=0}^{n} b_i x^i \in R[x]$ setze

$$\begin{aligned} t_\alpha(h) = h(x + \alpha) &= \sum_{i=0}^{n} b_i (x + \alpha)^i \\ &\overset{\alpha x = x\alpha}{=} \sum_{i=0}^{n} \sum_{j=0}^{n} a_i \binom{i}{j} \alpha^{i-j} x^j \\ &= \sum_{j=0}^{n} \left(\sum_{j=0}^{n} a_i \binom{i}{j} \alpha^{i-j} \right) x^j \in R[x] \end{aligned}$$

wobei $\binom{i}{j} \in R$ die Binomialkoeffizienten bezeichnet (mit $\binom{i}{j} = 0$ für $j > i$). Es gilt

$$f(\alpha) = 0 \Rightarrow s_\alpha(f)(0) = 0 \Rightarrow s_\alpha(f) = \tilde{g} \cdot x \Rightarrow f = s_{-\alpha}(\tilde{g}) \cdot (x - \alpha)$$

mit $\deg s_{-\alpha}(\tilde{g}) = \deg \tilde{g} = \deg f - 1$. □

Korollar 6.37. *Sei R kommutativ und nullteilerfrei. Ein Polynom $f \neq 0$ vom Grad n hat höchstens n Wurzeln in R.*

Beweis. Sei α eine Wurzel, so schreiben wir $f = g \cdot (x - \alpha)$ mit $\deg(g) = n - 1$. Ist nun $\beta \neq \alpha$ eine weitere Wurzel, so gilt

$$0 = f(\beta) = g(\beta) \cdot (\beta - \alpha) \Rightarrow g(\beta) = 0.$$

Die Aussage folgt über vollständige Induktion nach dem Grad. □

6.5 Das charakteristische Polynom

Definition 6.38. Sei $A \in M_{n,n}(K)$. Das Polynom

$$\chi_A(t) \colon = \det(tE - A) \in K[t]$$

heißt das *charakteristische Polynom* der Matrix A.

Lemma 6.39. *χ_A ist ein normiertes Polynom vom Grad n*

$$\chi_A(t) = t^n + c_{n-1}(A)t^{n-1} + \cdots + c_0(A)$$

und es gilt $c_0(A) = \chi_A(0) = (-1)^n|A|$ und $c_{n-1}(A) = -\mathrm{sp}(A)$.

Beweis. Die Leibnizformel zeigt, dass χ_A die Form

$$\chi_A(t) = (t - a_{11}) \cdots (t - a_{nn}) + (\text{Polynom vom Grad } \leq n-2)$$

hat. Daher ist χ_A normiert vom Grad n und $c_{n-1} = -a_{11} - a_{22} \cdots - a_{nn}$. Schließlich gilt

$$c_0 = \chi_A(0) = |0 \cdot E - A| = |-A| = (-1)^n|A|.$$

□

Lemma 6.40. *Ist $T \in \mathrm{Gl}_n(K)$, so gilt*

$$\chi_{TAT^{-1}} = \chi_A.$$

Beweis.

$$\begin{aligned} \chi_{TAT^{-1}} &= |tE - TAT^{-1}| \\ &= |T(tE - A)T^{-1}| \\ &= |T| \cdot |tE - A| \cdot |T^{-1}| \\ &= \chi_A. \end{aligned}$$

□

Bemerkung. Insbesondere gilt $c_i(TAT^{-1}) = c_i(A)$ für alle Koeffizienten c_i des charakteristischen Polynoms.

Unsere Regeln zur Determinantenberechnung wenden sich nun hier an und wir erhalten:

- $A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \Rightarrow \quad tE - A = \begin{pmatrix} t-\lambda_1 & & * \\ & \ddots & \\ 0 & & t-\lambda_n \end{pmatrix}$

$\Longrightarrow \chi_A(t) = (t - \lambda_1) \cdots (t - \lambda_n)$.

- $A = \begin{pmatrix} B & * \\ 0 & C \end{pmatrix}$ oder $A = \begin{pmatrix} B & 0 \\ * & C \end{pmatrix}$

$\Longrightarrow \chi_A(t) = \chi_B(t) \cdot \chi_C(t)$.

Sei nun $f = c_0 + c_1 t + \cdots + c_r t^r \in K[t]$ ein Polynom. Wir können Matrizen $A \in M_{m,m}(K)$ in f einsetzen durch die Regel

$$f(A) = c_0 E + c_1 A + c_2 A^2 + \cdots + c_r A^r \in M_{n,n}(K).$$

und erhalten einen unitären Ringhomomorphismus

$$K[t] \to M_{m,m}(K), \qquad f \mapsto f(A).$$

Bemerkung. Insbesondere gilt für $f, g \in K[t]$ und $A \in M_{n,n}(K)$:

$$f(A) \cdot g(A) = (f \cdot g)(A) = (g \cdot f)(A) = g(A) \cdot f(A),$$

d. h. die Matrizen $f(A)$ und $g(A)$ *kommutieren*.

Satz 6.41 (Cayley-Hamilton). *Es gilt* $\chi_A(A) = 0$.

Beweis. Sei $D \in M_{n,n}(K[t])$ die Adjunkte zu $(tE - A)$ also

$$D \cdot (tE - A) = \det(tE - A) \cdot E = \chi_A(t) \cdot E \qquad (*)$$

in $M_{n,n}(K[t]) = M_{n,n}(K)[t]$. In der Definition der Adjunkten treten Determinanten von $(n-1) \times (n-1)$-Untermatrizen auf, also sind die Einträge von D Polynome in $K[t]$ vom Grad $\leq n-1$. Wir schreiben

$$D = \sum_{i=0}^{n-1} D_i t^i \quad D_i \in M_{n,n}(K).$$

Desweiteren sei $\chi_A(t) = \sum_{i=0}^{n} a_i t^i$, mit $a_i \in K$. Ein Koeffizientenvergleich in $(*)$ liefert

$$a_i E = D_{i-1} - D_i A$$

wobei wir $D_{-1} = 0$ und $D_n = 0$ ergänzen. Es folgt:

$$\chi_A(A) = \sum_{i=0}^{n} a_i A^i = \sum_{i=0}^{n} (D_{i-1} - D_i A) A^i$$
$$= -D_0 A + D_0 A - D_1 A^2 + \cdots + D_{n-1} A^n - D_n A^{n+1} = 0.$$

□

6.6 Endomorphismen

Sei V ein n-dimensionaler K-Vektorraum und $\alpha \colon V \to V$ ein Endomorphismus, $\alpha \in \mathrm{End}(V)$. Sei A die Darstellungsmatrix von α bzgl. einer Basis $(v_1, \ldots, v_n)$. Nach 6.40 hängt das charakteristische Polynom nicht von der Wahl der Basis ab und wir erhalten, dass die folgenden Objekte wohldefiniert sind:

Definition 6.42.

$$\mathrm{sp}(\alpha) \stackrel{\mathrm{df}}{=} \mathrm{sp}(A),$$
$$\det(\alpha) \stackrel{\mathrm{df}}{=} \det(A),$$
$$\chi_\alpha(t) \stackrel{\mathrm{df}}{=} \chi_A(t)$$

wobei A die α bzgl. irgendeiner Basis darstellende Matrix ist.

