# Lineare Algebra II, SS 2021

Ulrich Görtz

Version vom 13. Mai 2021.

Ulrich Görtz

Universität Duisburg-Essen
Fakultät für Mathematik
45117 Essen

ulrich.goertz@uni-due.de

Ich freue mich über Kommentare und Berichtigungen.

Ich bedanke mich für Bemerkungen/Korrekturen bei Fereshteh Fattahi, Lukas Fußangel, Jesco Nevihosteny.



Gesetzt in der Schrift Vollkorn[3] von F. Althausen mit LuaLaTeX, TikZ und anderen TEX-Paketen. Einige Abbildungen wurden mit IPE[4] erstellt. Die HTML-Version wird mit plasTeX[5] erzeugt.

[1] https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode.de
[2] https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de
[3] http://vollkorn-typeface.com/
[4] http://ipe.otfried.org/
[5] https://github.com/plastex/plastex

# Inhaltsverzeichnis

KAPITEL 14

# Einleitung

Diese Vorlesung ist die Fortsetzung der Linearen Algebra 1, und entsprechend baut das Skript auf dem Skript zur Linearen Algebra 1 auf.

Die Vorlesung Linearen Algebra 2 lässt sich grob in drei Themenbereiche unterteilen,

- erstens die Fortsetzung des Studiums der Eigenwerttheorie, insbesondere die Frage, wann ein Endomorphismus diagonalisierbar ist und welche »Normalform« der darstellenden Matrix im nicht-diagonalisierbare Fall erreicht werden kann,
- zweitens die Konstruktion des »Quotienten« eines Vektorraums nach einem Unterraum (und analoger Konstruktionen für Gruppen und Ringe) und
- drittens das Studium von Bilinearformen über den reellen Zahlen (und Sesquilinearformen über den komplexen Zahlen).

Im Rest dieser Einleitung sollen diese drei Themen etwas genauer beleuchtet werden.

## 14.1. Die Jordansche Normalform

Seien $K$ ein Körper, $V$ ein endlichdimensionaler $K$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus. Wir haben in der Linearen Algebra 1 definiert, wann $f$ diagonalisierbar heißt, und auch gesehen, dass nicht jeder Endomorphismus diagonalisierbar ist.

Es ist möglich und wichtig, noch bessere Kriterien dafür zu entwickeln, wann ein Endomorphismus diagonalisierbar ist, und für Endomorphismen, die diese Eigenschaft nicht haben, ebenfalls »möglichst einfache« darstellende Matrizen bezüglich geeigneter Basen zu suchen.

Für solche Endomorphismen, die überhaupt eine darstellende Matrix in oberer Dreiecksgestalt besitzen, werden wir im Satz über die Jordansche Normalform zeigen, dass sich eine darstellende Matrix finden lässt, die höchstens auf der Diagonale und auf der direkt über der Diagonale liegenden Nebendiagonale Einträge hat.

Um das zu beweisen, werden wir die ersten Wochen der Vorlesung darauf verwenden, die Theorie des »Polynomrings« über einem Körper zu entwickeln und zeigen, dass es in diesen Ringen ganz ähnlich wie im Ring der ganzen Zahlen eine »Primfaktorzerlegung« gibt. Auch wenn es erst später ab der vierten Vorlesungswoche wirklich sichtbar werden wird, wie die Verbindung zur Linearen Algebra hergestellt wird, stellt sich diese Theorie als essenziell für das weitere heraus.

## 14.2. Quotienten und andere Universalkonstruktionen

Um zu erklären, was es mit der Quotientenkonstruktion auf sich hat, betrachten wir die folgende Situation: Sei $K$ ein Körper, $V$ ein Vektorraum und $U \subseteq V$ ein Untervektorraum. Wenn $f\colon V \to W$ ein Vektorraumhomomorphismus mit Kern $U$ ist, dann werden Vektoren $v, v'$ unter $f$ genau dann auf dasselbe Element von $W$ abgebildet, wenn die Differenz $v - v'$ in $U$ liegt. Vektoren, die sich »nur um ein Element aus $U$ unterscheiden«, werden also unter $f$ »identifiziert«.

Aber gibt es zu gegebenem $U$ überhaupt immer einen Homomorphismus, der $U$ als Kern hat? Wir haben in der Linearen Algebra I gesehen, dass das jedenfalls dann immer der Fall ist, wenn $V$ endlichdimensional ist. Allerdings mussten wir, um ein solches $f$ zu erhalten, einen Komplementärraum zu $U$ wählen. Dass hier eine Wahl erforderlich ist, ist etwas unschön, und an dieser Stelle entsteht auch die Einschränkung auf den endlichdimensionalen Fall, weil wir den Basisergänzungssatz benötigen, den wir nur für endlichdimensionale Vektorräume bewiesen hatten. Die Aussage gilt aber allgemein, und die Konstruktion des Quotienten $V/U$ und der zughörigen »kanonischen Projektion« $V \to V/U$ ist eine abstrakte Konstruktion eines Vektorraumhomomorphismus mit Kern $U$.

Insofern kann man argumentieren, dass man diese Konstruktion schon viel früher in der Vorlesung hätte behandeln können, auch schon vor der Einführung der Begriffe der Basis und der Dimension. Andererseits hat man durch die Wahl eines Komplementärraums (jedenfalls im endlichdimensionalen Fall) einen guten »Ersatz« für den Quotienten, und das ist der Grund, warum es auch nicht schadet, die allgemeine Konstruktion erst etwas später zu machen.

Eine sehr ähnliche Konstruktion ist die des Restklassenringes $\mathbb{Z}/n$ zusammen mit der kanonischen Projektion $\mathbb{Z} \to \mathbb{Z}/n$, die wir in der Linearen Algebra I kennengelernt haben (Abschnitt I.4.2.1). Diese Konstruktion werden wir mit dem Begriff des Quotienten eines Rings nach einem Ideal weiter verallgemeinern.

Ist $G$ eine Gruppe und $H \subseteq G$ eine Untergruppe, dann kann man sich ebenso die Frage stellen, ob es einen Gruppenhomomorphismus $f\colon G \to G'$ mit $\operatorname{Ker}(f) = H$ gibt. Dies ist allerdings nicht immer der Fall! Wenn wir über Quotienten von Gruppen sprechen, werden wir klären, welche zusätzliche Bedingung $H$ erfüllen muss.

Wir werden auch besprechen, was es bedeutet, dass der Quotient (beispielsweise eines Vektorraums nach einem Untervektorraum) durch eine »universelle Eigenschaft« charakterisiert werden kann. Mit ähnlichen universellen Eigenschaften lassen sich viele Konstruktionen charakterisieren, die wir schon gesehen haben (zum Beispiel auch das Produkt und die direkte Summe von Vektorräumen, der Kern und das Bild von linearen Abbildungen, ...), und dieser Begriff ist oft nützlich, wenn man in anderen Kontexten das richtige »Analogon« zu einem dieser Begriffe sucht.

## 14.3. Euklidische und unitäre Vektorräume

Ein »euklidischer Vektorraum« ist ein endlichdimensionaler Vektorraum über den reellen Zahlen, in dem wir eine zusätzliche Struktur zur Verfügung haben, die uns erlaubt, Abstände zwischen Punkten und die Länge von Vektoren zu messen, darüber zu sprechen, wann zwei Vektoren zueinander senkrecht sind, und den Winkel zwischen zwei Vektoren zu definieren. In Kapitel I.11 wird das für den Standardvektorraum $\mathbb{R}^n$ erklärt, aber in der Linearen Algebra 2 wollen wir eine entsprechende Theorie für beliebige (endlichdimensionale) $\mathbb{R}$-Vektorräume definieren.

Sei $V$ ein endlichdimensionaler Vektorraum über $\mathbb{R}$. Wie sich herausstellen wird, kann man alle die oben genannten geometrischen Begriffe (Abstand, Länge, Winkel) definieren, sobald ein sogenanntes *Skalarprodukt*

$$\beta\colon V \times V \to \mathbb{R}$$

gegeben ist, dass ist eine bilineare Abbildung (d.h. $\beta$ ist linear in jedem der beiden Faktoren, also eine multilineare Abbildung $V^2 \to \mathbb{R}$), für die außerdem $\beta(v,w) = \beta(w,v)$ für alle $v, w \in V$ und $\beta(v,v) > 0$ für alle $v \in V \setminus \{0\}$ gilt. Zum Beispiel kann man dann die Länge eines Vektors $v$ durch

$$\|v\| := \sqrt{\beta(v,v)}$$

definieren.

Für $V = \mathbb{R}^n$ ist durch $\beta((x_i)_i, (y_i)_i) := \sum_{i=1}^n x_i y_i$ ein solches Skalarprodukt gegeben, das sogenannte Standardskalarprodukt.

Es zeigt sich, dass mit einem kleinen Trick auch für Vektorräume über den komplexen Zahlen eine ganz ähnliche Theorie entwickelt werden kann, und es ist zum Beispiel für Anwendungen in der theoretischen Physik sehr nützlich, das zu tun. Würde man auf $\mathbb{C}^n$ das Standardskalarprodukt durch dieselbe Formel wie für $\mathbb{R}^n$ definieren, dann würde natürlich im allgemeinen nicht gelten, dass das Skalarprodukt eines Vektors $\neq 0$ mit sich selbst eine positive reelle Zahl ist. Wenn man die Formel stattdessen abändert zu

$$\beta((x_i)_i, (y_i)_i) := \sum_{i=1}^n \overline{x_i} y_i,$$

dann gilt aber $\beta(x, x) \in \mathbb{R}_{>0}$ für alle $x \in \mathbb{C}^n \setminus \{0\}$, so dass man dann wieder die Länge von $x$ durch $\|x\| := \sqrt{\beta(x, x)}$ definieren kann. Hier bezeichnet für eine komplexe Zahl $x = a + ib$, $a, b \in \mathbb{R}$, das Symbol $\overline{x} := a - ib$ die sogenannte komplex konjugierte Zahl. Dann gilt $x\overline{x} = a^2 + b^2 \geq 0$ und der Ausdruck ist nur für $x = 0$ gleich Null.

Um diese Idee umzusetzen, betrachtet man statt bilinearer Abbildungen im Fall eines komplexen Vektorraums $V$ sogenannte *Sesquilinearformen*, das sind Abbildungen

$$\beta \colon V \to V \to \mathbb{C},$$

die im zweiten Eintrag linear, aber im ersten Eintrag »semilinear bezüglich der komplexen Konjugation« sind, d.h. es gilt $\beta(xv + x'v', w) = \overline{x}\beta(v, w) + \overline{x'}\beta(v', w)$ für alle $x, x' \in \mathbb{C}$, $v, v', w \in V$. Die Symmetriebedingung ersetzt man entsprechend durch die Bedingung $\beta(w, v) = \overline{\beta(v, w)}$.

Dann man ganz parallel die Theorie der euklidischen Vektorräume ($\mathbb{R}$-Vektorräume mit einem Skalarprodukt) und der unitären Vektorräume ($\mathbb{C}$-Vektorräume mit einem Skalarprodukt im Sinne einer Sesquilinearform) entwickeln.

Man erhält damit eine Theorie, die nicht nur für geometrische Betrachtungen nützlich ist. Zum Beispiel werden wir als eine Konsequenz des Spektralsatzes für selbstadjungierte Abbildungen (Theorem 19.42) beweisen können, dass jede Matrix $A \in M_n(\mathbb{R})$, die *symmetrisch* ist (d.h. $A = A^t$), diagonalisierbar ist.

KAPITEL 15

# Ringe

## 15.1. Definition und erste Eigenschaften

Wir beginnen mit der Definition einer weiteren algebraischen Struktur, der sogenannten *Ringe*, in denen eine Addition und Multiplikation existiert, wo wir aber anders als bei Körpern nicht verlangen, dass jedes Element $\neq 0$ ein multiplikatives Inverses hat. Die Definition hat verschiedene »Versionen«, je nachdem, ob gefordert wird, dass die Multiplikation ein neutrales Element hat (das werden wir immer verlangen) und/oder kommutativ ist. Zwei wichtige Beispiele von Ringen sind der Ring $\mathbb{Z}$ der ganzen Zahlen und der Ring $M_n(K)$ der quadratischen Matrizen der Größe $n \in \mathbb{N}$ über einem Körper $K$.

DEFINITION 15.1. (1) Ein *Ring* ist eine Menge $R$ zusammen mit Verknüpfungen

$$+\colon R \times R \to R \,(\text{Addition}) \text{ und } \quad \cdot\colon R \times R \to R \,(\text{Multiplikation}),$$

so dass gilt:

(a) $(R, +)$ ist eine kommutative Gruppe,

(b) die Multiplikation $\cdot$ ist assoziativ.

(c) es gelten die Distributivgesetze

$$a(b+c) = a \cdot b + a \cdot c, \quad (a+b)c = a \cdot c + b \cdot c$$

für alle $a, b, c \in R$.

(2) Wenn die Multiplikation von $R$ kommutativ ist, dann nennt man $R$ auch einen *kommutativen Ring*.

(3) Wenn die Multiplikation von $R$ ein neutrales Element besitzt, so wird dieses mit 1 bezeichnet, und man nennt $R$ einen *Ring mit Eins*.

⊣

Wir nutzen dieselben Konventionen wie im Fall von Körpern: Der Multiplikationspunkt kann ausgelassen werden, wenn keine Missverständnisse dadurch entstehen können. Es gilt »Punkt- vor Strichrechnung«. Für die additive Gruppe $(R, +)$ verwenden wir die üblichen Bezeichnungen: Das neutrale Element der Addition in einem Ring bezeichnen wir mit 0, das additive Inverse von $a$ mit $-a$, und wir schreiben $a - b$ statt $a + (-b)$.

In diesem Skript verstehen wir, wenn nicht ausdrücklich etwas anderes gesagt wird, unter einem *Ring* immer einen *Ring mit Eins*. Dann ist das neutrale Element der Multiplikation eindeutig bestimmt, so dass die in der Definition festgelegte Bezeichnung 1 sinnvoll ist. In der Vorlesung treten sowohl kommutative als auch nicht-kommutative Ringe auf.

Für $a \in R$ und $n \in \mathbb{N}$ ist $a^n = a \cdot \dots \cdot a$ das $n$-fache Produkt von $a$ mit sich selbst. Für $n = 0$ verstehen wir das wie üblich als das leere Produkt, d.h. wir setzen $a^0 = 1$.

DEFINITION 15.2. Sei $R$ ein Ring. Ein Element $a \in R$ heißt eine *Einheit*, wenn $a$ ein multiplikatives Inverses besitzt, d.h., wenn $b \in R$ existiert mit $ab = ba = 1$. Die Menge aller Einheiten von $R$ bildet bezüglich der Multiplikation eine Gruppe, die wir die *Einheitengruppe* oder *multiplikative Gruppe von* $R$ nennen und mit $R^\times$ bezeichnen. ⊣

Ist $R$ ein Ring und $b \in R$ eine Einheit, so ist das multiplikative Inverse von $b$ eindeutig bestimmt und wird auch mit $b^{-1}$ bezeichnet. Im Fall kommutativer Ringe, wo also $ab^{-1} = b^{-1}a$ für alle $a \in R$ gilt, verwendet man auch gelegentlich die Bruchschreibweise $\frac{a}{b}$ für das Element $ab^{-1}$. Ist der Ring nicht kommutativ, so sollte man diese Schreibweise vermeiden, weil unklar bleibt, ob $ab^{-1}$ oder $b^{-1}a$ gemeint ist.

BEISPIEL 15.3. (1) Die ganzen Zahlen bilden bezüglich der üblichen Addition und Multiplikation einen kommutativen Ring. Es ist $\mathbb{Z}^\times = \{1, -1\}$.

(2) Ist $n \in \mathbb{N}_{>1}$, so ist $\mathbb{Z}/n$ mit der Addition und Multiplikation von Restklassen modulo $n$ ein kommutativer Ring. Das haben wir (ohne das Wort »Ring« zu verwenden) in Abschnitt I.4.2.1 nachgeprüft. Die Einheitengruppe $(\mathbb{Z}/n)^\times$ besteht aus den Restklassen aller derjenigen Zahlen $m \in \mathbb{Z}$, die zu $n$ teilerfremd sind, siehe Satz I.4.16.

(3) Jeder Körper ist ein kommutativer Ring. Ein Ring ist genau dann ein Körper, wenn er kommutativ ist und $R^\times = R \setminus \{0\}$ gilt. Insbesondere stimmt für einen Körper $K$ die neu eingeführte Schreibweise $K^\times$ mit der im vergangenen Semester eingeführten überein.

(4) Sei $K$ ein Körper, $n \in \mathbb{N}$. Dann ist $M_n(K)$ mit der Addition von Matrizen umd dem Matrizenprodukt ein Ring, der sogenannte *Matrizenring*. Ist $n \geq 2$, dann ist der Ring $M_n(K)$ nicht kommutativ.

(5) Ist $K$ ein Körper und $V$ ein $K$-Vektorraum, so ist $\mathrm{End}_K(V)$ mit der Addition von linearen Abbildungen und der Verkettung von linearen Abbildungen als Multiplikation ein Ring, der sogenannte *Endomorphismenring* von $V$. In diesem Ring entspricht die Potenz eines Elements $f$ also der entsprechend häufigen Verkettung des Endomorphismus $f$ mit sich selbst, zum Beispiel: $f^3 = f \circ f \circ f$.