Definition 6.43. Sei $\alpha \in \mathrm{End}(V)$. Ein $\lambda \in K$ heißt *Eigenwert von* α, wenn es einen Vektor $v \in V$, $v \neq 0$, mit $\alpha(v) = \lambda \cdot v$ gibt. Ist λ ein Eigenwert von α, so heißt der Untervektorraum

$$V_\lambda = \ker(\lambda \cdot \mathrm{id}_V - \alpha)$$

der *Eigenraum* zu λ und seine Elemente $\neq 0$, d.h. solche $v \neq 0$ mit $\alpha(v) = \lambda \cdot v$ heißen *Eigenvektoren* zum Eigenwert λ.

Bemerkung. Ideal wäre es, wenn man V in die direkte Summe von Eigenräumen zerlegen könnte. Dann hätte α bezüglich einer Basis von V Diagonalgestalt. Leider geht das nicht immer.

Beispiel. $\alpha = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{Gl}_2(\mathbb{R})$ entspricht der Drehung um $\frac{\pi}{2}$ in der Ebene und hat keinen Eigenwert.

Satz 6.44. *Die Eigenwerte von* $\alpha \in \mathrm{End}_K(V)$ *sind genau die Nullstellen von* $\chi_\alpha(t)$ *in* K.

Beweis. Es sei α bezüglich irgendeiner Basis durch die Matrix A dargestellt:

$$\lambda \text{ EW von } \alpha \Leftrightarrow \exists v \neq 0 : \alpha(v) = \lambda(v) \quad \Leftrightarrow \exists\, v \neq 0\ (\lambda\,\mathrm{id}_V - \alpha)v = 0 \Leftrightarrow$$
$$\ker(\lambda\,\mathrm{id}_V - \alpha) \neq 0 \Leftrightarrow \det(\lambda\,\mathrm{id}_V - \alpha) = 0 \Leftrightarrow \det(\lambda E - A) = 0 \Leftrightarrow \chi_A(\lambda) = 0.$$

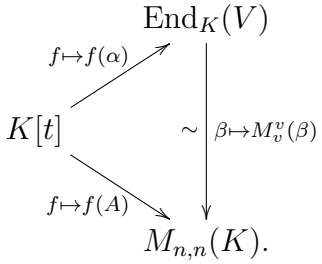
□

Bemerkung. Damit sehen wir auch algebraisch, dass die reelle Matrix $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ keinen Eigenwert hat, weil $\chi_A(t) = t^2+1$ keine reellen Nullstellen hat. Aber $\pm i$ sind komplexe Nullstellen, d.h. als komplexe 2×2-Matrix aufgefasst, besitzt A zwei Eigenwerte.

Sei nun $f = c_0 + c_1 t + \cdots + c_r t^r \in K[t]$, V ein n-dimensionaler K-Vektorraum und $\alpha \in \mathrm{End}(V)$. Wir setzen

$$f(\alpha) = c_0 \cdot \mathrm{id}_V + c_1\alpha + c_2\alpha \circ \alpha + \cdots + c_r \underbrace{\alpha \circ \alpha \cdots \circ \alpha}_{r-\text{mal}} \in \mathrm{End}(V).$$

Sei $A = M_v^v(\alpha)$ die Darstellungsmatrix von α bezüglich einer Basis $v = (v_1, \ldots, v_n)$ von V. Dann gilt $M_v^v(f(\alpha)) = f(A)$, d. h. wir erhalten ein kommutatives Diagramm unitärer Ringhomomorphismen



Satz 6.45 (Cayley-Hamilton für Endomorphismen).

$$\chi_\alpha(\alpha) = 0.$$

Beweis. Sei α bzgl. einer Basis durch die Matrix A dargestellt. Dann wird $\chi_\alpha(\alpha)$ durch $\chi_A(A) = 0$ (Cayley-Hamilton) dargestellt. □

6.7 Zerlegung in Eigenräume

Sei V ein n-dimensionaler K-Vektorraum und $\alpha \in \mathrm{End}(V)$.

Definition 6.46. Man sagt, eine Matrix $A = (a_{ij})$ habe *Diagonalgestalt*, wenn $a_{ij} = 0$ für $i \neq j$ gilt. Man schreibt dann

$$A = \mathrm{diag}(a_{11}, \ldots, a_{nn}).$$

Der Endomorphismus α von V heißt *diagonalisierbar*, wenn die Darstellungsmatrix von α bezüglich einer Basis $(v_1, \ldots, v_n)$ von Diagonalgestalt ist.

Lemma 6.47. *Der Endomorphismus α ist genau dann diagonalisierbar, wenn es eine Basis $(v_1, \ldots, v_n)$ von V bestehend aus Eigenvektoren zu α gibt.*

Beweis. Es ist $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ genau dann die Darstellungsmatrix von α bezüglich der Basis $(v_1, \ldots, v_n)$, wenn $\alpha(v_i) = \lambda_i v_i$ gilt. □

Bemerkung. Ist α diagonalisierbar, also die Darstellungsmatrix von α bezüglich einer Basis $(v_1, \ldots, v_n)$ von Diagonalgestalt $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, so gilt

$$\chi_\alpha(t) = \det\left(\begin{pmatrix} t-\lambda_1 & & 0 \\ & \ddots & \\ 0 & & t-\lambda_n \end{pmatrix}\right) = (t-\lambda_1)\cdots(t-\lambda_n),$$

d.h. χ_α zerfällt in das Produkt von Linearfaktoren.

Satz 6.48. *Es seien $\lambda_1, \ldots, \lambda_m$ paarweise verschiedene Eigenwerte von α und $v_1, \ldots, v_n \in V$ Eigenvektoren zu $\lambda_1, \ldots, \lambda_n$. Dann ist das System*

$$(v_1, \ldots, v_m)$$

linear unabhängig.

Beweis. Nach Voraussetzung gilt $(\alpha - \lambda_i \,\mathrm{id}_V)(v_j) = (\lambda_j - \lambda_i)v_j$. Setzt man

$$\beta_i = (\alpha - \lambda_1 \,\mathrm{id}_V) \circ \cdots \circ (\alpha - \lambda_{i-1} \,\mathrm{id}_V) \circ (\alpha - \lambda_{i+1} \,\mathrm{id}_V) \circ \cdots \circ (\alpha - \lambda_m \,\mathrm{id}_V),$$

so folgt

$$\beta_i(v_j) = \Big(\prod_{k \neq i}(\lambda_j - \lambda_k)\Big) \cdot v_j = \begin{cases} 0 & i \neq j \\ (\text{Skalar} \neq 0) \cdot v_j & i = j. \end{cases}$$

Gilt nun

$$a_1v_1 + \cdots + a_mv_m = 0, \quad a_1, \ldots, a_m \in K,$$

so erhält man für jedes $i = 1, \ldots, m$ durch Anwendung von β_i die Gleichung $a_i = 0$. Daher ist das System $(v_1, \ldots, v_m)$ linear unabhängig. □

Satz 6.49. *Sei V ein n-dimensionaler K-Vektorraum und $\alpha \in \mathrm{End}(V)$. Zerfällt das charakteristische Polynom von α in paarweise verschiedene Linearfaktoren, d.h.*

$$\chi_\alpha(t) = (t - \lambda_1) \ldots (t - \lambda_n)$$

mit $\lambda_i \neq \lambda_j$ für $i \neq j$, so gibt es eine Basis von V aus Eigenvektoren von α. Insbesondere wird α bezüglich einer Basis von V durch eine Diagonalmatrix dargestellt.