(6) Die einelementige Menge $R = \{0\}$ ist (mit der einzig möglichen Addition $0 + 0 = 0$ und Multiplikation $0 \cdot 0 = 0$) ein Ring, der sogenannte *Nullring*. Dies ist der einzige Ring, in dem $1 = 0$ gilt, denn in jedem Ring gilt $1 \cdot a = a$ für alle $a$ nach Definition des Elements 1 und $0 \cdot a = 0$.

(7) Sind $R_1, R_2$ Ringe, so ist $R_1 \times R_2$ mit der komponentenweisen Addition und Multiplikation ein Ring, das sogenannte *Produkt* von $R_1$ und $R_2$. Das bedeutet

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2), \quad (x_1, x_2) \cdot (y_1, y_2) = (x_1 y_1, x_2 y_2).$$

Das Nullelement ist $(0, 0)$, das Einselement ist $(1, 1)$. Ist allgemeiner $I$ irgendeine Menge und sind $R_i, i \in I$, Ringe, so ist das Produkt $\prod_{i \in I} R_i$ mit der komponentenweisen Addition und Multiplikation ein Ring, das Produkt der Ringe $R_i$.

(8) Ist $R$ ein Ring und $X$ eine Menge, so bildet die Menge $\mathrm{Abb}(X, R)$ aller Abbildungen von $X$ nach $R$ einen Ring mit

$$(f + g)(x) = f(x) + g(x),$$
$$(f \cdot g)(x) = f(x) \cdot g(x).$$

Das Null- und Einselement sind die konstanten Abbildungen $x \mapsto 0$ und $x \mapsto 1$. Wenn $R$ kommutativ ist, dann ist auch dieser Ring kommutativ.

Wir können $\mathrm{Abb}(X, R)$ identifizieren mit dem Produkt $R^X = \prod_{x \in X} R$. Dabei entspricht eine Abbildung $f\colon X \to R$ dem Element $(f(x))_x \in R^X$.

◊

In jedem Ring $R$ gilt $0 \cdot a = 0 = a \cdot 0$ und $(-1)a = -a = a \cdot (-1)$ für alle $a \in R$. Das folgt aus dem Distributivgesetz. Aus $ab = ac$ folgt allerdings im allgemeinen nicht, dass $b = c$ ist; ebenso impliziert $ab = 0$ nicht unbedingt, dass $a = 0$ oder $b = 0$ gilt. (Geben Sie für beide Aussagen Beispiele im Matrizenring $M_n(K)$.) Vergleiche aber Definition 15.28, Lemma 15.32.

DEFINITION 15.4. Seien $R, S$ Ringe. Ein *Ringhomomorphismus* von $R$ nach $S$ ist eine Abbildung $f\colon R \to S$, so dass gilt:

(a) für alle $x, y \in R$ ist $f(x+y) = f(x) + f(y)$,

(b) für alle $x, y \in R$ ist $f(xy) = f(x)f(y)$,

(c) es gilt $f(1) = 1$.

⊣

BEMERKUNG 15.5. Ist $f\colon R \to S$ ein Ringhomomorphismus, so gilt $f(0) = 0$ und $f(-x) = -f(x)$ für alle $x \in R$. Ferner induziert $f$ einen Gruppenhomomorphismus $R^\times \to S^\times$ zwischen den Einheitengruppen, denn aus $ab = 1$ folgt $f(a)f(b) = f(ab) = f(1) = 1$, also $f(a) \in S^\times$. ◊

Wie man leicht nachprüft, ist die Verkettung von Ringhomomorphismen wieder ein Ringhomomorphismus. Für jeden Ring $R$ ist die identische Abbildung $\mathrm{id}_R$ ein Ringhomomorphismus.

BEISPIEL 15.6. Sei $R$ ein Ring. Dann gibt es einen *eindeutig bestimmten* Ringhomomorphismus $\varphi\colon \mathbb{Z} \to R$. Denn nach Definition eines Ringhomomorphismus muss $\varphi(1) = 1$ gelten, wobei links die ganze Zahl 1 und rechts das Element $1 \in R$ gemeint sind. Es folgt für alle $n \in \mathbb{N}_{\geq 1}$, dass

$$\varphi(n) = 1 + \cdots + 1,$$

wobei in der Summe rechts das Element $1 \in R$ zu sich selbst addiert wird, und die Summe aus $n$ Summanden besteht. Schließlich hat man $\varphi(-n) = -\varphi(n)$, so dass $\varphi$ auch auf den negativen ganzen Zahlen eindeutig festgelegt ist. Es ist nicht schwer zu überprüfen, dass es sich bei dieser Abbildung tatsächlich um einen Ringhomomorphismus handelt.

Wir haben diese Abbildung in dem speziellen Fall, dass $R$ ein Körper ist, schon im Abschnitt I.4.2.2 betrachtet; siehe auch Ergänzung 18.29.

Wie bei Körpern bezeichnen wir das Bild der ganzen Zahl $n$ unter diesem Ringhomomorphismus oft auch einfach wieder mit $n$. In diesem Sinne können wir $n$ als Element jedes Rings $R$ auffassen. Allerdings kann, wie schon bei Körpern, dann $m = n$ in $R$ gelten, auch wenn die ganzen Zahlen $m$ und $n$ unterschiedlich sind. ◊

BEISPIEL 15.7. Wir können nun Lemma I.4.13 eleganter formulieren: Die natürliche Abbildung $\mathbb{Z} \to \mathbb{Z}/n$ ist ein Ringhomomorphismus (und dies ist der Ringhomomorphismus aus Beispiel 15.6 für den Ring $\mathbb{Z}/n$). ◊

BEISPIEL 15.8. Sei $K$ ein Körper.

(1) Sei $V$ ein $K$-Vektorraum. Sei $\mathrm{End}_{\mathrm{Gp}}(V)$ die Menge aller Gruppenendomorphismen $V \to V$ der additiven Gruppe $(V, +)$. Mit der üblichen Summe von Abbildungen als Addition und der Verkettung von Abbildung als Multiplikation ist $\mathrm{End}_{\mathrm{Gp}}(V)$ ein (im allgemeinen nicht-kommutativer) Ring. Das Einselement ist die Abbildung $\mathrm{id}_V$.

Für $a \in K$ ist die Skalarmultiplikation mit $a$ ein Gruppenendomorphismus $V \to V$, also ein Element von $\mathrm{End}_{\mathrm{Gp}}(V)$. Hier benutzen wir eines der Distributivgesetze für die Skalarmultiplikation auf $V$.

Wir erhalten so einen Ringhomomorphismus $K \to \mathrm{End}_{\mathrm{Gp}}(V)$. Die Kompatibilität mit der Addition entspricht »dem anderen« Distributivgesetz, die Kompatibilität mit der Multiplikation $V$ dem »Assoziativgesetz«. Dass Skalarmultiplikation mit $1 \in K$ die identische Abbildung ist, ist ein weiteres der Vektorraumaxiome.

(2) Sei nun $V$ eine kommutative Gruppe, die wir additiv schreiben, und sei $\varphi\colon K \to \mathrm{End}_{\mathrm{Gp}}(V)$ ein Ringhomomorphismus. Dann erhalten wir durch $a \cdot v := \varphi(a)(v)$ eine Skalarmultiplikation und damit die Struktur eines $K$-Vektorraums auf $V$.

◊

Mit dem Begriff des Homomorphismus erhalten wir wie üblich auch einen Begriff von Isomorphismen zwischen Ringen:

DEFINITION 15.9. Ein *Ringisomorphismus* ist ein Ringhomomorphismus $f\colon R \to S$, derart dass ein Ringhomomorphismus $g\colon S \to R$ existiert, der eine Umkehrabbildung zu $f$ ist, d.h. so dass $g \circ f = \mathrm{id}_R$ und $f \circ g = \mathrm{id}_S$ gilt. ⊣

Wie bei Gruppen und Vektorräumen beweist man:

LEMMA 15.10. *Sei $f\colon R \to S$ ein Ringhomomorphismus. Die Abbildung $f$ ist genau dann bijektiv, wenn $f$ ein Isomorphismus ist.*

DEFINITION 15.11. Sei $S$ ein Ring. Eine Teilmenge $R \subseteq S$ heißt *Unterring*, wenn $R$ eine Untergruppe der additiven Gruppe von $S$ ist, für alle $x, y \in R$ auch das Produkt $xy$ in $R$ liegt, und das Einselement von $S$ in $R$ liegt. ⊣

BEISPIEL 15.12. Seien $K$ ein Körper, $V$ ein endlichdimensionaler $K$-Vektorraum und $\mathscr{B}$ eine Basis von $V$. Dann ist die Abbildung $\mathrm{End}_K(V) \to M_n(K), f \mapsto M_{\mathscr{B}}^{\mathscr{B}}(f)$, ein Ringisomorphismus. ◊

ERGÄNZUNG 15.13. Seien wie in Beispiel 15.12 $K$ ein Körper und $V$ ein endlichdimensionaler $K$-Vektorraum. Man kann zeigen, dass jeder Isomorphismus $\mathrm{End}_K(V) \to M_n(K)$ die Form $f \mapsto M_{\mathscr{B}}^{\mathscr{B}}(f)$ für eine Basis $\mathscr{B}$ von $V$ hat. Das ist ein Spezialfall des Satzes von Skolem und Noether[I]. □ Ergänzung 15.13

Ist $R \subseteq S$ ein Unterring, so ist $R$ mit der Addition und Multiplikation von $S$ selbst ein Ring und die Inklusionsabbildung $R \to S, x \mapsto x$, ist ein injektiver Ringhomomorphismus. Ist andererseits $\iota\colon R \to S$ ein injektiver Ringhomomorphismus, so ist $\iota(R)$ ein Unterring von $S$ und die Abbildung $R \to \iota(R)$ ein Ringisomorphismus.

BEISPIEL 15.14. Zwei eng verwandte Beispielklassen von Ringen, die im weiteren Verlauf der Vorlesung eine große Rolle spielen werden, sind die folgenden.

(1) Seien $K$ ein Körper und $A \in M_n(K)$. Dann ist

$$K[A] := \left\{ \sum_{i=0}^{n} a_i A^i;\ n \in \mathbb{N},\ a_i \in K \right\}$$

ein Unterring von $M_n(K)$. Der Ring $K[A]$ ist kommutativ.

[I] https://de.wikipedia.org/wiki/Satz_von_Skolem-Noether

(2) Seien $K$ ein Körper, $V$ ein $K$-Vektorraum und $f \in \mathrm{End}_K(V)$. Dann ist

$$K[f] := \left\{ \sum_{i=0}^{n} a_i f^i ;\ n \in \mathbb{N},\ a_i \in K \right\}$$

ein Unterring des Endomorphismenrings $\mathrm{End}_K(V)$. Hierbei bezeichnet $f^i$ die $i$-te Potenz von $f$ *im Ring* $\mathrm{End}_K(V)$, d.h. die $i$-fache Verkettung von $f$ mit sich selbst. Der Ring $K[f]$ ist kommutativ.

Ist $V$ endlichdimensional und $\mathscr{B}$ eine Basis von $V$, dann schränkt sich der Isomorphismus $\mathrm{End}_K(V) \xrightarrow{\sim} M_n(K)$ aus Beispiel 15.12 ein zu einem Isomorphismus $K[f] \xrightarrow{\sim} K[A]$.

◊

## 15.2. Ideale

Definition 15.15. Sei $f\colon R \to R'$ ein Ringhomomorphismus. Dann heißen

$$\mathrm{Im} f := f(R)$$

das Bild und

$$\mathrm{Ker} f := f^{-1}(\{0\})$$

der Kern des Ringhomomorphismus $f$. ⊣

Weil ein Ringhomomorphismus $f$ insbesondere ein Homomorphismus der zugehörigen additiven Gruppen ist, folgt aus Lemma I.8.24, dass $f$ genau dann injektiv ist, wenn $\mathrm{Ker}(f) = \{0\}$ gilt.

Es ist leicht zu sehen, dass in dieser Situation $\mathrm{Im} f$ wieder ein Ring ist. Weil meist $1 \notin \mathrm{Ker} f$ gilt, ist der Kern eines Ringhomomorphismus in der Regel kein Ring in unserem Sinne, allerdings stets ein sogenanntes Ideal:

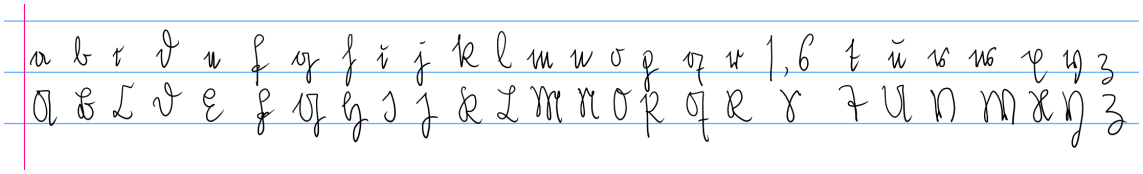
Definition 15.16. Sei $R$ ein Ring. Eine Teilmenge $\mathfrak{a} \subseteq R$ heißt *Ideal* von $R$, falls $\mathfrak{a}$ eine Untergruppe von $(R, +)$ ist und falls für alle $a \in \mathfrak{a}$ und $x \in R$ gilt: $xa \in \mathfrak{a}$ und $ax \in \mathfrak{a}$. ⊣

Bemerkung 15.17. Für die Bezeichnung von Idealen werden häufig Frakturbuchstaben benutzt (vor allem $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ und für Ideale mit speziellen Eigenschaften auch $\mathfrak{m}$, $\mathfrak{n}$, $\mathfrak{p}$, $\mathfrak{q}$). Daher hier eine Liste.

| a | b | c | d | e | f | g | h | i | j | k | l | m |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathfrak{a}$ | $\mathfrak{b}$ | $\mathfrak{c}$ | $\mathfrak{d}$ | $\mathfrak{e}$ | $\mathfrak{f}$ | $\mathfrak{g}$ | $\mathfrak{h}$ | $\mathfrak{i}$ | $\mathfrak{j}$ | $\mathfrak{k}$ | $\mathfrak{l}$ | $\mathfrak{m}$ |
| $\mathfrak{A}$ | $\mathfrak{B}$ | $\mathfrak{C}$ | $\mathfrak{D}$ | $\mathfrak{E}$ | $\mathfrak{F}$ | $\mathfrak{G}$ | $\mathfrak{H}$ | $\mathfrak{I}$ | $\mathfrak{J}$ | $\mathfrak{K}$ | $\mathfrak{L}$ | $\mathfrak{M}$ |

| n | o | p | q | r | s | t | u | v | w | x | y | z |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathfrak{n}$ | $\mathfrak{o}$ | $\mathfrak{p}$ | $\mathfrak{q}$ | $\mathfrak{r}$ | $\mathfrak{s}$ | $\mathfrak{t}$ | $\mathfrak{u}$ | $\mathfrak{v}$ | $\mathfrak{w}$ | $\mathfrak{x}$ | $\mathfrak{y}$ | $\mathfrak{z}$ |
| $\mathfrak{N}$ | $\mathfrak{O}$ | $\mathfrak{P}$ | $\mathfrak{Q}$ | $\mathfrak{R}$ | $\mathfrak{S}$ | $\mathfrak{T}$ | $\mathfrak{U}$ | $\mathfrak{V}$ | $\mathfrak{W}$ | $\mathfrak{X}$ | $\mathfrak{Y}$ | $\mathfrak{Z}$ |

Und noch einmal handgeschrieben (in einer Annäherung der Sütterlin-Schreibschrift[2]; für das kleine »s« gibt es zwei Formen, je nachdem, wo im Wort es steht):



◊

[2] https://de.wikipedia.org/wiki/S%C3%BCtterlinschrift

Beispiel 15.18. (1) In jedem Ring sind $\{0\}$ (das *Nullideal*) und $R$ (das sogenannte *Einsideal*) Ideale.

(2) Ist $\mathfrak{a}$ ein Ideal eines Rings $R$, das eine Einheit von $R$ enthält, so gilt $1 \in R$ und folglich $\mathfrak{a} = R$.

(3) Ist $K$ ein Körper, so sind $\{0\}$ und $K$ die einzigen Ideale von $K$. Ist andersherum $R$ ein kommutativer Ring, in dem $\{0\}$ und $R$ die einzigen Ideale sind, dann ist $R$ (warum?) ein Körper.

(4) Ist $f\colon R \to S$ ein Ringhomomorphismus, dann ist $\operatorname{Ker}(f) \subseteq R$ ein Ideal. Wir wissen bereits, dass es sich um eine Untergruppe von $(R,+)$ handelt, da $f$ insbesondere ein Gruppenhomomorphismus ist. Außerdem gilt für $x \in R$, $a \in \operatorname{Ker}(f)$, dass $f(xa) = f(x)f(a) = 0$, also $xa \in \operatorname{Ker}(f)$, und genauso zeigt man $ax \in \operatorname{Ker}(f)$. Wir werden später sehen, dass für jeden Ring $R$ und jedes Ideal $\mathfrak{a} \subseteq R$ ein Ringhomomorphismus $R \to S$ mit Kern $\mathfrak{a}$ existiert. (Siehe Abschnitt 18.4.)