Beweis. In diesem Fall sind $\lambda_1, \ldots, \lambda_n$ paarweise verschiedene Eigenwerte. Sind $v_1, \ldots, v_n$ assoziierte Eigenvektoren, so ist $(v_1, \ldots, v_n)$ nach 6.48 ein linear unabhängiges System und wegen $n = \dim V$ eine Basis. □

Wie macht man das explizit?
Betrachte den Endomorphismus α des $\mathbb{R}^2$ der bzgl. der kanonischen Basis durch die Matrix $A = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}$ gegeben ist.

$$\begin{aligned} \chi_\alpha(t) = \det \begin{pmatrix} t-1, 2 \\ -1, t-4 \end{pmatrix} &= (t-1)(t-4) + 2 \\ &= t^2 - 5t + 6 \\ &= (t-2)(t-3) \end{aligned}$$

Suche Eigenvektoren. Betrachte das homogene lineare Gleichungssystem

$$\begin{array}{c} (2E - A)x = 0 \\ \| \\ \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \end{array}$$

Nichttriviale Lösung: $(2, -1)^t$ ist Eigenvektor zum Eigenwert $\lambda = 2$.
Probe: $\begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ -2 \end{pmatrix}$

$$\begin{array}{c} (3E - A)x = 0 \\ \| \\ \begin{pmatrix} 2 & 2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \end{array}$$

Lösung: $(-1, 1)^t$

Probe: $\begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \end{pmatrix}$.

Also hat α bzgl. der Basis $((2,-1)^t,(-1,1)^t)$ die Darstellungsmatrix $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$

Testrechnung: Der Basiswechsel von$((1,0)^t,(0,1)^t)$ zu $((2,-1)^t,(-1,1)^t)$ ist durch die Transformationsmatrix $T = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}^{-1}$ gegeben.

berechnen:
$$\left(\begin{array}{cc|cc} 2 & -1 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & 1 & 1 \\ -1 & 1 & 0 & 1 \end{array}\right)$$
$$\rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{array}\right) \rightsquigarrow T = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

Wenden wir den Basiswechselsatz an, so müßten wir erhalten: $TAT^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$.

Probe:
$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 3 & 6 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \quad \text{o.k.}$$

6.8 Trigonalisierbarkeit

Definition 6.50. Sei V ein n-dimensionaler K-Vektorraum und $\alpha \in \mathrm{End}(V)$. Der Endomorphismus α heißt *trigonalisierbar*, wenn es eine Basis gibt bzgl. derer α durch eine obere Dreiecksmatrix, d.h. durch eine Matrix der Form

$$\begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix}$$

dargestellt wird.

Satz 6.51. *α ist genau dann trigonalisierbar, wenn $\chi_\alpha(t)$ vollständig in Linearfaktoren zerfällt.*

Beweis. Ist α bzgl. einer Basis durch $\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ gegeben, so gilt

$$\chi_\alpha(t) = \det(\begin{pmatrix} t-\lambda_1 & & * \\ & \ddots & \\ 0 & & t-\lambda_n \end{pmatrix}) = (t-\lambda_1)\cdots(t-\lambda_n),$$

Die andere Richtung beweisen wir per Induktion nach n. Der Fall $n = 1$ ist klar, Sei $\chi_\alpha = (t-\lambda_1)\ldots(t-\lambda_n)$ und sei v_1 ein Eigenvektor zu λ_1. Wir ergänzen v_1 zu einer Basis $(v_1,\ldots,v_n)$ von V. Bzgl. dieser Basis hat α die Gestalt $A =$

$\left(\begin{array}{c|c} \lambda_1 & * \\ \hline 0 & A' \end{array}\right)$ mit einer $(n-1)\times(n-1)$ Matrix A'. Sei $V' = \mathrm{Lin}(v_2,\ldots,v_n)$. Dann gilt

$$V = K\cdot v_1 \oplus V'.$$

Außerdem gilt $\chi_A(t) = (t-\lambda_1)\cdot\chi_{A'}(t)$, d. h.

$$(t-\lambda_1)(\chi_{A'}(t) - (t-\lambda_2)\cdots(t-\lambda_n)) = 0$$

Da $K[t]$ nullteilerfrei ist, folgt

$$\chi_{A'}(t) = (t-\lambda_2)\cdots(t-\lambda_n),$$

also zerfällt auch $\chi_{A'}(t)$ in Linearfaktoren. Sei α' der auf V' bzgl. der Basis $(v_2,\ldots,v_n)$ durch A' dargestellte Endomorphismus. Nach Induktionsvoraussetzung gibt es eine Basis $(v_2',\ldots,v_n')$ von V' bzgl. derer α' durch eine obere Dreiecksmatrix B' dargestellt wird. Dann wird α bzgl. der Basis $(v_1, v_2',\ldots,v_n')$ von V durch die Matrix $\left(\begin{array}{c|c} \lambda_1 & * \\ \hline 0 & B' \end{array}\right)$ dargestellt, ist also trigonalisierbar. □

Korollar 6.52. *Über $K = \mathbb{C}$ ist jeder Endomorphismus eines endlich-dimensionalen Vektorraums trigonalisierbar.*

Beweis. Über $\mathbb{C}$ zerfällt jedes Polynom in Linearfaktoren (Hauptsatz der Algebra). □

Sei nun K wieder allgemein und λ ein Eigenwert von $\alpha \in \mathrm{End}(V)$.

Definition 6.53.

(i) Die *algebraische Vielfachheit* $\mu_{\mathrm{alg}}(\lambda)$ ist die Vielfachheit von λ als Nullstelle von $\chi_\alpha(t)$, d. h. $\chi_\alpha(t) = g(t)(t-\lambda)^{\mu_{\mathrm{alg}}(\lambda)}$ mit $g(\lambda)\neq 0$.

(ii) Die *geometrische Vielfachheit* $\mu_{\mathrm{geo}}(\lambda)$ ist gleich $\dim V_\lambda$.

Satz 6.54. *Es gilt*

$$\mu_{\mathrm{geo}}(\lambda) \le \mu_{\mathrm{alg}}(\lambda).$$

Beweis. Sei $r = \mu_{\mathrm{geo}}(\lambda)$ und $v_1,\ldots,v_r$ eine Basis von V_λ. Ergänzen wir zu einer Basis $(v_1,\ldots,v_r,v_{r+1},\ldots,v_n)$, so wird α durch eine Matrix der Form $A = \left(\begin{array}{c|c} \lambda E_r & * \\ \hline 0 & A' \end{array}\right)$ dargestellt. Also gilt $\chi_\alpha(t) = \chi_A(t) = (t-\lambda)^r\cdot\chi_{A'}(t)$. Dies impliziert $\mu_{\mathrm{alg}}(\lambda) \ge r$. □

Satz 6.55. *Für einen Endomorphismus α auf dem n-dimensionalen K-Vektorraum V gilt*

$$\alpha \text{ diagonalisierbar} \iff \sum_{\lambda \text{ EW von } \alpha} \mu_{\mathrm{geo}}(\lambda) = n$$

$$\alpha \text{ trigonalisierbar} \iff \sum_{\lambda \text{ EW von } \alpha} \mu_{\mathrm{alg}}(\lambda) = n.$$

Beweis. Es seien $\lambda_1, \ldots, \lambda_r$ die (verschiedenen) Eigenwerte von α. Gilt

$$\sum_{i=1}^{r} \mu_{\text{alg}}(\lambda) = n = \deg \chi_\alpha(t),$$

so folgt

$$\chi_\alpha(t) = (t-\lambda_1)^{\mu_{\text{alg}}(\lambda_1)} \cdots (t-\lambda_r)^{\mu_{\text{alg}}(\lambda_r)},$$

und nach 6.51 ist α trigonalisierbar. Ist umgekehrt α trigonalisierbar, so zerfällt χ_α nach 6.51 in Linearfaktoren und es folgt $\sum_{\lambda\ EW} \mu_{\text{alg}}(\lambda) = n$.