◊

Beispiel 15.19. Wir betrachten den Ring $\mathbb{Z}$ der ganzen Zahlen. Ist $d \in \mathbb{Z}$, so ist die Menge

$$(d) := \{xd;\ x \in \mathbb{Z}\}$$

aller Vielfachen von $d$ ein Ideal (und wir werden in Satz 15.39 sehen, dass im Ring $\mathbb{Z}$ alle Ideale diese Form haben). ◊

Ergänzung 15.20. Der Begriff *Ideal* geht auf Ernst Kummer[3] zurück, der ihn im Bereich der Zahlentheorie einführte und als Abkürzung für »ideale Zahlen« verstand. Dort treten Ringe auf, in denen das Analogen der eindeutigen Primzaktorzerlegung zwar nicht mehr für die Elemente des Rings gilt, aber wo man eine analoge Aussage für die Ideale des Rings beweisen kann. Siehe auch Ergänzung 15.55. □ Ergänzung 15.20

Der Durchschnitt von Idealen ist wieder ein Ideal. Wir erhalten so den Begriff des von einer Teilmenge von $R$ erzeugen Ideals.

Definition 15.21. Sei $R$ ein Ring und sei $M \subseteq R$ eine Teilmenge. Wir schreiben $(M)$ für den Durchschnitt aller Ideale von $R$, die $M$ als Teilmenge enthalten, und nennen $(M)$ das *von der Teilmenge $M$ erzeugte Ideal*. Es handelt sich dabei um das kleinste Ideal von $R$, das $M$ enthält, das heißt: Ist $\mathfrak{a} \subseteq R$ ein Ideal mit $M \subseteq \mathfrak{a}$, so gilt $(M) \subseteq \mathfrak{a}$. ⊣

Im Fall $M = \{x_1, \ldots, x_n\}$ schreibt man auch $(x_1, \ldots, x_n)$ statt $(\{x_1, \ldots, x_n\})$. Der Fall von Idealen, die von einem einzigen Element erzeugt werden, ist besonders wichtig; diese Ideale nennt man *Hauptideale*. Ist $R$ ein kommutativer Ring und $a \in R$, so gilt

$$(a) = \{xa;\ x \in R\}.$$

Es ist $(0) = \{0\}$ das Nullideal und $(1) = R$ das Einsideal von $R$.

In einem kommutativen Ring kann man die Elemente eines Ideals der Form $(x_1, \ldots, x_n)$ ähnlich explizit beschreiben wie die Elemente eines von einer Menge erzeugten Untervektorraums in einem Vektorraum. Es gilt

$$(x_1, \ldots, x_n) = \left\{ \sum_{i=1}^{n} a_i x_i;\ a_i \in R \right\},$$

denn die rechte Seite ist, wie man nachrechnet, ein Ideal, und es ist klar, dass sie in der linken Seite enthalten ist.

[3] https://de.wikipedia.org/wiki/Ernst_Eduard_Kummer

## 15.3. Der Polynomring über einem (kommutativen) Ring

Ist $K$ ein Körper und $A$ eine quadratische Matrix in $M_n(K)$, dann möchten wir die Polynomfunktion $K \to K, \lambda \mapsto \det(A - \lambda E_n)$, untersuchen (bzw. die Funktion $\lambda \mapsto \det(\lambda E_n - A)$, die sich später als etwas »schöner« erweist und sich von der vorgenannten Funktion nur um den Faktor $(-1)^n$ unterscheidet), um die Eigenwerte von $A$ zu untersuchen. Der Ring der Polynomfunktionen $K \to K$ hat aber (im Fall endlicher Körper) einige unschöne Eigenschaften (es ist kein Integritätsring im Sinne von Definition 15.28 unten). Es ist daher nützlich, eine Variante dieses Rings einzuführen, den sogenannten Polynomring.

Sei $R$ ein kommutativer Ring. Wir wollen den *Polynomring* über $R$ definieren, wobei wir uns ein Polynom als einen »formalen Ausdruck« der Form

$$\sum_{i=0}^{n} a_i X^i, \quad a_i \in R,$$

vorstellen, also als eine Linearkombination von Potenzen der »Unbestimmten« $X$ mit Koeffizienten $a_i \in R$. Dabei sollen zwei Polynome genau dann gleich sein, wenn alle Koeffizienten gleich sind (wobei wir erlauben, zusätzliche Summanden $0 \cdot X^r$ hinzuzufügen, um auch zwei Polynome vergleichen zu können, in denen die Summationsgrenzen unterschiedlich sind). Der Begriff des Polynoms wird sich daher im allgemeinen Fall vom Begriff der Polynomfunktion (Abschnitt I.4.3) unterscheiden, siehe Bemerkung 15.27.

Es ist auch klar, wie wir mit Polynomen »rechnen« möchten, d.h. wie die Addition und Multiplikation von Polynomen vonstatten gehen sollte: Polynome werden »koeffizientenweise« addiert, d.h.

$$\sum_{i=0}^{n} a_i X^i + \sum_{i=0}^{n} b_i X^i = \sum_{i=0}^{n} (a_i + b_i) X^i,$$

wobei man beachte, dass wir uns durch »Auffüllen mit Nullen« immer auf den Fall zurückziehen können, dass beide Summen denselben Summationsbereich haben. Die Multiplikation ist eindeutig dadurch festgelegt, dass das Distributivgesetz gelten soll, und dass

$$X^i \cdot X^j = X^{i+j} \qquad \text{für alle } i, j \geq 0$$

gelten soll. Es folgt dann

$$\left( \sum_{i=0}^{m} a_i X^i \right) \cdot \left( \sum_{i=0}^{n} b_i X^i \right) = \sum_{i=0}^{n} \left( \sum_{j+k=i} a_j b_k \right) X^i$$

für das Produkt von zwei allgemeinen Polynomen. Dabei ist $0 \leq j \leq m$, $0 \leq k \leq n$.

Ein technisches Problem bei der ganzen Sache ist, wie man das Symbol $X$ in die Definition einbaut, bzw. was $X$ eigentlich »ist«. Die Lösung, die wir wählen, ist, das $X$ zunächst einmal zu vergessen. Ein Polynom soll ja durch seine Koeffizienten festgelegt sein und wir müssen nur beschreiben, wie mit Tupeln von Koeffizienten gerechnet werden soll. Danach können wir das Element $X$ des Polynomrings definieren als das Polynom mit Koeffizienten $a_i = 0$ für alle $i \neq 1$ und $a_1 = 1$. In der Tupelschreibweise schreiben wir die Koeffizienten in der Reihenfolge $(a_0, a_1, a_2, \dots)$.

Definition 15.22. Der *Polynomring* $R[X]$ über $R$ in der Unbestimmten $X$ ist der Ring aller Folgen $(a_i)_{i \in \mathbb{N}}$ mit nur endlich vielen Einträgen $\neq 0$, mit elementweiser Addition und der Multiplikation

$$(a_i)_i \cdot (b_i)_i = \left( \sum_{j+k=i} a_j b_k \right)_i .$$

Dies ist ein kommutativer Ring mit $1 = (1, 0, 0, \dots)$ (und $0 = (0, 0, \dots)$). Die Elemente von $R[X]$ heißen *Polynome*.

Wir setzen $X := (0, 1, 0, 0, \dots) \in R[X]$ und erhalten dann

$$(a_0, a_1, a_2, \dots) = \sum_{i \geq 0} a_i X^i,$$

wobei nur endlich viele $a_i$ von Null verschieden sein dürfen. Insbesondere können wir jedes Element von $R[X]$ in eindeutiger Weise in der Form $\sum_{i \geq 0} a_i X^i$ schreiben (fast alle $a_i = 0$). ⊣

Es ist nicht schwer nachzurechnen, dass für diese Verknüpfungen tatsächlich alle Ringaxiome erfüllt sind. Der Ring $R[X]$ ist ein kommutativer Ring.

Die Abbildung $R \to R[X], a \mapsto (a, 0, 0, \dots)$ ist ein injektiver Ringhomomorphismus und wir fassen vermöge dieses Homomorphismus Elemente von $R$ als Elemente von $R[X]$ auf. Diese Elemente heißen *konstante Polynome*.

An Stelle von $X$ kann man natürlich auch andere Buchstaben verwenden, um die Unbestimmte zu bezeichnen, wir können also auch von den Polynomringen $R[x]$, $R[T]$, usw. sprechen.

Bemerkung 15.23. Achtung: Ist $S$ ein Ring, $R \subseteq S$ ein Unterring und $\alpha \in S$, dann verwendet man die eckigen Klammern auch mit einer etwas anderen (allgemeineren) Bedeutung, und zwar bezeichnet $R[\alpha]$ dann nicht den Polynomring in der Unbestimmten $\alpha$ (was ja auch problematisch wäre, weil dann $\alpha$ zwei verschiedene Bedeutungen hätte), sondern den Unterring von $S$, der aus allen polynomialen Ausdrücken in $\alpha$ besteht:

$$R[\alpha] = \left\{ \sum_{i=0}^{n} a_i \alpha^i;\ n \in \mathbb{N}, a_i \in R \right\} \subseteq S.$$

Beispiele dafür sind die Ringe $K[A]$ und $K[f]$ aus Beispiel 15.14. Ein anderes Beispiel ist der Körper $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2};\ a, b \in \mathbb{Q}\} \subseteq \mathbb{R}$ – hier sind die höheren Potenzen von $\sqrt{2}$ (warum?) verzichtbar. Allgemeiner verwendet man diese Notation auch, wenn $\varphi\colon R \to S$ ein (nicht notwendig injektiver) Ringhomomorphismus ist, für $\alpha \in S$ schreibt man dann

$$R[\alpha] = \left\{ \sum_{i=0}^{n} \varphi(a_i) \alpha^i;\ n \in \mathbb{N}, a_i \in R \right\} \subseteq S.$$

Mit der in Satz 15.24 eingeführten Terminologie ist also $R[\alpha] \subseteq S$ das Bild des Einsetzungshomomorphismus $R[X] \to S, f \mapsto f(\alpha)$, der durch $X \mapsto \alpha$ und $\varphi\colon R \to S$ gegeben ist.

Wenn man möchte, dann kann man die Schreibweise $R[X]$ als Spezialfall der hier beschriebenen Notation betrachten, denn der Ring $R[X]$ besteht ja genau aus allen polynomialen Ausdrücken in $X$ mit Koeffizienten in $R$. ◊

Allgemeiner kann man Polynomringe in mehr als einer Unbestimmten definieren, etwa $R[X_1, X_2, \dots, X_n]$ oder sogar $R[X_i,\ i \in I]$ für eine beliebige Menge $I$. Man kann dabei den Fall, dass die Indexmenge $I$ unendlich viele Elemente hat, zulassen; es werden aber nur endliche Summen und Produkte der Unbestimmten und ihrer Potenzen gebildet, d.h., dass in jedem einzelnen Polynom nur endlich viele der Unbestimmten $X_i$ auftreten können.

Ist $f = \sum_{i=0}^{n} a_i X^i \in R[X]$ ein Polynom mit Koeffizienten in $R$ und $x \in R$, so können wir $x$ für die Unbestimmte $X$ »einsetzen«: Wir definieren

$$f(x) := \sum_{i=0}^{n} a_i x^i \in R.$$

Im folgenden Satz wird das noch etwas verallgemeinert und präzisiert. Erstens können wir nicht nur Elemente aus $R$ einsetzen, sondern Elemente aus einem Ring $S$, sobald wir »wissen, wie die Koeffizienten (aus $R$) als Elemente von $S$ aufgefasst« werden sollen. Formal verlangen wir, dass ein Ringhomomorphismus $R \to S$ gegeben ist. Zweitens erhalten wir für fixiertes $x$

auf diese Weise einen *Ringhomomorphismus* $R[X] \to R$ (bzw. im allgemeineren Fall $R[X] \to S$), das heißt $(f+g)(x) = f(x) + g(x)$, $(fg)(x) = f(x)g(x)$ und für $f = 1$ gilt $f(x) = 1$.

SATZ 15.24 (Einsetzungshomomorphismus). *Sei $R$ ein kommutativer Ring, $\varphi\colon R \to S$ ein Ringhomomorphismus und $x \in S$. Dann existiert ein eindeutig bestimmter Ringhomomorphismus $\Phi\colon R[X] \to S$ mit $\Phi(a) = \varphi(a)$ für alle $a \in R$ und $\Phi(X) = x$, nämlich*

$$\sum_{i=0}^{n} a_i X^i \mapsto \sum_{i=0}^{n} \varphi(a_i) x^i.$$

BEWEIS. Aus den Bedingungen $\Phi(a) = \varphi(a)$ für alle $a \in R$ und $\Phi(X) = x$ ergibt sich, weil $\Phi$ ein Ringhomomorphismus ist, die angegebene Formel für das Bild eines beliebigen Polynoms unter $\Phi$. Es ist also nur noch zu zeigen, dass diese Formel wirklich einen Ringhomomorphismus beschreibt. Das folgt aus einer einfachen direkten Rechnung, die ausnutzt, dass $\varphi$ ein Ringhomomorphismus ist. □

Wir schreiben in der Situation des Satzes auch $f(x) = \Phi(f)$.

Die Abbildung $\varphi$ wird oftmals nicht explizit angegeben, wenn »klar« ist, um welche Abbildung es sich handelt. Die drei (für uns) wichtigsten Fälle sind

(1) $R = S$ und $\varphi = \mathrm{id}_R$,

(2) $R \subseteq S$ ist ein Unterring und $\varphi$ ist die Inklusionsabbildung $R \to S, x \mapsto x$.

(3) $R = K$ ist ein Körper, $S = M_n(K)$ der Matrizenring ($n \in \mathbb{N}$), und $\varphi\colon K \to M_n(K)$ ist gegeben durch $a \mapsto aE_n$.

BEISPIEL 15.25. (1) Sei $K = \mathbb{Q}$, $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ und $f = X^2 - 5X + 5$. Dann ist (mit $\varphi$ wie in Punkt (3) der vorhergehenden Liste)

$$f(A) = A^2 - 5E_2A + 5E_2 = A^2 - 5A + 5E_2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} - \begin{pmatrix} 5 & 10 \\ 15 & 20 \end{pmatrix} + \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 7 & 0 \\ 0 & 7 \end{pmatrix}.$$

(2) Die Ringe $K[A]$ und $K[f]$ aus Beispiel 15.14 sind gerade die Bilder der Einsetzungshomomorphismen

$$K[X] \to M_n(K),\ X \mapsto A, \quad \text{und} \quad K[X] \to \mathrm{End}_K(V),\ X \mapsto f.$$

Dass es sich um Ringhomomorphismen handelt, besagt, dass die Multiplikation von Polynomen der Multiplikation in $M_n(A)$ (also dem Matrizenprodukt) bzw. in $\mathrm{End}_K(V)$ (also der Verkettung von Endomorphismen) entspricht. Zum Beispiel wird unter dem rechten Ringhomomorphismus das Polynom $X^2 - 1$ auf den Endomorphismus $f^2 - \mathrm{id}_V$ abgebildet:

$$f^2 - \mathrm{id}_V\colon V \to V, \quad v \mapsto f(f(v)) - v.$$

◊

DEFINITION 15.26. Sei $R$ ein kommutativer Ring, $f = \sum_{i=0}^{N} a_i X^i \in R[X]$ mit $a_N \neq 0$. Dann heißt $a_N$ der *Leitkoeffizient* von $f$ und $N$ der Grad von $f$, in Zeichen $\deg f$. Das Element $a_0$ heißt der *Absolutkoeffizient* (oder: das *absolute Glied*) von $f$. Ein *normiertes* Polynom ist ein Polynom, dessen Leitkoeffizient gleich 1 ist.

Wir setzen formal $\deg 0 = -\infty$. (Dass das eine gute Idee ist, ergibt sich in Kürze aus Lemma 15.30.) Es ist also für $f \in R[X]$ der Grad $\deg(f)$ genau dann $\geq 0$, wenn $f \neq 0$ gilt. ⊣

Ein Polynom vom Grad 1 heißt auch *lineares Polynom*, unter einem *quadratischen Polynom* versteht man ein Polynom vom Grad 2. Manchmal spricht man auch von *kubischen Polynomen* im Sinne von Polynomen vom Grad 3.

BEMERKUNG 15.27. Sei $R$ ein Ring. Ist $f \in R[X]$ ein Polynom, so erhalten wir die Abbildung $R \to R, x \mapsto f(x)$. Abbildungen dieser Form nennen wir *Polynomfunktionen*. Die Polynomfunktionen bilden einen Unterring $\mathrm{Pol}(R)$ des Rings $\mathrm{Abb}(R, R)$ (siehe Beispiel 15.3).

Die Abbildung

$$R[X] \to \mathrm{Pol}(R),$$

die $f \in R[X]$ abbildet auf die zugehörige Polynomfunktion $x \mapsto f(x)$, ist ein Ringhomomorphismus vom Polynomring $R[X]$ in den Ring der Polynomfunktionen $R \to R$, der nach Definition von $\mathrm{Pol}(R)$ surjektiv, aber im allgemeinen nicht injektiv ist. Ist $R$ ein Körper mit unendlich vielen Elementen, so ist dieser Ringhomomorphismus ein Isomorphismus, siehe Korollar I.4.28.