Ist α diagonalisierbar, so zerfällt V in die direkte Summe der Eigenräume, daher gilt

$$n = \dim V = \sum_{\lambda\ EW} \dim V_\lambda = \sum_{\lambda\ EW} \mu_{\text{geo}}(\lambda).$$

Gelte umgekehrt diese Formel. Es seien $\lambda_1, \ldots, \lambda_r$ die (verschiedenen) Eigenwerte von α. Wir betrachten den natürlichen Homomorphismus

$$\phi : \bigoplus_{i=1}^{r} V_{\lambda_r} \longrightarrow V,\ (v_1, \ldots, v_r) \longmapsto v_1 + \ldots + v_r$$

Wir zeigen, dass ϕ ein Isomorphismus ist. Nach Voraussetzung haben Quelle und Ziel die gleiche Dimension, also gzz., dass ϕ injektiv ist. Dies folgt aus 6.48. Mit Hilfe des Isomorphismus ϕ erhalten wir eine Basis von V aus Eigenvektoren von α, also ist α diagonalisierbar nach 6.47. □

6.9 Der euklidische Algorithmus

Sei K ein Körper.

Satz 6.56 (Division mit Rest)**.** *Seien $f, g \in K[x]$, $g \neq 0$. Dann existieren eindeutig bestimmte Polynome $q, r \in K[x]$ mit*

$$f = q \cdot g + r\ ,\ \deg(r) < \deg(g).$$

Das Polynom r heißt der Rest der Division von f durch g.

Beweis.
Eindeutigkeit: Sei $f = q_1 \cdot g + r_1 = q_2 \cdot g + r_2$. Dann gilt

$$(q_1 - q_2) \cdot g = r_2 - r_1.$$

Wegen $\deg(r_2 - r_1) < \deg(g)$ folgt $q_1 - q_2 = 0$ also auch $r_2 - r_1 = 0$.
Existenz: Per Induktion nach $\deg(f)$. Ist $\deg(f) < \deg(g)$, setze $q = 0$, $f = r$. Sonst sei

$$f = ax^{n+k} + \text{niedrigere Potenzen}, g \qquad = bx^n + \text{niedrigere Potenzen},$$

$n, k \geq 0$, $a, b \neq 0$. Dann gilt

$$\deg\left(f - \frac{a}{b}x^k g\right) < \deg f.$$

Nach Induktionsannahme gibt es q_1, r_1 mit

$$f - \frac{a}{b}x^k g = q_1 \cdot g + r_1, \quad \deg(r_1) < \deg(g).$$

Wir erhalten die Darstellung

$$f = \left(q_1 + \frac{a}{b}x^k\right) \cdot g + r_1$$

□

Beispiel. $f = x^5 + 3x^4 - 4x^3 + 2$, $g = x^2 + 1$. Man findet g und r wie beim schriftlichen Dividieren ganzer Zahlen.

$$\begin{array}{l}
x^5 + 3x^4 - 4x^3 + 0x^2 + 0x + 2 \colon (x^2 + 1) = x^3 + 3x^2 - 5x - 3 \\
\underline{x^5 + x^3} \\
\quad 3x^4 - 5x^3 + 0x^2 + 0x + 2 \\
\quad \underline{3x^4 + 3x^2} \\
\quad\quad -5x^3 - 3x^2 + 0x + 2 \\
\quad\quad \underline{-5x^3 - 5x} \\
\quad\quad\quad -3x^2 + 5x + 2 \\
\quad\quad\quad \underline{-3x^2 - 3} \\
\quad\quad\quad\quad 5x + 5, \quad \text{Rest.}
\end{array}$$

Seien $f_1, f_2 \in K[x]$ beide nicht 0. Der *euklidische Algorithmus* ist die Folge von Divisionen mit Rest

$$\begin{array}{rl}
f_1 = q_1 f_2 + f_3 & \deg f_3 < \deg f_2 \\
f_2 = q_2 f_3 + f_4 & \deg f_4 < \deg f_3 \\
\vdots & \\
f_{n-1} = q_{n-1} f_n + 0 & \text{(weil die Grade abnehmen).}
\end{array}$$

Definition 6.57. Wir sagen f *teilt* g (Notation: $(f \mid g)$) wenn $g = f \cdot h$ für ein Polynom h gilt.

Satz 6.58. *Seien $f_1, f_2 \in K[x]$ beide $\neq 0$ und $d = f_n$ wie im euklidischen Algorithmus. Dann gilt*

(i) $d \mid f_1$, $d \mid f_2$

(ii) Aus $g \mid f_1$ und $g \mid f_2$ folgt $g|d$

(iii) Es gibt Polynome $p, q \in K[x]$ mit

$$pf_1 + qf_2 = d.$$

Beweis.
(i) Steige den Algorithmus aufwärts:
$d|f_n \Rightarrow d|f_{n-1} \quad f_{n-2} = qf_{n-1} + f_n \Rightarrow d|f_{n-2}$ usw. $\Rightarrow d|f_2$, $d|f_1$
(ii) Steige ab: $g|f_1$, $g|f_2 \Rightarrow g|f_3 \Rightarrow \cdots \Rightarrow g|f_n = d$.
(iii) Den Algorithmus absteigend sieht man, dass jedes f_k eine Darstellung der Form $f_k = p_k f_1 + q_k f_2$ hat. □

Definition 6.59. Ein Polynom d mit den Eigenschaften (i) und (ii) aus 6.58 heißt *ein größter gemeinsamer Teiler* von f_1 und f_2.

Satz 6.60. *Sind $f_1, f_2 \in K[x]$ beide nicht 0, so existiert ein größter gemeinsamer Teiler. Dieser ist bis auf einen konstanten Faktor $\neq 0$ eindeutig bestimmt.*

Beweis. Die Existenz folgt aus 6.58. Sind d_1, d_2 beide g.g.T., so gilt $d_1|d_2$ und $d_2|d_1$, also $\deg(d_1) \leq \deg(d_2) \leq \deg(d_1) \Rightarrow \deg d_1 = \deg d_2$ und aus $d_1|d_2$ folgt $d_2 = \alpha \cdot d_1$, $\alpha \in K$, $\alpha \neq 0$. □

Bemerkung. Unter allen größten gemeinsamen Teilern gibt es genau ein normiertes Polynom. Dies nennt man *den* größten gemeinsamen Teiler ($ggT(f_1, f_2)$). Dieser kann in der Form

$$ggT(f_1, f_2) = pf_1 + qf_2$$

dargestellt werden.