Über einem endlichen Körper $K$ hat es gewisse Vorteile, mit dem Ring $K[X]$ zu arbeiten, der – wie wir in den nachfolgenden Abschnitten sehen werden – eine relativ einfache Struktur hat. Insbesondere gilt für $f, g \in K[X]$ mit $f, g \neq 0$, dass auch das Produkt $fg \neq 0$ ist. Diese wichtige Eigenschaft besprechen wir im folgenden Abschnitt über *Integritätsringe*. ◊

## 15.4. Integritätsbereiche

**15.4.1. Definition.** Sei $R$ ein Ring. In diesem Abschnitt betrachten wir nur kommutative Ringe. Wir haben schon Beispiele von Ringen gesehen, in denen so genannte Nullteiler existieren – Elemente $x$, so dass $xy = 0$ für ein $y \neq 0$ – die von 0 verschieden sind. Das ist sozusagen eine unangenehme Eigenschaft, und wir werden uns daher an vielen Stellen auf nullteilerfreie Ringe einschränken, also auf Ringe, in denen 0 der einzige Nullteiler ist. Wir machen dafür die folgende Definition.

DEFINITION 15.28. Ein kommutativer Ring $R$ heißt *Integritätsbereich* (oder *Integritätsring*), wenn $R \neq \{0\}$ und für alle $x, y \in R$ mit $xy = 0$ gilt: $x = 0$ oder $y = 0$. ⊣

BEISPIEL 15.29. Der Ring $\mathbb{Z}$ und alle Körper sind Integritätsbereiche. Der Ring $\mathbb{Z}/n$ ist genau dann ein Integritätsring, wenn $n$ eine Primzahl ist. In diesem Fall ist $\mathbb{Z}/n$ ja sogar ein Körper. Andernfalls können wir $n = ab$ mit $1 < a, b < n$ schreiben, und dann gilt in $\mathbb{Z}/n$, dass $a, b \neq 0$ aber $ab = 0$ ist. ◊

LEMMA 15.30. *Sei $R$ ein kommutativer Ring und seien $f, g \in R[X]$. Dann gilt*

(1) $\deg(f + g) \leq \max(\deg f, \deg g)$,

(2) $\deg(fg) \leq \deg f + \deg g$, *und falls $R$ ein Integritätsbereich ist, so gilt sogar die Gleichheit.*

Wie wir sehen werden, gelten die Aussagen des Lemmas (mit unserer Definition $\deg(0) = -\infty$) auch für den Fall, dass $f$ oder $g$ das Nullpolynom ist, wenn man mit $-\infty$ in der »offensichtlichen« Weise rechnet, das heißt es gelte

$$-\infty \leq -\infty, \quad -\infty \leq n \text{ für alle } n \in \mathbb{N},$$

und

$$-\infty + (-\infty) = -\infty, \quad -\infty + n = n + (-\infty) = -\infty \text{ für alle } n \in \mathbb{N}.$$

Insbesondere ist dann $\max(-\infty, n) = n$ für alle $n \in \mathbb{N} \cup \{-\infty\}$.

BEWEIS. Es ist klar, dass für $f = 0$ oder $g = 0$ beide Aussagen richtig sind (und das erklärt, warum es sinnvoll ist, dem Nullpolynom auf diese formale Art den Grad $-\infty$ zuzuweisen).

Nun gelte $f \neq 0$ und $g \neq 0$. Wir schreiben

$$f(X) = \sum_{i=0}^{m} a_i X^i, \qquad g(X) = \sum_{i=0}^{n} b_i X^i$$

mit $a_m \neq 0$ und $b_n \neq 0$. Ist $m \neq n$, so ist der Grad von $f + g$ gleich der größeren der beiden Zahlen $m$ und $n$. Ist $m = n$, dann ist ebenfalls $\deg(f + g) = \max(m, n)$, es sei denn, es gilt $a_m = -b_n$. Im letzteren Fall ist $\deg(f + g) < \max(m, n)$. Damit ist Teil (1) bewiesen.

Für Teil (2) müssen wir nur beobachten, dass

$$f(X)\, g(X) = \sum_{i=0}^{m+n} \left( \sum_{j+k=i} a_j b_k \right) X^i$$

gilt, und daher jedenfalls $\deg(fg) \leq m + n = \deg f + \deg g$ ist. Weil $j, k \geq 0$ gilt, hat die Summe für $i = m + n$ nur den einen Summanden $a_m b_n$. Ist $R$ ein Integritätsring, so ist das Produkt $a_m b_n \neq 0$, und es folgt $\deg(fg) = m + n$. □

KOROLLAR 15.31. *Sei $R$ ein Integritätsring. Dann ist auch $R[X]$ ein Integritätsring. Es gilt $R[X]^\times = R^\times$.*

BEWEIS. Es folgt aus Lemma 15.30, dass das Produkt von zwei Polynomen $f, g \in R[X] \setminus \{0\}$ nicht $= 0$ sein kann. Es ist auch klar, dass $R[X]$ nicht der Nullring ist, sofern $R$ nicht der Nullring ist. Also ist $R[X]$ ein Integritätsring.

Ist $f \in R[X]^\times$, so existiert $g \in R[X]$ mit $fg = 1$, also ist $\deg(fg) = 0$. Aus Lemma 15.30 folgt dann $\deg(f) = \deg(g) = 0$ (hier benutzen wir erneut, dass $R$ ein Integritätsring ist!), also sind $f$ und $g$ konstante Polynome und es folgt $f \in R^\times$. □

Machen Sie sich klar, dass für einen endlichen Körper $K$ der Ring $\mathrm{Pol}(K)$ der Polynomfunktionen $K \to K$ (siehe Bemerkung 15.27) *kein* Integritätsring ist.

**15.4.2. Teilbarkeit in Integritätsringen.** Eine wichtige Eigenschaft von Integritätsringen ist die sogenannte Kürzungsregel.

LEMMA 15.32. *Ist $R$ ein Integritätsring, und sind $a, b, c \in R$ mit $a \neq 0$ und $ab = ac$, so folgt $b = c$.*

BEWEIS. Aus $ab = ac$ folgt $a(b - c) = ab - ac = 0$, also $b - c = 0$, weil wir $a \neq 0$ vorausgesetzt haben und $R$ ein Integritätsring ist. □

Wir wollen nun den Begriff des Teilers, den wir vom Ring der ganzen Zahlen her kennen, für allgemeine Integritätsringe definieren.

DEFINITION 15.33. Sei $R$ ein Integritätsring. Seien $a, b \in R$

(1) Wir sagen, $a$ sei ein *Teiler* von $b$ (oder $b$ sei durch $a$ *teilbar*, in Zeichen $a \mid b$), falls $c \in R$ existiert mit $ac = b$. Es ist äquivalent zu sagen, dass $b$ ein *Vielfaches* von $a$ sei. Wenn $a$ kein Teiler von $b$ ist, dann schreiben wir $a \nmid b$.

(2) Wir nennen $a, b$ zueinander *assoziiert*, falls $c \in R^\times$ existiert mit $ac = b$.

⊣

Da das Element $c$ in Teil (2) der Definition eine Einheit sein muss, können wir die Gleichung $ac = b$ auch umschreiben als $bc^{-1} = a$; wie die Sprechweise suggeriert, kommt es also nicht auf die Reihenfolge von $a$ und $b$ an. (Die Relation »assoziiert zu« ist symmetrisch, Abschnitt I.3.14, Definition I.3.67, siehe auch Definition 15.64 unten.)

LEMMA 15.34. *Seien $R$ ein Integritätsring und $a, b \in R$.*

(1) *Es sind äquivalent:*
   (i) $a \mid b$,
   (ii) $b \in (a)$,
   (iii) $(b) \subseteq (a)$.

(2) *Es sind äquivalent:*
   (i) *$a$ und $b$ sind assoziiert zueinander,*
   (ii) $a \mid b$ *und* $b \mid a$,
   (iii) $(a) = (b)$.

BEWEIS. Der Beweis von Teil (1) ist einfach. In Teil (2) ist klar, dass für assoziierte Elemente $a$ und $b$ gilt, dass $(a) = (b)$ ist. Wegen Teil (1) ist das äquivalent zu der Bedingung, dass $a \mid b$ und $b \mid a$. Gilt umgekehrt $(a) = (b)$, etwa $b = ca$ und $a = db$, so folgt $a = cda$ und damit $(1 - cd)a = 0$. Weil $R$ ein Integritätsring ist, folgt $a = 0$ (also auch $b = 0$) oder $1 - cd = 0$, und das impliziert, dass $c$ und $d$ Einheiten von $R$ sind, also dass $a$ und $b$ zueinander assoziiert sind. □

Grundlegende Eigenschaften der Teilbarkeit wie die folgenden lassen sich dann leicht beweisen:

$$a \mid b,\ b \mid c \quad \Longrightarrow \quad a \mid c$$

und

$$a \mid b,\ a \mid c \quad \Longrightarrow \quad a \mid (b + c)$$

für alle $a, b, c \in R$.

Es stellt sich heraus, dass der Begriff des Integritätsrings so allgemein ist, dass keine allgemeine »vernünftige« Theorie von Teilbarkeit zu erwarten ist (konkret: im allgemeinen gibt es kein analoges Ergebnis zur eindeutigen Primfaktorzerlegung, die wir in $\mathbb{Z}$ haben). Besonders gut verhalten sich Integritätsringe, in denen wir eine Division mit Rest, ähnlich wie in $\mathbb{Z}$, haben.

Im Ring der ganzen Zahlen können wir *Division mit Rest* durchführen: Sind $a$ und $b$ ganze Zahlen, so existieren $q, r \in \mathbb{Z}$ mit $a = qb + r$ und $|r| < |b|$. Dabei sind $q$ und $r$ sogar eindeutig bestimmt: Es ist $q$ die größte ganze Zahl, die $\leq \frac{a}{b}$ ist, und $r = a - qb$. Die Division mit Rest ist eine essenzielle Eigenschaft des Rings der ganzen Zahlen, aus der sich viele nützliche Eigenschaften folgern lassen, und es ist daher naheliegend zu untersuchen, ob es in anderen Ringen eine ähnliche »Division mit Rest« gibt. (Siehe auch Ergänzung I.3.44.)

SATZ 15.35 (Polynomdivision). *Sei $R$ ein kommutativer Ring und seien $f, g \in R[X]$, so dass der Leitkoeffizient von $g$ in $R^\times$ liegt. Dann existieren eindeutig bestimmte Polynome $q, r \in R[X]$ mit* $\deg r < \deg g$ *und so dass*

$$f = qg + r.$$

Für uns ist vor allem der Fall wichtig, dass $R$ ein Körper ist. In diesem Fall ist die Bedingung, dass der Leitkoeffizient von $g$ eine Einheit ist, dazu äquivalent, dass $g \neq 0$ gilt.

BEWEIS. Wir führen Induktion nach dem Grad von $f$. Die Voraussetzung an $g$ impliziert insbesondere, dass $g \neq 0$, also $\deg(g) \in \mathbb{N}$ ist. Ist $\deg(f) < \deg(g)$, so können wir einfach $q = 0, r = f$ setzen.

Sei nun $m := \deg(f) \geq \deg(g) =: n$. Insbesondere ist dann $f \neq 0$. Sei $a \in R$ der Leitkoeffizient von $f$ und $b \in R^\times$ der Leitkoeffizient von $g$. Dann ist

$$h := f - ab^{-1}X^{m-n}g$$

ein Polynom vom Grad $< m$, denn $f$ und $ab^{-1}X^{m-n}g$ sind Polynome vom Grad $m$ mit demselben Leitkoeffizienten $a$. Nach Induktionsvoraussetzung können wir $h$ in der Form $q_1 g + r$ mit $\deg(r) < \deg(g)$ schreiben. Wir setzen dann $q := q_1 + ab^{-1}X^{m-n}$ und erhalten

$$f = h + ab^{-1}X^{m-n}g = q_1 g + r + ab^{-1}X^{m-n}g = qg + r.$$

Die Eindeutigkeit kann man folgendermaßen begründen: Ist

$$f = q_1 g + r_1 = q_2 g + r_2$$

mit $\deg(r_1), \deg(r_2) < \deg(g)$, dann folgt

$$r_2 - r_1 = (q_1 - q_2)g,$$

und das ist aus Gradgründen nur möglich, wenn $q_1 - q_2 = 0$ ist. Also ist $q_1 = q_2$ und damit auch $r_1 = r_2$.

Vergleiche auch Lemma I.4.26 und Beispiel I.4.27. □

DEFINITION 15.36. Ein Integritätsring $R$ heißt *euklidischer Ring*, falls eine Abbildung

$$\delta\colon R \setminus \{0\} \to \mathbb{N}$$

(eine sogenannte *Gradabbildung*) existiert, so dass für alle $a, b \in R, b \neq 0$, Elemente $q, r \in R$ existieren, so dass $r = 0$ oder $\delta(r) < \delta(b)$ und $a = qb + r$. ⊣

Es wird in der Definition nicht verlangt, dass $q$ und $r$ für gegebene $a$ und $b$ eindeutig bestimmt sind.

BEISPIEL 15.37. (1) Der Ring $\mathbb{Z}$ ist euklidisch, als Gradfunktion können wir den Absolutbetrag verwenden: $\delta(a) = |a|$. Das folgt daraus, dass wir im Ring $\mathbb{Z}$ die Division mit Rest haben.

(2) Sei $K$ ein Körper. Dann ist der Polynomring $K[X]$ mit der Gradfunktion $\delta(f) = \deg(f)$ ein euklidischer Ring. Dies folgt daraus, dass wir im Ring $K[X]$ die Polynomdivision durchführen können.

(3) Der Ring $\mathbb{Z}[i] = \{a + ib;\ a, b \in \mathbb{Z}\}$ ist euklidisch (siehe die Übungsaufgaben).

◊

> Menschen, die von der Algebra nichts wissen, können sich auch nicht die wunderbaren Dinge vorstellen, zu denen man mit Hilfe der genannten Wissenschaft gelangen kann.
>
> Gottfried Wilhelm Leibniz
>
> Fundort: http://www.mathe.tu-freiberg.de/~hebisch/cafe/zitate.html

DEFINITION 15.38. (1) Ein Ideal $\mathfrak{a}$ in einem Ring $R$ heißt *Hauptideal*, wenn ein Element $a \in R$ existiert, so dass $\mathfrak{a} = (a) := \{xa;\ x \in R\}$.

(2) Ein Integritätsring $R$ heißt *Hauptidealring*, wenn jedes Ideal in $R$ ein Hauptideal ist.

⊣

Der Erzeuger eines Hauptideals ist in der Regel nicht eindeutig bestimmt. Ist $R$ ein Integritätsring, so folgt aus Lemma 15.34, dass Elemente $a, b$ genau dann dasselbe Hauptideal erzeugen, wenn sie zueinander assoziiert sind.

SATZ 15.39. *Jeder euklidische Ring ist ein Hauptidealring. Insbesondere gilt:*

(1) *Der Ring $\mathbb{Z}$ ist ein Hauptidealring.*

(2) *Ist $K$ ein Körper, dann ist der Polynomring $K[X]$ in einer Unbestimmten über $K$ ein Hauptidealring.*

BEWEIS. Sei $R$ ein euklidischer Ring mit Gradfunktion $\delta$. Sei $\mathfrak{a} \subseteq R$ ein Ideal. Ist $\mathfrak{a}$ das Nullideal, dann handelt es sich trivialerweise um ein Hauptideal: $\mathfrak{a} = (0)$. Andernfalls sei $a \in \mathfrak{a} \setminus \{0\}$ ein Element, für das der Wert $\delta(a)$ minimal ist. Wir wollen zeigen, dass $\mathfrak{a} = (a)$ gilt. Die Inklusion $\supseteq$ ist klar, weil $a$ nach Definition in $\mathfrak{a}$ liegt.

Sei nun $x \in \mathfrak{a}$. Wir benutzen jetzt, dass $R$ euklidisch ist und schreiben $x = qa + r$ mit $r = 0$ oder $\delta(r) < \delta(a)$. Ist $r = 0$, so folgt $x = qa \in (a)$, wie gewünscht. Der Fall $r \neq 0$, $\delta(r) < \delta(a)$ kann gar nicht eintreten, denn es ist $r = x - qa \in \mathfrak{a}$, und $a$ war so gewählt, dass kein Element aus $\mathfrak{a} \setminus \{0\}$ unter $\delta$ einen kleineren Wert als $\delta(a)$ annimmt. □

BEISPIEL 15.40. Der Ring $\mathbb{Z}[X]$ ist kein Hauptidealring (zum Beispiel ist das Ideal $(2, X)$ kein Hauptideal – warum?). Insbesondere ist $\mathbb{Z}[X]$ kein euklidischer Ring: Die Funktion deg ist keine Gradabbildung mit den in der Definition euklidischer Ringe geforderten Eigenschaften, *und es gibt auch keine andere Abbildung* $\mathbb{Z}[X] \setminus \{0\} \to \mathbb{N}$, die diese Eigenschaften hat.