Korollar 6.61. *Haben die nicht verschwindenden Polynome f und g nur konstante gemeinsame Teiler, so existieren Polynome p, q mit*

$$pf + qg = 1.$$

Beweis. $ggT(f, g) = 1$ □

Definition 6.62. Ein Polynom f, $\deg(f) \geq 1$, heißt *irreduzibel*, wenn aus $f = g \cdot h$ folgt, dass g oder h konstant ist. Ansonsten heißt f reduzibel.

Beispiele.

(i) Jedes Polynom $ax + b$, $a \neq 0$, vom Grad 1 ist irreduzibel. Über $\mathbb{C}$ hat jedes irreduzible Polynom Grad 1 (Hauptsatz der Algebra).

(ii) Sei $\lambda \in \mathbb{C}$ nicht reell. Dann ist

$$(x - \lambda)(x - \bar{\lambda}) = x^2 - 2\,\mathrm{Re}(\lambda)x + |\lambda|^2$$

irreduzibel als Polynom in $\mathbb{R}[x]$, aber reduzibel als Polynom in $\mathbb{C}[x]$.

(iii) Jedes Polynom $f \in \mathbb{R}[x]$ vom Grad $\deg f > 2$ ist reduzibel, denn entweder hat f eine reelle Nullstelle $\lambda \in \mathbb{R}$ und $(x-\lambda) \mid f$ oder $f = \sum a_i x^i$ hat eine komplexe Nullstelle $\lambda \neq \bar{\lambda}$. In diesem Fall gilt

$$0 = f(\lambda) = \overline{f(\lambda)} = \sum a_i \bar{\lambda}^i = f(\bar{\lambda}),$$

d. h. auch $\bar{\lambda}$ ist eine Nullstelle und $(x^2 - 2\,\mathrm{Re}(\lambda)x + |\lambda|^2) \mid f$.

(iv) $x^3+2 \in \mathbb{Q}[x]$ ist irreduzibel, aber $x^3+2 = (x-1)^3 \in \mathbb{Z}/3\mathbb{Z}[x]$ ist reduzibel.

Lemma 6.63. *Ist f irreduzibel und $f \mid (g \cdot h)$, so gilt $f \mid g$ oder $f \mid h$.*

Beweis. Da f irreduzibel ist, sind die Polynome der Form a und af, $a \in K$, $a \neq 0$, die einzigen Teiler von f. Gilt nun $f \nmid g$, so folgt $ggT(f,g) = 1$. Also existieren $p, q \in K[x]$ mit $p \cdot f + q \cdot g = 1 \Rightarrow p \cdot f \cdot h + q \cdot g \cdot h = h \Rightarrow f|h$. □

Bemerkung. Ein nicht konstantes Polynom f, so dass $f \mid (g \cdot h) \Rightarrow (f \mid g) \vee (f \mid h)$, nennt man auch *prim*. Das Lemma zeigt, dass in $K[x]$ wie in $\mathbb{Z}$ die Begriffe „irreduzibel" und „prim" äquivalent sind.

Satz 6.64. *Jedes Polynom $f \neq 0$ besitzt eine bis auf die Reihenfolge der Faktoren eindeutige "Primfaktorzerlegung":*

$$f = a \cdot p_1 \dots p_k\ ,\ a = e(f), e(p_i) = 1,\ i = 1, \dots, k$$

mit irreduziblen normierten Faktoren p_i.

Beweis.
Existenz: Induktion nach $\deg f$. Sei $\deg f = 0$. Setze $k = 0$, $a = f$. Sei $\deg(f) = n$. Ist f irreduzibel, so schreibt man $f = e(f)e(f)^{-1}f$ und $e(f)^{-1}f$ ist normiert und irreduzibel. Ansonsten gilt $f = g \cdot h$ mit $\deg(g), \deg(h) < n$ und die Induktionsvoraussetzung für g und h liefert auch eine Darstellung für f.

Eindeutigkeit: $a = e(f)$ ist eindeutig. Angenommen es gilt $p_1 \dots p_k = q_1 \dots q_\ell$ mit irreduziblen, normierten Polynomen $p_1, \dots, p_k, q_1, \dots, q_\ell$. Zu zeigen:
$k = \ell$ und $p_1, \dots, p_k$, und $q_1, \dots, q_\ell$ unterscheiden sich nur in der Reihenfolge.

Aus 6.63 erhalten wir $p_k | q_i$ für ein $i,\ 1 \leq i \leq \ell$. O.B.d.A. $i = \ell$. Da q_ℓ irreduzibel ist, folgt $p_k = q_\ell$. Also

$$p_k(p_1 \dots p_{k-1} - q_1 \dots q_{\ell-1}) = 0$$

Aus 6.34 folgt: $p_1 \dots p_{k-1} = q_1 \dots q_{\ell-1}$. Dann weiter ... □

Bemerkung. Der Beweis, dass in $\mathbb{Z}$ jedes Element eine eindeutige Primfaktorzerlegung $a = (\pm 1) \cdot p_1 \dots p_k$ hat geht genauso.

Der Polynomring $K[t]$ ist Unterring eines Körpers, nämlich des Körpers $K(t)$ *der rationalen Funktionen*. Der Übergang ist der gleiche wie der von $\mathbb{Z}$ zu $\mathbb{Q}$:

Wir betrachten die Äquivalenzrelation auf $K[t] \times (K[t] \setminus \{0\})$, die gegeben ist durch

$$(f_1, g_1) \sim (f_2, g_2) \Leftrightarrow f_1 g_2 = f_2 g_1.$$

(Übungsaufgabe: man verifiziere (Ä1)–(Ä3)). Die Äquivalenzklasse eines Paares (f, g), $g \neq 0$, wird mit

$$\frac{f}{g}$$

bezeichnet.

Definition 6.65. Die Äquivalenzklassen heißen *rationale Funktionen*. Die Menge der rationalen Funktionen wird mit $K(t)$ bezeichnet.

Rechnen tut man, wie man es mit Brüchen gewöhnt ist:

$$\frac{f_1}{g_1} \cdot \frac{f_2}{g_2} = \frac{f_1 f_2}{g_1 g_2}$$

$$\frac{f_1}{g_1} + \frac{f_2}{g_2} = \frac{f_1 g_2 + f_2 g_1}{g_1 g_2}.$$

(Übungsaufgabe: man verifiziere die Wohldefiniertheit).

Satz 6.66. *$K(t)$ ist ein kommutativer Ring mit Nullelement $0/1$ und Einselement $1/1$. Es gilt $f/g = 0/1$ genau dann wenn $f = 0$ gilt. Daher hat $f/g \neq 0/1$ das Inverse g/f, weshalb $K(t)$ ein Körper ist. Die Zuordnung*

$$K[t] \longrightarrow K(t), \quad f \longmapsto \frac{f}{1}$$

ist ein injektiver Ringhomomorphismus („bettet $K[t]$ in $K(t)$ ein“).

Beweis. Klar. □

Satz 6.67 (Partialbruchzerlegung)**.** *Sei $f, g \in K[t]$ mit $ggT(f, g) = 1$ und g normiert. Sei $g = p_1^{v_1} \dots p_m^{v_m}$ die Primfaktorzerlegung von g mit paarweise verschiedenen normierten irreduziblen p_i vom Grad n_i. Dann gibt es eindeutige*

$$h, a_{11}, \dots, a_{1v_1}, a_{21}, \dots, a_{2v_2}, \dots, a_{mv_m} \in K[t]$$

mit $\deg a_{ij} < n_i$, so dass

$$\frac{f}{g} = h + \sum_{i=1}^{m} \sum_{j=1}^{v_i} \frac{a_{ij}}{p_i^j}$$

Beweis.
Existenz: Induktion nach m und v_i. Für $m = 0$ ist $g = 1$ und $h = f$. Für $m = 1$, $v_1 = 1$ gilt $g = p_1$ und $f = hg + a$ mit $\deg a < n_1$, also $\frac{f}{g} = h + \frac{a}{p_1}$. Sei nun $g = p_1^{v_1}$ mit $v_1 > 1$. Nach Induktionsannahme gilt

$$\frac{p_1 f}{g} = \frac{f}{p_1^{v_1-1}} = h' + \sum_{j=1}^{v_i-1} \frac{a_{1,j+1}}{p_1^j}$$

mit $\deg a_{1j} < n_1$. Schreibe $h' = hp_1 + a_{11}$, $\deg a_{11} < n_1$. Dann gilt

$$\frac{f}{g} = h + \sum_{j=1}^{v_i} \frac{a_{1j}}{p_1^j}.$$

Sei nun $m > 1$ und $g' = \frac{g}{p_1^{v_1}}$. Dann ist $ggT(p_1^{v_1}, g') = 1$. Also existieren $q, r \in K[t]$ mit $1 = qg' + rp_1^{v_1}$ und somit

$$\frac{f}{g} = \frac{fq}{p_1^{v_1}} + \frac{fr}{g'}$$

Wende nun die Induktionsannahme auf $\frac{fq}{p_1^{v_1}}$ und $\frac{fr}{g'}$ an.
Eindeutigkeit: Induktion über die Anzahl der Primfaktoren. Falls $m = 0$: klar. Falls $m = 1$, $v_1 = 1$: Eindeutigkeit der Division mit Rest 6.56. Ansonsten seien $h', a'_{ij} \in K[x]$ weitere Polynome mit den geforderten Eigenschaften. Dann gilt

$$p_1 | \frac{a_{ij} g}{p_i^j}, \quad p_1 | \frac{a'_{ij} g}{p_i^j} \qquad \text{für } (i,j) \neq (1, v_1)$$

also

$$p_1 | f - \frac{a_{1v_1}}{p_1^{v_1}}, \quad p_1 | f - \frac{a'_{1v_1}}{p_1^{v_1}}, \qquad \text{d. h. } p_1 | (a_{1v_1} - a'_{1v_1}) \frac{g}{p_1^{v_1}}.$$

Wegen $p_1 \nmid \frac{g}{p_1^{v_1}}$ folgt $p_1 | a_{1v_1} - a'_{1v_1}$, wegen $\deg a_{1v_1} < \deg p_1$, $\deg a'_{1v_1} < \deg p_1$ folgt $a_{1v_1} = a'_{1v_1}$. Setze $f' = \frac{f - \frac{a_{1v_1}}{p_1^{v_1}}}{p_1}$, $g' = \frac{g}{p_1}$. Nach Induktionsannahme ist die Darstellung

$$\frac{f'}{g'} = h + \sum_{j=1}^{v_i-1} \frac{a_{1j}}{p_1^j} + \sum_{i=2}^{m} \sum_{j=1}^{v_i} \frac{a_{ij}}{p_i^j}$$

eindeutig, d. h. $h = h'$ und $a_{ij} = a'_{ij}$ für alle i, j. □

Konkrete Berechnung:

Das Polynom h läßt sich aus der Division mit Rest von f durch g bestimmen. Wir können also annehmen, dass $\deg f < \deg g$ und $h = 0$. Schreibe $a_{ij} = \sum_{k=1}^{n_i} z_{ijk} x^k$. Durch Koeffizientenvergleich in

$$f = \sum_{i=1}^{n} \sum_{j=1}^{v_i} \frac{a_{ij} g}{p_i^j} \in K[t]$$

erhält man ein lineares Gleichungssystem in den Unbestimmten z_{ijk}, welches nach Satz 6.67 eindeutig lösbar ist.

Beispiel. Wir bestimmen die Partialbruchzerlegung von $\frac{x^2-2}{(x^2+1)(x+2)}$ in $\mathbb{R}[x]$. Koeffizientenvergleich in

$$x^2 - 2 = (ax + b)(x + 2) + c(x^2 + 1) = (a + c)x^2 + (2a + b)x + (2b + c)$$

ergibt das LGS in den Unbestimmten (a, b, c) mit erweiterter Koeffizientenmatrix

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

Anwenden des Gauß-Verfahrens:

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & -2 & -2 \\ 0 & 2 & 1 & -2 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 5 & 2 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{3}{5} \\ 0 & 1 & 0 & -\frac{6}{5} \\ 0 & 0 & 1 & \frac{2}{5} \end{array}\right)$$

Also gilt

$$\frac{x^2 - 2}{(x^2 + 1)(x + 2)} = \frac{1}{5}\left(\frac{3x - 6}{x^2 + 1} + \frac{2}{x + 2}\right).$$

Beispiel. Seien $\lambda_1, \ldots, \lambda_m \in K$ paarweise verschieden, $g = \prod_{i=1}^m (t - \lambda_i)^{v_i}$. Dann gilt

$$\frac{1}{g} = \sum_{i=1}^m \frac{a_i}{(t - \lambda_i)^{v_i}} \quad \text{mit} \quad a_i = \prod_{j \neq i} (\lambda_i - \lambda_j)^{-v_j}$$

Begründung: Multiplikation der Gleichung mit $\prod_{i=1}^m (t - \lambda_i)^{v_i}$ und Einsetzen von $\lambda_1, \ldots, \lambda_m$.

Sei $\deg f < \deg g$ teilerfremd zu g und $f = \sum_{j \in \mathbb{N}_0} b_{ij}(t - \lambda_i)^j$, $b_{ij} \in K$. Dann gilt

$$\frac{f}{g} = \sum_{i=1}^m \sum_{j=1}^{v_i} \frac{a_i b_{i,v_i - j}}{(t - \lambda_i)^j}.$$

6.10 Das Minimalpolynom

Sei K ein Körper, V ein K-Vektorraum mit $\dim V = n$ und $\alpha \in \mathrm{End}_K(V)$ ein Endomorphismus.