Insbesondere sehen wir, dass der Polynomring über einem Integritätsring $R$ nicht unbedingt ein euklidischer Ring. Wenn man das Studium der Ringtheorie noch ein kleines bosschen weiterführt, kann man zeigen, dass $R[X]$ genau dann ein Hauptidealring ist, wenn $R$ ein Körper ist. ◊

Es gibt auch Hauptidealringe, die nicht euklidisch sind, es ist aber nicht ganz einfach, hierfür Beispiele zu geben (siehe zum Beispiel [**Sch**] 6.10).

DEFINITION 15.41. Sei $R$ ein Integritätsring, seien $a, b \in R$.

(1) Ein Element $d \in R$ heißt *größter gemeinsamer Teiler* von $a, b$, wenn $d \mid a$, $d \mid b$, und für jedes Element $d'$, das $a$ und $b$ teilt, $d' \mid d$. Man schreibt oft $\mathrm{ggT}(a, b)$ für einen größten gemeinsamen Teiler von $a$ und $b$ (aber siehe die folgende Bemerkung – diese Notation ist nicht ganz unproblematisch!).

(2) Ein Element $d \in R$ heißt *kleinstes gemeinsames Vielfaches* von $a, b$, wenn $a \mid d$, $b \mid d$, und für jedes Element $d'$, das von $a$ und $b$ geteilt wird, $d \mid d'$. Man schreibt oft $\mathrm{kgV}(a, b)$ für ein kleinstes gemeinsames Vielfaches von $a$ und $b$ (aber siehe die folgende Bemerkung – diese Notation ist nicht ganz unproblematisch!).

(3) Die Elemente $a, b$ heißen *teilerfremd*, falls 1 ein größter gemeinsamer Teiler von $a$ und $b$ ist.

⊣

Man beachte, dass das Zeichen $>$ in der Definition des Begriffs des größten gemeinsamen Teilers nicht auftritt – in einem allgemeinen Integritätsring steht uns ja keine Anordnung der Elemente zur Verfügung. Angewandt auf den Ring der ganzen Zahlen stimmt die obige Definition aber mit der üblichen Definition überein (siehe Lemma I.3.53). (Wenn Sie den Begriff der partiellen Ordnung kennen (Abschnitt I.3.14.3), dann ist die »richtige« Sichtweise, dass der größte gemeinsame Teiler von zwei Elementen das größte Element unter allen gemeinsamen Teilern bezüglich der durch Teilbarkeit gegebenen partiellen Ordnung ist (wenn ein solches größtes Element existiert). Siehe Beispiel I.3.81.)

BEMERKUNG 15.42. Sei $R$ ein Integritätsring.

(1) Sind $a, b \in R$ und erfüllen $d_1$ und $d_2$ die Eigenschaft eines größten gemeinsamen Teilers, dann gilt $d_1 \mid d_2$ und $d_2 \mid d_1$, also sind $d_1$ und $d_2$ zueinander assoziiert. Andererseits ist für jeden größten gemeinsamen Teiler $d$ von $a$ und $b$ und jede Einheit $u \in R^\times$ offenbar auch $ud$ ein größter gemeinsamer Teiler von $a$ und $b$. Ähnlich verhält es sich mit dem kleinsten gemeinsamen Vielfachen.

Weil größter gemeinsamer Teiler und kleinstes gemeinsames Vielfaches nur bis auf Multiplikation mit Einheiten aus $R$ eindeutig bestimmt sind, ist es eine ungenaue Notation, $d = \operatorname{ggT}(a, b)$ zu schreiben (und entsprechend für $\operatorname{kgV}(a, b)$).

Zum Beispiel sind im Ring $\mathbb{Z}$ sowohl 2 als auch $-2$ ein größter gemeinsamer Teiler von $-6$ und 14.

(2) Im allgemeinen müssen ein größter gemeinsamer Teiler bzw. ein kleinstes gemeinsames Vielfaches zweier Elemente nicht existieren. Selbst wenn ein größter gemeinsamer Teiler $d$ von $a, b \in R$ existiert, kann man $d$ im allgemeinen nicht in der Form $xa + yb$ ausdrücken (wie es im Ring der ganzen Zahlen möglich ist, siehe Lemma I.3.53 bzw. den folgenden Punkt (3)). Im allgemeinen folgt aus der Bedingung, dass 1 ein größter gemeinsamer Teiler von $a$ und $b$ ist, also *nicht*, dass das von $a$ und $b$ erzeugte Ideal das Einsideal ist.

(3) Ein Element $d \in R$ ist genau dann ein gemeinsamer Teiler von $a$ und $b$, wenn $(a, b) \subseteq (d)$ gilt (siehe Lemma 15.34). Wenn $(a, b) = (d)$ ein Hauptideal ist, dann folgt mit demselben Lemma, dass $d$ ein größter gemeinsamer Teiler von $a$ und $b$ ist.

Wir sehen insbesondere, dass in einem Hauptidealring ein größter gemeinsamer Teiler zweier Elemente immer existiert. Außerdem erzeugen in diesem Fall Elemente $a$ und $b$ genau dann das Einsideal, wenn 1 größter gemeinsamer Teiler von $a$ und $b$ ist.

(4) Ist $R$ sogar euklidisch, dann kann man den größten gemeinsamen Teiler von $a$ und $b$ mit dem euklidischen Algorithmus (Bemerkung 15.43) berechnen.

Siehe auch Bemerkung 15.54. ◊

BEMERKUNG 15.43 (Der euklidische Algorithmus). Ist $R$ ein Hauptidealring und sind $a, b \in R$, so ist $(a, b)$ ein Hauptideal. In euklidischen Ringen kann man mit dem sogenannten *Euklidischen Algorithmus* recht leicht ein Element $d \in R$ berechnen, für das $(a, b) = (d)$ gilt. Wie in Bemerkung 15.42 erläutert, bedeutet das genau, dass $d$ ein ggT von $a$ und $b$ ist. Wir nehmen dazu an, dass $a, b \neq 0$ ist, denn sonst ist nichts zu tun.

Der Algorithmus besteht darin, induktiv eine Folge $a_0, a_1, a_2, \ldots$ von Elementen in $R$ wie folgt zu definieren bzw. zu berechnen:

$$a_0 := a, \quad a_1 := b$$

und für $i > 1$ definieren wir $a_i$ durch Division von $a_{i-2}$ durch $a_{i-1}$ mit Rest, d.h. wir schreiben

$$a_{i-2} = q_{i-1} a_{i-1} + a_i.$$

Der Algorithmus bricht ab, sobald $a_{k+1} = 0$ ist, das Ergebnis ist dann $d := a_k$, wie wir nachfolgend begründen werden. Weil für die Gradfunktion $\delta$ von $R$ gilt, dass

$$\delta(a_1) > \delta(a_2) > \delta(a_3) > \cdots$$

(solange $a_i \neq 0$ gilt), ist das nach endlich vielen Schritten der Fall.

Dann folgt aus $a_{i-2} = q_i a_{i-1} + a_i$, dass $(a_{i-1}, a_i) = (a_{i-2}, a_{i-1})$ gilt, und aus der letzten Gleichung $a_{k-1} = q_k a_k$ folgt $a_{k-1} \in (a_k)$, also

$$(a_k) = (a_{k-1}, a_k) = (a_{k-2}, a_{k-1}) = \cdots = (a, b),$$

wir haben also tatsächlich einen Erzeuger des Hauptideals $(a, b)$ gefunden.

Oft ist es nützlich, dass der Algorithmus auch eine Möglichkeit liefert, eine Darstellung der Form $a_k = xa + yb$ zu berechnen. Dazu betrachten wir die Gleichungskette

$$\begin{aligned} a_k &= a_{k-2} - q_{k-1}a_{k-1} \\ &= a_{k-2} - q_{k-1}(a_{k-3} - q_{k-2}a_{k-2}) \\ &= -q_{k-1}a_{k-3} + (1 + q_{k-1}q_{k-2})a_{k-2} \\ &= -q_{k-1}a_{k-3} + (1 + q_{k-1}q_{k-2})(a_{k-4} - q_{k-3}a_{k-3}) \\ &= \cdots, \end{aligned}$$

aus der wir die gewünschte Darstellung $a_k = xa_0 + ya_1 = xa + yb$ erhalten. ◊

**15.4.3. Faktorielle Ringe.** Wir wollen nun eine Klasse von Ringen definieren und untersuchen, in der ein Analogon der eindeutigen Primfaktorzerlegung gilt, die wir vom Ring der ganzen Zahlen kennen (Satz I.3.56).

Eine Primzahl ist eine natürliche Zahl $p > 1$, die sich nicht als Produkt $ab$ mit $a, b \in \mathbb{Z}$, $1 < a, b < p$ schreiben lässt. Um diesen Begriff auf beliebige Integritätsringe zu übertragen, ist es sinnvoll, die Einschränkung auf Zahlen $> 1$ fallenzulassen und auch Zahlen $< -1$ zu betrachten, die sich nicht in nichttrivialer Weise als Produkt schreiben lassen. Das Nullelement und die Einheiten $1, -1 \in \mathbb{Z}^\times$ spielen eine Sonderrolle. Der Begriff, den man so erhält, ist der des »irreduziblen Elements«, Definition 15.44 (1). Oft ist eine andere Eigenschaft von Primzahlen aber wichtiger, nämlich die sogenannte *Primeigenschaft*. Wenn eine Primzahl $p$ ein Produkt teilt, dann teilt sie auch einen der Faktoren:

$$p \mid ab \quad \Longrightarrow \quad p \mid a \text{ oder } p \mid b.$$

Siehe Satz I.3.52 für einen Beweis. Wir haben diese Eigenschaft von Primzahlen in Abschnitt I.4.2.1 benutzt, um zu zeigen, dass der Restklassenring $\mathbb{Z}/p$ für eine Primzahl $p$ ein Körper ist. Diese Eigenschaft ist die Grundlage von Teil (2) der folgenden Definition. In allgemeinen Integritätsringen müssen diese Eigenschaften nicht zusammenfallen!

Definition 15.44. Sei $R$ ein Integritätsring.

(1) Ein Element $p \in R \setminus (R^\times \cup \{0\})$ heißt *irreduzibel*, falls für alle $a, b \in R$ mit $p = ab$ gilt: $a \in R^\times$ oder $b \in R^\times$.

(2) Ein Element $p \in R \setminus (R^\times \cup \{0\})$ heißt *prim* (oder *Primelement*), falls für alle $a, b \in R$ mit $p \mid ab$ gilt: $p \mid a$ oder $p \mid b$.

⊣

Ist $R$ ein Integritätsring und sind $p, a, b \in R$ mit $p = ab \neq 0$, dann ist $a$ genau dann eine Einheit in $R$, wenn $p$ und $b$ assoziiert sind. Denn wenn $a$ eine Einheit ist, so folgt direkt aus der Definition, dass $p$ und $b$ assoziiert zueinander sind. Und wenn $p$ und $b$ assoziiert sind, sagen wir $p = ub$ mit $u \in R^\times$, so folgt $ub = ab$ und mit der Kürzungsregel, dass $a = u \in R^\times$ ist.

Wir könnten also Teil (1) der Definition auch so formulieren, dass $p \in R \setminus (R^\times \cup \{0\})$ genau dann irreduzibel ist, wenn in jeder Darstellung $p = ab$ einer der Faktoren zu $p$ assoziiert ist.

SATZ 15.45. *Sei $R$ ein Integritätsring. Ist $p \in R$ prim, so ist $p$ irreduzibel. Ist $R$ ein Hauptidealring, so gilt auch die Umkehrung.*

BEWEIS. Sei zunächst $p$ prim. Wenn sich $p$ als Produkt $p = ab$ schreiben lässt, so folgt aus der Primeigenschaft $p \mid a$ oder $p \mid b$. Nehmen wir ohne Einschränkung an, dass der erste Fall eintritt. Andererseits impliziert $p = ab$ auch, dass $a$ ein Teiler von $p$ ist. Wir haben also $a \mid p$ und $p \mid a$, und es folgt, dass $a$ und $p$ zueinander assoziiert sind. Wie oben bemerkt, zeigt das die Irreduzibilität von $p$.

Sei nun $R$ ein Hauptidealring und $p \in R$ irreduzibel. Wir wollen zeigen, dass $p$ prim ist. Seien also $a, b \in R$ mit $p \mid ab$. Nehmen wir an, dass $p \nmid a$ gilt, also $a \notin (p)$. Dann ist $(p) \subsetneq (a, p)$ eine echte Teilmenge. Hier ist $(a, p)$ das von $a$ und $p$ erzeugte Ideal, das wir folgendermaßen explizit beschreiben können:

$$(a, p) = \{xa + yp;\ x, y \in R\}.$$

In der Tat ist klar, dass hier $\supseteq$ gilt, da $a$ und $p$ in $(a, p)$ liegen und wegen der Idealeigenschaft folglich auch alle Ausdrücke der Form $xa + yp$. Andererseits ist leicht zu sehen, dass die rechte Seite ein Ideal ist, und weil $(a, p)$ das *kleinste* Ideal ist, das $a$ und $p$ enthält, folgt die Gleichheit.

Weil $R$ ein Hauptidealring ist, ist das Ideal $(a, p)$ ein Hauptideal, es gibt also ein Element $d \in R$ mit $(a, p) = (d)$. Es folgt dann $d \mid p$ und wegen der Irreduzibilität von $p$ und weil $(p) \neq (d)$ ist, dass $(d) = R$ sein muss. Damit erhalten wir $1 \in (d) = (a, p)$, also existieren $x, y \in R$ mit $ax + yp = 1$. Wir sehen jetzt, dass $p \mid (1 - ax)$, also erst recht $p \mid (b - abx)$, und wegen $p \mid ab$ folgt nun $p \mid b$. □

LEMMA 15.46. *Sei $R$ ein Hauptidealring, und seien*

$$\mathfrak{a}_0 \subseteq \mathfrak{a}_1 \subseteq \mathfrak{a}_2 \subseteq \cdots$$

*Ideale von $R$, die ineinander enthalten sind. Man spricht von einer* aufsteigenden Kette *von Idealen in $R$.*

*Dann existiert $i \geq 0$, so dass $\mathfrak{a}_i = \mathfrak{a}_j$ für alle $j \geq i$. Man sagt, die Kette sei* stationär.

BEWEIS. Sei $R$ ein Hauptidealring und sei

$$\mathfrak{a}_0 \subseteq \mathfrak{a}_1 \subseteq \mathfrak{a}_2 \subseteq \cdots$$

eine aufsteigende Kette von Idealen in $R$. Dann ist auch die Vereinigung $\mathfrak{a} := \bigcup_{i \geq 0} \mathfrak{a}_i$ ein Ideal. In der Tat, für $x, y \in \mathfrak{a}$ existieren $i$ und $j$ mit $x \in \mathfrak{a}_i, y \in \mathfrak{a}_j$. Sei ohne Einschränkung $i \leq j$, also $\mathfrak{a}_i \subseteq \mathfrak{a}_j$. Dann gilt $x + y \in \mathfrak{a}_j \subseteq \mathfrak{a}$. Außerdem gilt für alle $z \in R$, dass $zx \in \mathfrak{a}_i \subseteq \mathfrak{a}$ ist.

Weil $R$ ein Hauptidealring ist, existiert ein Element $a \in R$ mit $\mathfrak{a} = (a)$. Dann muss aber $a$ in einem der Ideale $\mathfrak{a}_i$ liegen, es folgt $\mathfrak{a} = (a) \subseteq \mathfrak{a}_i$ und damit die Gleichheit $\mathfrak{a} = \mathfrak{a}_i$ und insbesondere $\mathfrak{a}_i = \mathfrak{a}_j$ für alle $j \geq i$. □

Ringe, die die Eigenschaft aus dem Lemma haben, in denen also jede aufsteigende Kette von Idealen stationär ist, heißen auch *noethersche* Ringe (nach der Mathematikerin Emmy Noether[4]).

Für $R = \mathbb{Z}$ bzw. $R = K[X]$ ($K$ ein Körper) kann man das Lemma noch einfacher beweisen, indem man den Absolutbetrag bzw. die Gradfunktion benutzt.

SATZ 15.47. *Sei $R$ ein Hauptidealring. Dann lässt sich jedes Element aus $R \setminus (R^\times \cup \{0\})$ als Produkt von Primelementen schreiben.*

[4] `https://de.wikipedia.org/wiki/Emmy_Noether`

BEWEIS. Wegen Satz 15.45 ist es äquivalent zu zeigen, dass sich jedes Element als Produkt von irreduziblen Elementen schreiben lässt. Angenommen, das wäre nicht der Fall, sei also $a_0 \in R \setminus (R^\times \cup \{0\})$ ein Element, das sich *nicht* als Produkt von irreduziblen Elementen schreiben lässt. Insbesondere kann dann $a_0$ nicht irreduzibel sein, es existiert also eine Produktdarstellung $a_0 = a_1 b_1$ mit Nicht-Einheiten $a_1, b_1$. Wenn diese Elemente beide als Produkt irreduzibler Elemente geschrieben werden könnten, dann bekämen wir auch eine entsprechende Darstellung für $a_0$. Das ist nicht möglich, wir können also (indem wir nötigenfalls $a_1$ und $b_1$ vertauschen) annehmen, dass auch $a_1$ sich nicht als Produkt von irreduziblen Elementen schreiben lässt.