Lemma 6.68. *Es gibt ein eindeutig bestimmtes normiertes Polynom $0 \neq \chi_\alpha^{\min} \in K[t]$, so dass $\chi_\alpha^{\min} | g$ für jedes $g \in K[t]$ mit $g(\alpha) = 0$.*

Beweis.
Existenz: Nach dem Satz von Cayley-Hamilton gilt $\chi_\alpha(\alpha) = 0$, d. h. die Menge

$$A = \{f \in K[t] \quad | \quad f(\alpha) = 0\}$$

besteht nicht nur aus dem Nullpolynom. Sei $0 \neq g \in A$ ein normiertes Polynom mit minimalen Grad. Ist $f \in A$ ein anderes Polynom, so gilt $f = qg + r$ für $q, r \in K[t]$ mit $\deg r < \deg g$ und

$$r(\alpha) = f(\alpha) - q(\alpha)g(\alpha) = 0,$$

als $r \in A$. Wegen der Minimalität von $\deg g$ folgt $r = 0$, d. h. $g|f$.
Eindeutigkeit: Seien $g_1, g_2 \in A$ zwei normierte Polynome mit $g_i|f$ für jedes $f \in A$. Dann gilt $g_1|g_2$ und umgekehrt, d. h. es gibt $q_1, q_2 \in K[t]$ mit $g_2 = q_1 g_1$, $g_1 = q_2 g_2$. Insbesondere gilt $\deg g_2 = \deg g_1$ und $\deg q_1 = \deg q_2 = 0$, also $q_1, q_2 \in K^\times$. Da g_1, g_2 normiert sind, folgt $q_1 = q_2 = 1$ und somit $g_1 = g_2$. □

Definition 6.69. Das Polynom $\chi_\alpha^{\min}$ heißt das *Minimalpolynom* von α.

Korollar 6.70. *Es gilt $\chi_\alpha^{\min}|\chi_\alpha$, insbesondere auch $\deg \chi_\alpha^{\min} \leq \dim V$.*

Beweis. Wegen $\chi_\alpha(\alpha) = 0$ gilt $\chi_\alpha^{\min}|\chi_\alpha$, also $\deg \chi_\alpha^{\min} \leq \deg \chi_\alpha = \dim V$. □

Satz 6.71. *Die Nullstellen von $\chi_\alpha^{\min}$ sind genau die Eigenwerte von α, d. h. die Polynome χ_α und $\chi_\alpha^{\min}$ haben dieselben Nullstellen (aber in der Regel unterschiedlichen Vielfachheiten).*

Beweis. Wegen $\chi_\alpha^{\min}|\chi_\alpha$ sind alle Nullstellen von $\chi_\alpha^{\min}$ auch Nullstellen von χ_α und damit Eigenwerte von α. Sei umgekehrt $\lambda \in K$ ein Eigenwert von α und $0 \neq v \in V$ ein Eigenvektor zu λ, d. h. $\alpha(v) = \lambda v$. Dann gilt

$$0 = \chi_\alpha^{\min}(\alpha)(v) = \chi_\alpha^{\min}(\lambda)v.$$

Wegen $v \neq 0$ folgt $\chi_\alpha^{\min}(\lambda) = 0$. □

Satz 6.72. *Der Endomorphismus α ist genau dann diagonalisierbar, wenn $\chi_\alpha^{\min}$ in paarweise verschiedene Linearfaktoren zerfällt.*

Beweis. Sei α diagonalisierbar und $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von α. Sei $f(t) = (t - \lambda_1) \cdots (t - \lambda_r)$ und v ein Eigenvektor zu Eigenwert λ_i. Dann gilt

$$f(\alpha)(v) = f(\lambda_i)(v) = 0.$$

Da α diagonalisierbar ist, hat V eine Basis $(v_1, \ldots, v_n)$ aus Eigenvektoren und $f(\alpha)(v_i) = 0$ für alle i. Also ist $f(\alpha) = 0$ der Null-Endomorphismus und $\chi_\alpha^{\min}|f$. Andererseits gilt $\chi_\alpha^{\min}(\lambda_i) = 0$ für alle i, also $f|\chi_\alpha^{\min}$. Da beide Polynome normiert sind, folgt $f = \chi_\alpha^{\min}$.

Es gelte $\chi_\alpha^{\min}(t) = (t-\lambda_1)\cdots(t-\lambda_r)$ mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_r$. Mittels Partialbruchzerlegung erhalten wir Konstanten $a_1, \ldots, a_r \in K$, so dass

$$\frac{1}{\chi_\alpha^{\min}(t)} = \sum_{i=1}^{r} \frac{a_i}{t-\lambda_i}.$$

Setze $f_i = \frac{a_i \chi_\alpha^{\min}(t)}{t-\lambda_i} \in K[t]$. Dann gilt

$$1 = \sum_{i=1}^{r} f_i, \qquad (t-\lambda_i) f_i = a_i \chi_\alpha^{\min},$$

also auch

$$\mathrm{id}_V = \sum_{i=1}^{r} f_i(\alpha), \qquad (\alpha - \lambda_i \,\mathrm{id}_V) \circ f_i(\alpha) = 0.$$

Insbesondere gilt für jedes $v \in V$

$$v = \sum_{i=1}^{r} f_i(\alpha)(v), \qquad f_i(\alpha)(v) \in V_{\lambda_i} = \ker(\alpha - \lambda_i \,\mathrm{id}_V).$$

Die lineare Abbildung

$$\phi\colon \bigoplus_{i=1}^{r} V_{\lambda_i} \to V, \qquad (v_1, \ldots, v_r) \mapsto v_1 + \cdots + v_r$$

ist also surjektiv. Wegen

$$\dim \bigoplus_{i=1}^{r} V_{\lambda_i} = \sum_{i=1}^{r} \mu_{\mathrm{geo}}(\lambda_i) \le \sum_{i=1}^{r} \mu_{\mathrm{alg}}(\lambda_i) \le \deg \chi_\alpha = \dim V$$

ist sie auch injektiv. Insbesondere ist α diagonalisierbar. □

Anwendungsbeispiele:

Beispiel. Sei $K = \mathbb{C}$ und $\alpha \in \mathrm{Gl}_K(V)$ ein Automorphismus endlicher Ordnung, d. h. es existiert ein $k \in \mathbb{N}$, so dass $\alpha^k = \mathrm{id}_V$. Dann ist α Nullstelle von $t^k - 1 \in \mathbb{C}[t]$, also $\chi_\alpha^{\min} \mid t^k - 1$. Nun gilt

$$t^k - 1 = \prod_{s=0}^{k-1} (t - \zeta_k^s), \qquad \zeta_k = e^{\frac{2\pi i}{k}} = \cos(\frac{2\pi}{k}) + i \sin(\frac{2\pi}{k})$$

d. h. $t^k - 1$ und somit auch $\chi_\alpha^{\min}$ zerfällt in paarweise verschiedene Linearfaktoren. Also ist α diagonalisierbar.
Sei $K = \mathbb{Z}/p\mathbb{Z}$ für eine Primzahl p und $\alpha \in \mathrm{Gl}_K(K^2)$ bezüglich der kanonischen Basis durch die Matrix $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ gegeben. Dann gilt $\alpha^p = \mathrm{id}$, aber α ist nicht diagonalisierbar: Es gilt $\chi_\alpha^{\min} = \chi_\alpha = (t-1)^2$.

Definition 6.73. Ein Endomorphismus $\alpha \in \mathrm{End}_K(V)$ heißt

(i) *nilpotent*, wenn $\alpha^m = 0$ für ein $m \in \mathbb{N}$.

(ii) *unipotent*, wenn $\alpha - \mathrm{id}_V$ nilpotent ist

(iii) *idempotent*, wenn $\alpha^2 = \alpha$.

Satz 6.74. *Sei $\alpha \in \mathrm{End}_K(V)$.*

(i) Ist α nilpotent, so ist α genau dann diagonalisierbar, wenn $\alpha = 0$.