Wenn wir in dieser Weise fortfahren, erhalten wir eine Folge von Elementen

$$a_i = a_{i+1} b_{i+1}, \quad i = 0, 1, 2, \ldots,$$

von $R$, die sämtlich keine Einheiten sind. In Termen von Idealen folgt, dass $(a_i) \subseteq (a_{i+1})$ für alle $i \geq 0$ gilt, wir erhalten also eine aufsteigende Kette

$$(a_0) \subseteq (a_1) \subseteq (a_2) \subseteq \cdots$$

von Idealen in $R$, die nach Lemma 15.46 stationär wird, es gibt also ein $i$ mit $(a_i) = (a_{i+1})$. Das impliziert aber, dass $b_{i+1}$ im Widerspruch zu unserer Konstruktion doch eine Einheit in $R$ ist. □

LEMMA 15.48. *Sei $R$ ein Integritätsring, seien $p_1, \ldots, p_r \in R$ prim und seien $q_1, \ldots, q_s \in R$ irreduzibel. Gilt*

$$p_1 \cdot \cdots \cdot p_r = q_1 \cdot \cdots \cdot q_s,$$

*so gilt $r = s$ und nach einer eventuellen Umnummerierung der $q_i$ gilt für alle $i = 1, \ldots, r$, dass $p_i$ und $q_i$ zueinander assoziiert sind.*

BEWEIS. Sei $p \in R$ ein Primelement, d.h. aus $p \mid ab$ folgt $p \mid a$ oder $p \mid b$ (für alle $a, b \in R$). Per Induktion folgt dann aus $p \mid a_1 \cdot \cdots \cdot a_n$ für Elemente $a_i \in R$, dass $p$ einen der Faktoren des Produkts $a_1 \cdot \cdots \cdot a_n$ teilt: Es existiert $i$ mit $p \mid a_i$.

Wir beweisen nun eine etwas allgemeinere Aussage als die des Lemmas, nämlich: Seien $u \in R^\times$, seien $p_1, \ldots, p_r \in R$ prim und seien $q_1, \ldots, q_s \in R$ irreduzibel. Gilt

$$p_1 \cdot \cdots \cdot p_r = u q_1 \cdot \cdots \cdot q_s,$$

so gilt $r = s$ und nach einer eventuellen Umnummerierung der $q_i$ gilt für alle $i = 1, \ldots, r$, dass $p_i$ und $q_i$ zueinander assoziiert sind.

Die Aussage des Lemmas folgt, indem wir $u = 1$ setzen.

Wir führen Induktion nach $r$. Der Fall $r = 0$, indem links das leere Produkt 1 steht, ist trivial, da irreduzible Elemente per Definition keine Einheiten sein können.

Für $r \geq 1$ gilt $p_1 \mid p_1 \cdot \cdots \cdot p_r = u q_1 \cdot \cdots \cdot q_s$, dass $p_1$ eines der $q_i$ teilt. (Dass $p \mid u$, ist unmöglich, da $u$ eine Einheit und $p_1$ keine Einheit ist.) Nach Umnummerierung der $q_i$ können wir annehmen, dass $p_1 \mid q_1$, etwa $q_1 = \varepsilon p_1$. Weil $q_1$ irreduzibel und $p_1$ als Primelement keine Einheit ist, folgt daraus, dass $\varepsilon \in R^\times$ und sodann, dass $q_1$ und $p_1$ zueinander assoziiert sind.

Es folgt auch (siehe Lemma 15.32), dass

$$p_2 \cdot \cdots \cdot p_r = (u\varepsilon^{-1}) q_2 \cdot \cdots \cdot q_s,$$

und per Induktionsvoraussetzung folgt die Behauptung.

Vergleiche auch den Beweis der Eindeutigkeit der Primfaktorzerlegung in $\mathbb{Z}$ in Satz I.3.56. □

Da Primelemente stets irreduzibel sind (Satz 15.45), zeigt Lemma 15.48, dass eine Zerlegung als ein Produkt in Primelemente immer bis auf Reihenfolge und Übergang zu assoziierten Elementen eindeutig ist.

DEFINITION 15.49. Ein Integritätsring $R$ heißt *faktoriell*, wenn sich jedes Element aus $R \setminus (R^\times \cup \{0\})$ als Produkt von Primelementen schreiben lässt. ⊣

Man sagt in der Situation dieser Definition auch, in $R$ gelte die »eindeutige Zerlegung in Primfaktoren«. Eine (etwas aus der Mode gekommene) alternative Bezeichnung für faktorielle Ringe ist *ZPE-Ringe* – das steht für »Zerlegung in Primelemente eindeutig«. Auf Englisch werden faktorielle Ringe oft als »UFD« bezeichnet, das ist die Abkürzung für »unique factorization domain«. Wir können Satz 15.47 nun wie folgt formulieren.

KOROLLAR 15.50. *Jeder Hauptidealring ist faktoriell.*

SATZ 15.51. *Sei $R$ ein Integritätsring. Dann sind äquivalent:*

(i) *Der Ring $R$ ist faktoriell.*

(ii) *Jedes Element aus $R \setminus (R^\times \cup \{0\})$ lässt sich als Produkt von irreduziblen Elementen schreiben, und jedes irreduzible Element von $R$ ist prim.*

BEWEIS. Es ist klar, dass (ii) $\Rightarrow$ (i) gilt. Für die Implikation (i) $\Rightarrow$ (ii) müssen wir zeigen, dass in einem faktoriellen Ring jedes irreduzible Element prim ist. Sei also $R$ faktoriell und $p \in R$ irreduzibel. Dann können wir $p$ als Produkt von Primelementen schreiben, etwa

$$p = p_1 \cdot \dots \cdot p_r.$$

Aus der Irreduzibilität folgt dann aber direkt, dass $r = 1$ und folglich $p = p_1$ ein Primelement sein muss. □

BEISPIEL 15.52. Da $\mathbb{Z}$ ein Hauptidealring ist, ist $\mathbb{Z}$ faktoriell. Wegen $\mathbb{Z}^\times = \{1, -1\}$ gilt auch die folgende, etwas präzisere Aussage: Jede ganze Zahl $a \in \mathbb{Z}$, $a \neq 0$, lässt sich schreiben als $a = \varepsilon p_1 \cdot \dots \cdot p_r$ mit $\varepsilon \in \{1, -1\}$ und (positiven) Primzahlen $p_i$. Dabei ist $\varepsilon$ eindeutig bestimmt (nämlich gleich dem Vorzeichen von $a$), und die $p_i$ sind eindeutig bestimmt bis auf die Reihenfolge. Siehe auch Satz I.3.56. ◊

Für die ganzen Zahlen kannten wir diese Aussage ja schon aus der Linearen Algebra 1. Im anderen wichtigen Beispiel für Hauptidealringe, das wir kennengelernt haben, ist sie hingegen neu, und wird in den kommenden beiden Kapitel eine wichtige Rolle spielen.

BEISPIEL 15.53. Sei $K$ ein Körper. Nach dem Gezeigten ist der Polynomring $R = K[X]$ faktoriell. Es gilt $R^\times = K^\times$ und wir erhalten: Jedes Polynom $f \in K[X], f \neq 0$, lässt sich schreiben als Produkt $f = uf_1 \cdot \dots \cdot f_r$, wobei $u \in K^\times, f_i \in K[X]$ irreduzibel und normiert.

Dabei ist $u$ eindeutig bestimmt ($u$ ist der Leitkoeffizient von $f$), und die $f_i$ sind eindeutig bestimmt bis auf ihre Reihenfolge. (Da die $f_i$ irreduzibel sind, gilt $\deg f_i > 0$ für alle $i$.) ◊

BEMERKUNG 15.54. Sei $R$ ein faktorieller Ring.

(1) Sei $P \subset R$ eine Menge von Primelementen mit der Eigenschaft, dass für jedes Primelement $q \in R$ genau ein $p \in P$ existiert, das zu $q$ assoziiert ist. Wir nennen dann $P$ ein Vertretersystem der Primelemente in $R$ bis auf Assoziiertheit. Wir können dann für ein Element $a \in R \setminus \{0\}$ die Primfaktorzerlegung in der Form

$$a = u \prod_{p \in P} p^{v_p(a)}$$

schreiben, wobei $u \in R^\times$ eine Einheit ist und $v_p(a) \in \mathbb{N}$ und $v_p(a) = 0$ für alle bis auf endlich viele $p \in P$ gilt (daher ist das Produkt ein endliches Produkt, wenn alle Faktoren, die $= 1$ sind, weggelassen werden, denn für $v_p(a) = 0$ ist $p^{v_p(a)} = p^0 = 1$). Ist $a$ eine

Einheit, so sind alle $v_p(a) = 0$, und umgekehrt. Bei dieser Schreibweise sind $u$ und alle Zahlen $v_p(a)$ eindeutig bestimmt.

Dann gilt $p^k \mid a$ genau dann, wenn $v_p(a) \geq k$ ist.

Im Fall $R = \mathbb{Z}$ wählt man als die Menge $P$ üblicherweise die Menge der (positiven) Primzahlen. Ist $R = K[X]$ der Polynomring über einem Körper, dann ist die übliche Wahl für $P$ die Menge der *normierten* primen Polynome. Man erhält dann genau die oben diskutierten Beispiele wieder.

(2) Seien nun $a, b \in R \setminus (R^\times \cup \{0\})$. Wir schreiben wie in Punkt (1) die Primfaktorzerlegungen als

$$a = u \prod_{p \in P} p^{v_p(a)}, \quad b = u' \prod_{p \in P} p^{v_p(b)}.$$

Es gilt $a \mid b$ genau dann, wenn $v_p(a) \leq v_p(b)$ für alle $p \in P$ gilt.

(3) Mit der Notation aus Punkt (2) ist

$$\prod_{p \in P} p^{\min(v_p(a), v_p(b))}$$

ein größter gemeinsamer Teiler von $a$ und $b$ in $R$, und

$$\prod_{p \in P} p^{\max(v_p(a), v_p(b))}$$

ein kleinstes gemeinsames Vielfaches von $a$ und $b$ in $R$ (Definition 15.41). Durch die Wahl von $P$ erhält man in dieser Art und Weise einen ausgezeichneten größten gemeinsamen Teiler und ein ausgezeichnetes kleinstes gemeinsames Vielfaches von $a$ und $b$. Jeder andere größte gemeinsame Teiler (bzw. jedes andere kleinste gemeinsame Vielfache) im Sinne von Definition 15.41 ist, wie in jedem Integritätsring, zu den oben genannten ggT/kgV assoziiert.

Insbesondere existieren ggT und kgV in faktoriellen Ringen immer. Allerdings folgt aus $\mathrm{ggT}(a, b) = 1$ nicht in jedem faktoriellen Ring, dass Elemente $x, y$ existieren mit $xa + yb = 1$ – in Hauptidealringen ist das aber richtig (Bemerkung 15.42), und nur in diesen »funktionieren« die Begriffe ggT und kgV wirklich gut.

◊

> Du wolltest doch Algebra, da hast du den Salat.
>
> Jules Verne, Reise um den Mond, 4. Kapitel
>
> Fundort: http://www.mathe.tu-freiberg.de/~hebisch/cafe/zitate.html

ERGÄNZUNG 15.55. Wir skizzieren zwei Beispiele von Integritätsringen, die nicht faktoriell sind.

(1) Die Teilmenge

$$\mathbb{Z}[i\sqrt{5}] := \{a + ib\sqrt{5};\ a, b \in \mathbb{Z}\} \quad \subseteq \mathbb{C}$$

ist ein Unterring. Man kann zeigen, dass dieser Integritätsring nicht faktoriell ist. Das Element 2 ist in diesem Ring irreduzibel, jedoch kein Primelement, denn es teilt das Produkt

$$(1 - i\sqrt{5})(1 + i\sqrt{5}) = 6 = 2 \cdot 3,$$

aber teilt weder $1 - i\sqrt{5}$ noch $1 + i\sqrt{5}$.

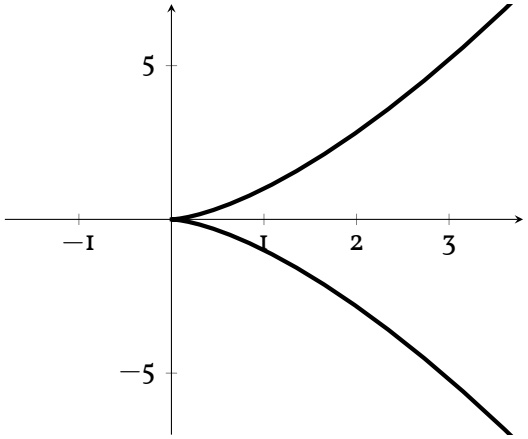
Dieser und ähnliche Ringe werden in der algebraischen Zahlentheorie genauer untersucht. Die Theorie auf den nicht-faktoriellen Fall auszudehnen ist dort sehr wichtig, und war der Ausgangspunkt dafür, den Begriff des Ideals einzuführen (siehe Ergänzung 15.20). Man kann zeigen, dass die Ideale im Ring $\mathbb{Z}[i\sqrt{5}]$ eine eindeutige »Zerlegung« in sogenannte Primideale (vgl. Ergänzung 15.75) zulassen, und dies ist oft ein guter Ersatz für die Zerlegung von Elementen des Rings als Produkt von Primelementen, die in diesem Ring eben nicht immer möglich ist.

(2) Sei $K$ ein Körper. Die Teilmenge

$$K[T^2, T^3] := \left\{ \sum_{i=0}^{n} a_i T^i;\ n \in \mathbb{N},\ a_i \in K,\ a_1 = 0 \right\} \subseteq K[T]$$

ist ein Unterring. Dieser Ring ist ein weiteres Beispiel eines Integritätsrings, der nicht faktoriell ist, denn $T^6 = (T^2)^3 = (T^3)^2$ hat zwei verschiedene Zerlegungen in irreduzible Elemente.

In der algebraischen Geometrie wird dieser Ring »in geometrischer Weise« interpretiert. Man kann eine Verbindung herstellen zu der hier abgebildeten »Kurve« in der Ebene (die Abbildung entspricht dem Fall $K = \mathbb{R}$), und dann in präziser Weise begründen, dass die Eigenschaft des obigen Rings, nicht faktoriell zu sein, damit zusammenhängt, dass die abgebildete Kurve am Ursprung nicht »glatt« ist, also an diesem Punkt auch »nach beliebig starkem Hereinzoomen« nicht wie eine Gerade aussieht.




Die Menge $\{(x, y)^t \in \mathbb{R}^2;\ y^2 = x^3\}$

Der Zusammenhang zwischen dem Ring $K[T^2, T^3]$ und der Gleichung $y^2 - x^3 = 0$ kommt daher, dass die Abbildung $K[X, Y] \to K[T], X \mapsto T^3, Y \mapsto T^2$, ein Ringhomomorphismus mit Bild $K[T^2, T^3]$ und Kern $(Y^2 - X^3)$ ist.
(Es ist in Ordnung, wenn Sie diese ganze Bemerkung etwas kryptisch finden ...)

□ Ergänzung 15.55

### 15.4.4. Nullstellen von Polynomen.

Sei $R$ ein Ring.

DEFINITION 15.56. Sei $f \in R[X]$. Ein Element $\alpha \in R$ heißt *Nullstelle* von $f$, falls $f(\alpha) = 0$. ⊣

Sei nun $R$ ein Integritätsring. Wir haben gesehen, dass dann auch $R[X]$ ein Integritätsring ist (Korollar 15.31).

LEMMA 15.57. *Ein Element $\alpha \in R$ ist genau dann Nullstelle eines Polynoms $f \in R[X]$, wenn $X - \alpha$ das Polynom $f$ teilt.*

BEWEIS. Wenn $f$ ein Vielfaches von $X-\alpha$ ist, dann ist natürlich $f(\alpha) = 0$. Ist andererseits $\alpha$ eine Nullstelle von $f$ und schreiben wir $f$ im Sinne der Division mit Rest als

$$f = q \cdot (X - \alpha) + r$$

mit $\deg(r) < 1$, dann ergibt Einsetzen von $\alpha$, dass $r(\alpha) = f(\alpha) = 0$. Weil $r$ ein konstantes Polynom ist, folgt $r = 0$, also $f = q \cdot (X - \alpha)$. □

Insbesondere sehen wir, dass ein Polynom vom Grad $n$ höchstens $n$ verschiedene Nullstellen haben kann (siehe auch Satz I.4.25).

Ein Polynom vom Grad 1 nennen wir auch ein *lineares Polynom*. Ein lineares Polynom, das $f$ teilt, nennen wir einen *Linearfaktor* von $f$. Ist $R = K$ ein Körper, so ist jedes lineare Polynom vom Grad 1 zu einem eindeutig bestimmten Polynom der Form $X - a$, $a \in K$ assoziiert. Über beliebigen Ringen ist diese Aussage natürlich nicht richtig; es kann dann auch lineare Polynome geben, die keine Nullstellen in dem Ring haben, zum Beispiel $R = \mathbb{Z}$ und $f = 2X - 1 \in \mathbb{Z}[X]$.

DEFINITION 15.58. Sei $R$ ein Integritätsring, $f \in R[X]$.

(1) Ist $\alpha \in R$, so gibt es eine eindeutig bestimmte natürliche Zahl $m \in \mathbb{N}$, so dass $(X-\alpha)^m \mid f$, aber $(X-\alpha)^{m+1} \nmid f$. Wir schreiben $\mathrm{mult}_\alpha(f) := m$. Das Element $\alpha$ ist genau dann eine Nullstelle von $f$, wenn $m \geq 1$. Wir sagen dann, $\alpha$ sei eine Nullstelle der *Vielfachheit* (oder: *Multiplizität*) $m$.

   Eine Nullstelle mit Vielfachheit 1 nennen wir auch *einfache Nullstelle*, eine mit Vielfachheit 2 entsprechend *doppelte Nullstelle* usw.

(2) Wir sagen, ein Polynom $f \in R[X] \setminus \{0\}$ *zerfalle vollständig in Linearfaktoren*, wenn $f$ Produkt von linearen Polynomen, d.h. von Polynomen vom Grad 1 ist.

⊣

DEFINITION 15.59. Ein Körper $K$ heißt *algebraisch abgeschlossen*, wenn jedes nichtkonstante Polynom, also jedes Polynom in $K[X] \setminus K$, eine Nullstelle in $K$ besitzt. ⊣

Per Induktion zeigt man, dass man algebraisch abgeschlossene Körper äquivalent dadurch charakterisieren kann, dass jedes nichkonstante Polynom vollständig in Linearfaktoren zerfällt.

Weder der Körper $\mathbb{Q}$ noch der Körper $\mathbb{R}$ sind algebraisch abgeschlossen (überlegen Sie sich Beispiele von nichtkonstanten Polynomen, die keine Nullstelle haben). Auch ein endlicher Körper kann nicht algebraisch abgeschlossen sein (warum?). Es ist auch gar nicht so einfach, Beispiele von algebraisch abgeschlossenen Körpern anzugeben. Das zugänglichste Beispiel ist der Körper $\mathbb{C}$.

THEOREM 15.60 (Fundamentalsatz der Algebra). *Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen.*

Dieses schwierige Theorem beweisen wir nicht im Rahmen der Vorlesung über lineare Algebra. Es wird üblicherweise auf verschiedene Arten in den Vorlesungen *Algebra* und *Funktionentheorie* bewiesen, kann aber auch mit Mitteln der Analysis I bewiesen werden. Siehe Ergänzung 16.32 für einen trickreichen Beweis, der nur sehr wenig Analysis benötigt und ansonsten mit linearer Algebra auskommt.

**15.4.5. Der chinesische Restsatz.** Sei $R$ ein Ring, $\mathfrak{a} \subset R$ ein Ideal. Für Elemente $x, y \in R$ schreiben wir

$$x \equiv y \mod \mathfrak{a}, \text{ wenn } x - y \in \mathfrak{a}.$$

In den meisten Fällen, die für uns relevant sind, ist $\mathfrak{a} = (a)$ ein Hauptideal; dann schreiben wir auch $x \equiv y \mod a$, und dies ist gerade äquivalent zu $a \mid x - y$. Man sagt, $x$ sei *kongruent* zu $y$ *modulo* $\mathfrak{a}$. Kongruenz ist eine »Äquivalenzrelation« (siehe Definition 15.64 unten).

Im folgenden Satz betrachten wir für Elemente $a, b$ eines (Integritäts-)Rings $R$ das von $a$ und $b$ erzeugte Ideal

$$(a, b) = \{xa + yb;\ x, y \in R\}$$

und betrachten die Bedingung, dass dieses gleich $R$ ist. Weil ein Ideal $\mathfrak{a}$ genau dann gleich dem ganzen Ring ist, wenn es die 1 enthält, ist das gewissermaßen eine abkürzende Schreibweise dafür, dass $x, y \in R$ existieren mit $xa + yb = 1$. Ist $R$ ein Hauptidealring (und das ist der Fall, der für uns später relevant sein wird), ist die Bedingung dazu äquivalent, dass 1 ein größter gemeinsamer Teiler von $a$ und $b$ ist (Bemerkung 15.42).

Satz 15.61 (Chinesischer Restsatz). *Seien $R$ ein Integritätsring und $a_1, \ldots, a_r \in R$, so dass $(a_i, a_j) = R$ für alle $i \neq j$. Sei $a = a_1 \cdot \cdots \cdot a_r$.*

*Seien $b_1, \ldots, b_r \in R$. Dann existiert ein Element $b \in R$, so dass*

$$b \equiv b_i \mod a_i \quad \textit{für alle } i = 1, \ldots, r$$

*gilt.*

*Ist $b'$ ein weiteres solches Element, so gilt $b \equiv b' \mod a$. (Wir sagen, die Lösung der vorgegebenen Kongruenzen sei* eindeutig bestimmt modulo $a$.*)*

Beweis. *Vorüberlegung.* Wir zeigen zuerst, dass unter der Voraussetzung, dass für alle $i \neq j$ die Elemente $a_i$ und $a_j$ das Einsideal erzeugen, auch für alle $i$ die Elemente $a_i$ und $a_i' := \prod_{j \neq i} a_j$ das Einsideal erzeugen. Sei zur Vereinfachung der Notation ohne Einschränkung $i = 1$. Jedenfalls existieren $x_j, y_j \in R, j = 2, \ldots, n$, so dass $x_j a_1 + y_j a_j = 1$. Daraus erhalten wir

$$\prod_{j=2}^{n} (x_j a_1 + y_j a_j) = 1,$$

und wenn wir den Ausdruck auf der linken Seite ausmultiplizieren, sind alle Summanden Vielfache von $a_1$, bis auf den Term $\prod_{j=2}^{n} y_j a_j$. Wir erhalten also tatsächlich einen Ausdruck der Form

$$xa_1 + y(a_2 \cdot \cdots \cdot a_n)$$

(mit $y = x_2 \cdot \cdots \cdot y_n$).

Nach dieser Vorüberlegung können wir für jedes $i \in \{1, \ldots, n\}$ Elemente $x_i, y_i \in R$ finden, so dass

$$x_i a_i + y_i a_i' = 1,$$

also $y_i a_i' \equiv 1 \mod a_i$. Nach Definition der $a_i'$ ist auch klar, dass $y_i a_i' \equiv 0 \mod a_j$ für alle $j \neq i$ gilt. Wir setzen nun

$$b = \sum_{i=1}^{n} b_i y_i a_i'.$$

In der Tat gilt dann für jedes $i$, dass

$$b \equiv b_i y_i a_i' \equiv b_i \mod a_i,$$

wie gewünscht. Damit ist die Existenzaussage bewiesen.

Seien nun $b, b' \in R$ mit $b \equiv b_i \mod a_i$ und $b' \equiv b_i \mod a_i$ für alle $i$. Es folgt $b - b' \in (a_i)$ für alle $i$, also $b - b' \in \bigcap_{i=1}^{n} (a_i)$. Es genügt also zu zeigen, dass $\bigcap_{i=1}^{n} (a_i) = (a)$ gilt (wobei die

Inklusion $\supseteq$ klar ist; allerdings ist das auch die Inklusion, die uns hier nicht interessiert). Mithilfe der Vorüberlegung können wir das per Induktion beweisen und uns damit auf den Fall $n = 2$ zurückziehen. Dann haben wir Elemente $a_1, a_2 \in R$ gegeben, die das Einsideal erzeugen, etwa $x_1a_1 + x_2a_2 = 1$, und wollen für $c \in (a_1) \cap (a_2)$ zeigen, dass $c \in (a_1a_2)$ gilt. Wir können $c = y_1a_1 = y_2a_2$ und damit

$$x_2c = x_2y_1a_1 = x_2y_2a_2 = y_2(1 - x_1a_1)$$

schreiben, also $y_2 = x_2y_1a_1 + y_2x_1a_1 \in (a_1)$. Es folgt, dass $c = y_2a_2$ ein Vielfaches von $a_1a_2$ ist, wie wir zeigen wollten. □

Man kann den Satz noch etwas allgemeiner fassen und mit fast demselben Beweis abhandeln, siehe Ergänzung 18.37.

Beispiel 15.62. Wir betrachten das folgende Beispiel. Sei $R = \mathbb{Z}$ der Ring der ganzen Zahlen. Wir wollen eine ganze Zahl $x$ finden, so dass

$$\begin{aligned} x &\equiv 3 \mod 5, \\ x &\equiv 2 \mod 6, \\ x &\equiv 1 \mod 11. \end{aligned}$$

Es folgt aus dem chinesischen Restsatz, dass solche Zahlen existieren, und dass je zwei Lösungen modulo $5 \cdot 6 \cdot 11 = 330$ kongruent sind.

Um ein $x$ zu finden, können wir die Schritte aus dem allgemeinen Beweis nachvollziehen. Wir schreiben zuerst die 1 als »Linearkombination« einer der Zahlen 5, 6, 11 und dem Produkt der anderen Zahlen, d.h.

$$\begin{aligned} 1 &= 5x_1 + 66y_1 \\ 1 &= 6x_2 + 55y_2 \\ 1 &= 11x_3 + 30y_3. \end{aligned}$$

Diese Darstellungen lassen sich mit dem euklidischen Algorithmus (Bemerkung 15.43) finden. Im konkreten Fall gilt zum Beispiel

$$\begin{aligned} 1 &= 5 \cdot (-13) + 66 \cdot 1 \\ 1 &= 6 \cdot (-9) + 55 \cdot 1 \\ 1 &= 11 \cdot 11 + 30 \cdot (-4). \end{aligned}$$

Dann können wir

$$x = 3 \cdot 66 \cdot 1 + 2 \cdot 55 \cdot 1 + 1 \cdot 30 \cdot (-4) = 188$$

setzen. Hier ist in jedem Summanden der erste Faktor die rechte Seite der Kongruenz die $x$ erfüllen soll, und dann kommt das Produkt aus der obigen Darstellung der 1. In der Tat hat 188 bei Division durch 5 den Rest 3, bei Division durch 6 den Rest 2 und bei Division durch 11 den Rest 1. ◊

Ergänzung 15.63. Die Aussage des chinesischen Restsatzes findet man bereits in dem Buch »Sun Zi Suanjing« ds chinesischen Mathematikers Sun Zi[5] (um 3. Jh.) – daher der Name.
□ Ergänzung 15.63

[5] https://de.wikipedia.org/wiki/Sun_Zi_(Mathematiker)

## 15.5. Der Quotientenkörper eines Integritätsrings

Wir wollen in diesem Abschnitt zu einem Integritätsring $R$ einen Körper $K$ konstruieren, der $R$ als Unterring enthält. Unser Modell dafür ist der Fall der ganzen Zahlen $\mathbb{Z}$, die als Unterring im Körper $\mathbb{Q}$ der rationalen Zahlen enthalten sind. Im allgemeinen Fall imitieren wir die Konstruktion der Bruchzahlen aus ganzen Zahlen.

Ein unmittelbarer Nutzen dieser Konstruktion wird für uns sein, dass wir den Begriff der Determinante auch für Matrizen über (Integritäts-)Ringen einführen können (Abschnitt 15.6) und einige der Ergebnisse der Theorie über Körpern auf den Fall von Ringen übertragen können. Im weiteren Verlauf der Vorlesung werden wir dann Determinanten von Matrizen benutzen, deren Einträge in einem Polynomring liegen, um das »charakteristische Polynom« einer Matrix zu definieren (Kapitel 16).

Wir beginnen damit, den Begriff der Äquivalenzrelation einzuführen, der in dieser Vorlesung noch an mehreren Stellen eine Rolle spielen wird. Siehe auch Abschnitt I.3.14.2, Definition I.3.67, wo dieser Begriff schon im Rahmen der Ergänzungen vorgestellt wurde.

DEFINITION 15.64. Sei $M$ eine Menge.

(1) Eine *Relation* auf $M$ ist eine Teilmenge $\mathcal{R} \subseteq M \times M$. (Elemente $x, y \in M$ »stehen in der gegebenen Relation zueinander«, wenn $(x, y) \in \mathcal{R}$ gilt.)

(2) Eine Relation $\mathcal{R}$ auf $M$ heißt *Äquivalenzrelation*, wenn gilt
   (a) (Reflexivität) Für alle $x \in M$ ist $(x, x) \in \mathcal{R}$.
   (b) (Symmetrie) Für alle $x, y \in M$ ist $(x, y) \in \mathcal{R}$ genau dann, wenn $(y, x) \in \mathcal{R}$.
   (c) (Transitivität) Für alle $x, y, z \in M$ mit $(x, y) \in \mathcal{R}$, $(y, z) \in \mathcal{R}$ gilt $(x, z) \in \mathcal{R}$.

⊣

Äquivalenzrelationen bezeichnet man oft mit dem Symbol $\sim$, d.h. man schreibt dann $x \sim y$ statt $(x, y) \in \mathcal{R}$. Aber auch die Symbole $=, \neq, \equiv, <, \leq, \mid$ bezeichnen Relationen. Welche davon sind Äquivalenzrelationen?

DEFINITION 15.65. Sei $\sim$ eine Äquivalenzrelation auf $M$. Die Teilmengen von $M$ der Form $[m] := \{m' \in M;\ m' \sim m\}$ für ein $m \in M$ heißen die *Äquivalenzklassen* bezüglich $\mathcal{R}$.

Die Menge aller Äquivalenzklassen bezeichnen wir mit $M/\sim$. ⊣

Zwei Äquivalenzklassen in $M$ sind entweder disjunkt oder gleich. (Warum?)

BEISPIEL 15.66. Beispiele für Äquivalenzrelationen.

(1) Sei $X$ eine Menge. Die Gleichheit von Elementen auf $X$ definiert eine Äquivalenzrelation. Jede Äquivalenzklasse besteht aus genau einem Element von $X$.

(2) Sei $R$ ein Integritätsring. Die Relation, dass zwei Elemente aus $R$ zueinander assoziiert sind (Definition 15.33), ist eine Äquivalenzrelation. Siehe auch Bemerkung 15.54. Dort wird – mit der nun neu eingeführten Terminologie – aus jeder der Äquivalenzklassen bezüglich dieser Äquivalenzrelation genau ein Element ausgewählt. Man spricht auch von einem *Vertretersystem* der Äquivalenzklassen.

(3) Sei $n > 0$ eine natürliche Zahl. Kongruenz modulo $n$ ist eine Äquivalenzrelation. Die Menge der Äquivalenzklassen ist die zugrundeliegende Menge des Restklassenrings $\mathbb{Z}/n$. Siehe Beispiel I.3.70 und Beispiel I.3.73.

◊

Überlegen Sie sich auch Beispiele für Relationen auf einer Menge $X$ (also Teilmengen von $X \times X$), die keine Äquivalenzrelationen sind. Können Sie jeweils ein Beispiel finden, das genau eine der drei Bedingungen reflexiv, symmetrisch, transitiv nicht erfüllt?

Sei $R$ ein Integritätsring, und $M = R \times (R \setminus \{0\})$. Wenn Sie Schwiergkeiten haben, der folgenden Diskussion zu folgen, dann sollten Sie zuerst alles im speziellen Fall $R = \mathbb{Z}$ durchgehen und dabei im Hinterkopf behalten, dass das Ziel ist, den Körper $\mathbb{Q}$ zu konstruieren.

Wir betrachten die folgende Äquivalenzrelation auf $M$:

$$(a,b) \sim (c,d) \quad \Leftrightarrow \quad ad = bc.$$

Siehe auch Beispiel I.3.72.

Es ist nicht schwer zu überprüfen, dass es sich hier tatsächlich um eine Äquivalenzrelation handelt. Reflexivität und Symmetrie sind offensichtlich. Für die Transitivität seien Paare mit $(a,b) \sim (c,d)$ und $(c,d) \sim (e,f)$ gegeben. Es folgt

$$adf = bcf = bde, \quad \text{also } d(af - be) = 0$$

und weil $d \neq 0$ und $R$ ein $R$ ein Integritätsring ist, dass $af - be = 0$. Das bedeutet genau, dass $(a,b) \sim (e,f)$ gilt.

Satz 15.67. *Sei $K := M/\sim$ die Menge der Äquivalenzklassen. Wir schreiben $\frac{a}{b}$ für die Äquivalenzklasse eines Elementes $(a,b) \in M$. Es gilt dann also*

$$\frac{a}{b} = \frac{c}{d} \quad \Leftrightarrow \quad ad = bc.$$

*Dann ist $K$ mit der Addition*

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$

*und der Multiplikation*

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$$

*ein Körper, der sogenannte* Quotientenkörper *von $R$, den wir auch mit* $\mathrm{Quot}(R)$ *bezeichnen.*

*Die Abbildung $R \to K, a \mapsto \frac{a}{1}$ ist ein injektiver Ringhomomorphismus. Man schreibt oft $a$ statt $\frac{a}{1}$ und fasst $R$ als Teilmenge von $K$ auf.*

Eine andere gebräuchliche Bezeichnung für den Quotientenkörper eines Integritätsrings $R$ ist $\mathrm{Frac}(R)$ (als Abkürzung für die englische Bezeichnung »field of fractions«).

Beweis. Zunächst ist nachzuprüfen, dass die angegebenen Vorschriften überhaupt Abbildungen definieren, dass sie also *wohldefiniert* sind. Denn wir haben dabei jeweils Repräsentanten der Äquivalenzklassen benutzt, und müssen begründen, dass eine andere Wahl von Repräsentanten derselben Äquivalenzklassen dasselbe Ergebnis liefern.