(ii) Ist α unipotent, so ist α genau dann diagonalisierbar, wenn $\alpha = \mathrm{id}_V$.

(iii) Ist α idempotent, so hat α bezüglich einer Basis die Darstellungsmatrix

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

für geeignetes $s \leq n$.

Beweis.
(i): Ist $\alpha^m = 0$, so ist α Nullstelle von $t^m \in K[t]$. Also gilt $\chi_\alpha^{\min} = t^k$ für ein $k \leq \min(n, m)$ und $\chi_\alpha^{\min}$ zerfällt genau dann in paarweise verschiedene Linearfaktoren, wenn $k = 1$ und $\alpha = 0$.
(ii): Dasselbe Argument mit $(t-1)^m$.
(iii): Ist α idempotent, so gilt $\chi_\alpha^{\min} \mid t^2 - t = t(t-1)$. Insbesondere zerfällt $\chi_\alpha^{\min}$ in paarweise verschiedene Linearfaktoren und α hat nur 1 oder 0 als mögliche Eigenwerte. □

Wie weit sind wir mit unserer Klassifikation der Matrizen bis auf Ähnlichkeit?

Satz 6.75. *Jede Matrix $A \in M_{3,3}(\mathbb{C})$ ist eindeutig bis auf Ähnlichkeit durch ihr charakteristisches Polynom χ_A und ihr Minimalpolynom $\chi_A^{\min}$ bestimmt.*

Beweis. Schreibe

$$\chi_A = (t - \lambda_1)(t - \lambda_2)(t - \lambda_3)$$

mit $\lambda_i \in \mathbb{C}$. Sind die λ_i paarweise verschieden, so ist jede Matrix B mit $\chi_B = \chi_A$ ähnlich zu der Diagonalmatrix $\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$, also auch zu A. Gilt $\lambda_3 = \lambda_2$, so müssen wir zusätzlich annehmen, dass auch $\chi_B^{\min} = \chi_A^{\min}$. Falls $\lambda_3 \neq \lambda_1$, so gilt

(a_1) $\chi_A^{\min} = (t - \lambda_1)(t - \lambda_2)$ oder

(a_2) $\chi_A^{\min} = (t - \lambda_1)(t - \lambda_2)^2$.

Falls $\lambda_3 = \lambda_2 = \lambda_1$, so gilt

(b_i) $\chi_A^{\min} = (t - \lambda_1)^i$

mit $1 \leq i \leq 3$.
Im Fall (a_1) und (b_1) sind wieder sowohl A als auch B ähnlich zu der Diagonalmatrix $\operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$.
Im Fall (a_2) sei $V'_{\lambda_2} = \ker(A - \lambda_2 E_3)^2$. Da $\chi_A^{\min} \neq (t - \lambda_2)^2$ ist $A - \lambda_2 E \neq 0$ und $\dim V'_{\lambda_2} \leq 2$. Mittels Partialbruchzerlegung finden wir $r, s \in \mathbb{C}^\times$ mit

$$\frac{1}{\chi_A^{\min}} = \frac{r}{t - \lambda_1} + \frac{s}{(t - \lambda_2)^2}$$

Setze

$$f_1 = \frac{r\chi_A^{\min}}{t - \lambda_1}, \qquad f_2 = \frac{s\chi_A^{\min}}{(t - \lambda_2)^2} \in \mathbb{C}[t].$$

Dann gilt für jedes $v \in \mathbb{C}^3$

$$v = f_1(A)v + f_2(A)v, \qquad f_1(A)v \in V_{\lambda_1}, \qquad f_2(A)v \in V'_{\lambda_2}$$

Wegen

$$\dim V_{\lambda_1} + \dim V'_{\lambda_2} \leq 3$$

folgern wir, dass V'_{λ_2} ein Komplement zu V_{λ_1} in $\mathbb{C}^3$ sein muss. Insbesondere gilt $\dim V'_{\lambda_2} = 2$. Sei $v_1 \in V_{\lambda_1}$ ein Eigenvektor. Da A nicht diagonalisierbar ist, gibt es ein $0 \neq v_3 \in V'_{\lambda_2}$ mit $v_3 \notin V_{\lambda_2}$. Setze $v_2 = (A - \lambda E_3)v_3$. Dann gilt $0 \neq v_2 \in V_{\lambda_2}$ und (v_1, v_2, v_3) ist eine Basis von $\mathbb{C}^3$ mit $Av_1 = \lambda_1 v_1$, $Av_2 = \lambda_2 v_2$, $Av_3 = \lambda_2 v_2 + v_3$, d. h. A ist ähnlich zu der Matrix

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$

Dasselbe Argument zeigt, dass auch B ähnlich zu dieser Matrix ist, also auch zu A.
Im Fall (b_2) gilt $\operatorname{im}(A - \lambda E_3) \subset \ker(A - \lambda E_3) = V_{\lambda_1} \neq \mathbb{C}^3$, also

$$3 - \dim V_{\lambda_1} = \dim \operatorname{im}(A - \lambda E_3) \leq \dim V_{\lambda_1} < 3.$$

Dies geht nur, wenn $\dim V_{\lambda_1} = 2$. Wähle $v_3 \in \mathbb{C}^3 \setminus V_{\lambda_1}$. Dann ist $0 \neq v_2 = (A - \lambda E_3)v_3$ ein Eigenvektor. Ergänze zu einer Basis (v_1, v_2) von V_{λ_1}. Dann ist (v_1, v_2, v_3) eine Basis von $\mathbb{C}^3$ und A und B ähnlich zu der Matrix

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}.$$

Im Fall (b_3) gilt $V'_{\lambda_1} = \ker(A - \lambda_1 E_3)^2 \neq \mathbb{C}^3$. Wähle $v_3 \in \mathbb{C}^3 \setminus V'_{\lambda_1}$ und setze $v_2 = (A - \lambda_1 E_3)v_1$, $v_3 = (A - \lambda_1 E_3)v_1$. Dann ist das System (v_1, v_2, v_3) linear unabhänging: Multiplikation der Gleichung

$$0 = rv_1 + sv_2 + tv_3$$

mit $(A - \lambda_1 E_3)^2$,$(A - \lambda_1 E_3)$ und 1 zeigt nacheinander $r = s = t = 0$. Damit ist das System auch eine Basis und A, B ähnlich zu der Matrix

$$\begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}.$$

□

Das im Beweis benutzte Argument ließe sich noch weiter ausbauen. So könnte man zeigen, dass eine Matrix $A \in M_{n,n}(K)$ für beliebiges n und beliebigen Körper K eindeutig bis auf Ähnlichkeit durch ihr charakteristisches Polynom χ_A und Minimalpolynom $\chi_A^{\min}$ bestimmt ist, sobald jeder Primfaktor von χ_A entweder in der Primfaktorzerlegung von $\chi_A^{\min}$ oder von $\frac{\chi_A}{\chi_A^{\min}}$ nur einfach vorkommt.

Das Minimalpolynom und das charakteristische Polynom reichen jedoch nicht aus, um zu zeigen, dass die beiden Matrizen

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in M_{4,4}(\mathbb{C})$$

nicht zueinander ähnlich sind. Hierfür brauchen wir noch feinere Invarianten, die wir im nächsten Semester kennnenlernen werden.