Seien also $\frac{a}{b} = \frac{a'}{b'}$ und $\frac{c}{d} = \frac{c'}{d'}$. Dann gilt $ab' = a'b$ und $cd' = c'd$ und daher

$$\frac{ad + bc}{bd} = \frac{adb'd' + bcb'd'}{bdb'd'} = \frac{a'd' + b'c'}{b'd'}$$

und

$$\frac{ac}{bd} = \frac{a'c'}{b'd'}$$

Wir erhalten also tatsächlich Abbildungen $+$ und $\cdot$ von $K \times K$ nach $K$.

Die Körperaxiome sind leicht nachzurechnen, die Rechnungen laufen genauso ab, wie man die Körperaxiome für den Körper $\mathbb{Q}$ aus den entsprechenden Rechenregeln für ganze Zahlen beweisen würde. Wir behandeln daher nur beispielhaft einige der Axiome.

Für das Assoziativgesetz der Addition rechnen wir

$$\left(\frac{a}{b}+\frac{c}{d}\right)+\frac{e}{f}=\frac{ad+bc}{bd}+\frac{e}{f}=\frac{(ad+bc)f+bde}{bdf}=\frac{adf+bcf+bde}{bdf}=\frac{a}{b}+\left(\frac{c}{d}+\frac{e}{f}\right).$$

Das neutrale Element der Addition ist $\frac{0}{1}$, das Negative von $\frac{a}{b}$ ist $\frac{-a}{b}$, denn

$$\frac{a}{b}+\frac{-a}{b}=\frac{ab-ba}{b^2}=\frac{0}{1}.$$

Das Assoziativgesetz der Multiplikation ist leicht einzusehen. Das neutrale Element der Multiplikation ist $\frac{1}{1}$. Ein Element $\frac{a}{b}$ mit $a \in R, b \in R \setminus \{0\}$ ist genau dann gleich dem Nullelement $\frac{0}{1}$, wenn $a = 0$ ist. Für $a, b \in R \setminus \{0\}$ ist $\frac{b}{a}$ das multiplikative Inverse von $\frac{a}{b}$. Das Distributivgesetz zu überprüfen, lassen wir als Übungsaufgabe.

Es bleibt nun noch, die Abbildung $\iota\colon R \to K, a \mapsto \frac{a}{1}$ anzuschauen. Weil

$$\iota(a+b)=\frac{a+b}{1}=\frac{a\cdot 1+1\cdot b}{1}=\frac{a}{1}+\frac{b}{1}=\iota(a)+\iota(b)$$

und

$$\iota(ab)=\frac{ab}{1}=\iota(a)\iota(b)$$

und offensichtlich $\iota(1)=\frac{1}{1}=1_K$ gilt, handelt es sich um einen Ringhomomorphismus. Gilt $\frac{a}{1}=\frac{b}{1}$, so folgt $a\cdot 1=1\cdot b$, also $a=b$, mithin ist $\iota$ injektiv. □

Der Satz zeigt, dass für jeden Integritätsring $R$ ein injektiver Ringhomomorphismus von $R$ in einen Körper existiert. Ist $R$ ein Ring, der kein Integritätsring ist, kann es einen injektiven Ringhomomorphismus von $R$ in einen Körper offenbar nicht geben.

Die zu Beginn des Beweises diskutierte Wohldefiniertheit ist eine konzeptionelle Schwierigkeit, die mit dem Begriff der Äquivalenzrelation verbunden ist. Machen Sie sich die Problematik daran bewusst, dass zum Beispiel die Vorschrift $\left(\frac{a}{b},\frac{c}{d}\right)\mapsto\frac{a+c}{1}$ für rationale Zahlen $\frac{a}{b},\frac{c}{d}\in\mathbb{Q}$ nicht wohldefiniert ist – sie definiert keine Abbildung $\mathbb{Q}\times\mathbb{Q}\to\mathbb{Q}$. Suchen Sie andere Beispiele von wohldefinierten/nicht wohldefinierten Zuordnungsvorschriften.

> Die ganzen Zahlen hat Gott gemacht, alles andere ist Menschenwerk.
>
> L. Kronecker

Beispiel 15.68. Der Quotientenkörper von $\mathbb{Z}$ ist der Körper $\mathbb{Q}$ der rationalen Zahlen. Hierzu ist nicht viel zu sagen, denn wir haben ja die allgemeine Konstruktion des Quotientenkörpers genau an die Regeln der üblichen Bruchrechnung angelehnt. ◊

Beispiel 15.69. Sei $K$ ein Körper. Der Polynomring $K[X]$ ist, wie wir in Korollar 15.31 gesehen haben, ein Integritätsring. Sein Quotientenkörper wird mit $K(X)$ bezeichnet und heißt der *Körper der rationalen Funktionen über K (in einer Unbestimmten).*

Seine Elemente sind Brüche der Form $\frac{f}{g}$, wobei $f$ und $g$ Polynome in $K[X]$ sind, und $g \neq 0$ gilt. Auch wenn $g$ nicht das Nullpolynom sein darf, kann $g$ natürlich Nullstellen in $K$ haben. Ein Element von $K(X)$ definiert daher im allgemeinen *nicht* durch Einsetzen von Elementen aus $K$ eine Abbildung $K \to K$. Die Nullstellen von $g$ sind sozusagen Polstellen, die man aus $K$ herausnehmen müsste, um den Definitionsbereich einer solchen Abbildung zu erhalten. ◊

BEMERKUNG 15.70. Sei $R$ ein faktorieller Ring und $K$ der Quotientenkörper von $R$. In Bemerkung 15.54 hatten wir die Primfaktorzerlegung eines Elements $a \in R \setminus \{0\}$ in der Form

$$a = u \prod_{p \in P} p^{v_p(a)}$$

geschrieben, wobei wir ein Vertretersystem $P$ der Primelemente in $R$ bis auf Assoziiertheit gewählt hatten, und die $v_p(a)$ natürliche Zahlen sind, von denen für gegebenes $a$ höchstens endlich viele von Null verschieden sind, und wo $u \in R^\times$ eine Einheit von $R$ ist.

Das können wir nun auf Elemente von $K^\times$ ausdehnen. Für $a \in K^\times$ erhalten wir eine (eindeutig bestimmte) Zerlegung

$$a = u \prod_{p \in P} p^{v_p(a)}$$

wo nun die $v_p(a) \in \mathbb{Z}$ ganze Zahlen sind (von denen wieder alle bis auf endlich viele verschwinden) und wieder $u \in R^\times$ ist. ◊

## 15.6. Determinanten über Ringen

Sei $R$ ein kommutativer Ring. Wir bezeichnen mit $M_{m\times n}(R)$ die Menge aller $m \times n$-Matrizen mit Einträgen in $R$, d.h.

$$M_{m\times n}(R) = \{(a_{ij})_{i=1,\dots,m,\, j=1,\dots,n};\ a_{ij} \in R\}.$$

Addition von Matrizen gleicher Größe, Multiplikation einer Matrix mit einem Skalar aus $R$ und das Produkt von Matrizen zueinander passender Größen definieren wir durch dieselben Formeln wie im Fall von Körpern. Es ist dann $M_{m\times n}(R)$ eine kommutative Gruppe bezüglich der Addition, es gilt das Assoziativgesetz für die Multiplikation (immer unter der Voraussetzung, dass alle Größen zueinander passen) und es gelten die Distributivgesetze. Diese Aussagen kann man mit denselben Rechnungen überprüfen, die wir in der Linearen Algebra 1 für Matrizen über einem Körper durchgeführt haben (Abschnitt I.5.3).

Wir schreiben wie gehabt $M_n(R) = M_{n\times n}(R)$ für die Menge der quadratischen Matrizen. Aus dem oben Gesagten folgt, dass es sich hierbei mit der Addition und Multiplikation von Matrizen um einen Ring handelt.

Die Leibniz-Formel ergibt über jedem Ring $R$ Sinn, und wir erhalten eine Abbildung

$$M_n(R) \to R, \quad A = (a_{ij})_{i,j} \mapsto \det(A) := \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}.$$

Wir nennen $\det(A)$ die Determinante der Matrix $A$.

Weil wir unten die folgende einfache Tatsache benötigen, halten wir sie als Lemma fest.

LEMMA 15.71. *Sei $n \in \mathbb{N}$ und $\varphi\colon R \to S$ ein Ringhomomorphismus. Indem wir $\varphi$ auf jeden Eintrag anwenden, erhalten wir einen Ringhomomorphismus $M_n(R) \to M_n(S)$, den wir ebenfalls mit $\varphi$ bezeichnen. Dann gilt für alle Matrizen $A \in M_n(R)$, dass*

$$\varphi(\det(A)) = \det(\varphi(A))$$

*ist.*

BEWEIS. Der Beweis ist (hoffentlich) nicht schwierig für Sie – überlegen Sie sich, warum die Aussage des Lemmas richtig ist! □

Man kann die Theorie der Determinante auch recht allgemein von Anfang an über Ringen entwickeln, aber man müsste dann an einigen Stellen vorsichtiger argumentieren, weil man über einem allgemeinen Ring beispielsweis nicht jede Matrix mit dem Gauß-Algorithmus auf Zeilenstufenform bringen kann. Das hatten wir aber im Kapitel über Determinanten in der Linearen Algebra I benutzt. Um nicht alles noch einmal durchgehen zu müssen, wählen wir eine etwas andere Strategie und führen die Ergebnisse, die wir benötigen, auf den Fall von Körpern zurück.

Sei dazu $R$ ein Integritätsring, $K$ sein Quotientenkörper. Wir können dann $M_n(R)$ als Teilmenge von $M_n(K)$ betrachten. Für $A \in M_n(R)$ ist dann die Determinante $\det(A)$, die wir gerade definiert haben, gleich der Determinante, die wir aus der Theorie über Körpern erhalten, wenn wir $A$ als Element von $M_n(K)$ betrachten. Es gelten, wie über jedem Körper, auch über $K$ die üblichen Rechenregeln, zum Beispiel:

SATZ 15.72. *Seien $A, B \in M_n(R)$. Dann gilt $\det(AB) = \det(A)\det(B)$. (Da beide Seiten dieser Gleichung Elemente von $R$ sind, gilt diese Gleichheit auch in $R$.)*

Zu einer Matrix $A \in M_n(R)$ können wir die Komplementärmatrix $A^{\mathrm{ad}}$ bilden (siehe Abschnitt I.9.3), die wieder in $M_n(R)$ liegt. Die Cramersche Regel Satz I.9.32 besagt, dass

$$AA^{\mathrm{ad}} = A^{\mathrm{ad}}A = \det(A)E_n$$

gilt. Alle hier auftretenden Matrizen liegen in $M_n(R)$, und für die Gleichheit spielt es keine Rolle, ob wir die Matrizen als Elemente von $M_n(R)$ oder von $M_n(K)$ auffassen. Daraus erhalten wir (vergleiche Korollar I.9.33) das folgende Korollar.

KOROLLAR 15.73. *Sei $A \in M_n(R)$. Es existiert genau dann eine Matrix $B \in M_n(R)$ mit $AB = BA = E_n$ (also ein multiplikatives Inverses von $A$ in dem Ring $M_n(R)$), wenn $\det(A) \in R^\times$.*

ERGÄNZUNG 15.74. Es ist nicht schwer zu zeigen, dass beide Sätze auch über beliebigen kommutativen Ringen gelten. Für den Determinantenproduktsatz kann man folgendermaßen vorgehen.

Als Vorüberlegung bemerken wir, dass für einen Ringhomomorphismus $f\colon R_1 \to R_2$ und eine Matrix $A = (a_{ij})_{i,j} \in M_n(R_1)$ gilt, dass $f(\det(A)) = \det(f(A))$, wenn wir mit $f(A)$ die Matrix bezeichnen, die aus $A$ durch Anwenden von $f$ auf jeden Eintrag von $A$ entsteht. Diese Gleichheit folgt direkt aus der Definition der Determinante durch die Leibniz-Formel.

Sei $R$ ein kommutativer Ring, und seien $A = (a_{ij})_{i,j}, B = (b_{ij})_{i,j} \in M_n(R)$.

Wir betrachten nun den Ring $\mathbb{Z}[X_{i,j}, Y_{i,j},\ i,j = 1,\dots,n]$, also den Polynomring über $\mathbb{Z}$ in $2n^2$ Unbestimmten $X_{i,j}$, $Y_{i,j}$. Wir erhalten einen (eindeutig bestimmten) Einsetzungshomomorphismus

$$\varphi\colon \mathbb{Z}[X_{i,j}, Y_{i,j},\ i,j = 1,\dots,n] \to R, \quad X_{i,j} \mapsto a_{ij},\ Y_{i,j} \mapsto b_{ij}.$$

Die Bilder der Elemente von $\mathbb{Z}$ unter $\varphi$ sind eindeutig festgelegt, denn $1 \in \mathbb{Z}$ muss auf $1 \in R$ abgebildet werden, und daraus ergeben sich die Bilder aller ganzen Zahlen daraus, dass $\varphi$ insbesondere ein Homomorphismus der zugrundeliegenden additiven Gruppen ist. (Vergleiche Beispiel 15.6.)

Wir schreiben $\tilde{A} = (X_{i,j})_{i,j}, \tilde{B} = (Y_{i,j})_{i,j} \in M_n(\mathbb{Z}[X_{i,j}, Y_{i,j}])$. Weil $\mathbb{Z}[X_{i,j}, Y_{i,j}]$ ein Integritätsring ist, gilt $\det(\tilde{A}\tilde{B}) = \det(\tilde{A})\det(\tilde{B})$, wie wir oben begründet haben.

Auf diese Gleichheit können wir den Ringhomomorphismus $\varphi$ anwenden. Mit Lemma 15.71 erhalten wir dann

$$\det(A)\det(B) = \varphi(\det(\tilde{A}))\varphi(\det(\tilde{B})) = \varphi(\det(\tilde{A}\tilde{B})) = \det(AB).$$

Im Fall der Cramerschen Regel können wir ähnlich argumentieren. Zunächst folgt aus dem Produktsatz, dass die Determinante einer über $R$ invertierbaren Matrix eine Einheit in $R$

ist. Sei nun andererseits $A \in M_n(R)$ eine Matrix mit $\det(A) \in R^\times$. Wie über einem Körper können wir zu $A$ die Komplementärmatrix $A^{\mathrm{ad}}$ bilden. Durch Reduktion auf den Fall des Integritätsrings $\mathbb{Z}[X_{ij}]$ genau wie beim Beweis des Produktsatzes sehen wir, dass das Produkt von $A$ und $A^{\mathrm{ad}}$ die Matrix $\det(A)E_n$ ist. Es folgt nun aus der Invertierbarkeit von $\det(A)$, dass auch $A$ invertierbar ist, und genauer erhalten wir die Formel $A^{-1} = \det(A)^{-1}A^{\mathrm{ad}}$.

Man nennt diese Methode die »Reduktion auf den universellen Fall«. □ Ergänzung 15.74

## 15.7. Ergänzungen *

Ergänzung 15.75 (Primideale). Die Primeigenschaft (Definition 15.44) kann man nicht nur für Elemente, sondern auch für Ideale in einem Ring definieren (und zwar auch in Ringen, die keine Integritätsringe sind).

Definition 15.76. Sei $R$ ein Ring. Ein Ideal $\mathfrak{p} \subset R$ heißt *Primideal*, wenn $\mathfrak{p} \neq R$ gilt und wenn für alle $x, y \in R$ gilt: Falls $xy \in \mathfrak{p}$, dann ist $x \in \mathfrak{p}$ oder $y \in \mathfrak{p}$. ⊣

Ist $R$ ein Integritätsring und $p \in R \setminus \{0\}$, so sieht man mit Lemma 15.34 leicht, dass $p$ genau dann ein Primelement ist, wenn das Hauptideal $(p)$ ein Primideal ist.

Andererseits kann zwar 0 per Definition kein Primelement sein, aber das Nullideal kann ein Primideal sein, genauer gilt:

Lemma 15.77. *Sei $R$ ein Ring. Dann sind äquivalent:*

(i) *Der Ring $R$ ist ein Integritätsring.*

(ii) *Das Nullideal in $R$ ist ein Primideal.*

Mit etwas mehr Arbeit kann man die folgende Aussage zeigen:

Satz 15.78. *Sei $f\colon R \to S$ ein Ringhomomorphismus.*

(1) *Wenn $S$ ein Integritätsring ist, dann ist $\operatorname{Ker}(f)$ ein Primideal in $R$.*

(2) *Wenn $f$ surjektiv ist und $\operatorname{Ker}(f)$ ein Primideal ist, dann ist $S$ ein Integritätsring.*

Sei nun $K$ ein Körper. Sei $f\colon \mathbb{Z} \to K$ der eindeutig bestimmte Ringhomomorphismus von $\mathbb{Z}$ nach $K$, siehe Beispiel 15.6. Der obige Satz sagt, dass $\mathfrak{p} := \operatorname{Ker}(f)$ ein Primideal von $\mathbb{Z}$ ist.

Ist $\mathfrak{p} \neq 0$, dann wird das Hauptideal $\mathfrak{p}$ von einer ganzen Zahl $p \neq 0$ erzeugt, von der wir ohne Einschränkung annehmen können, dass sie positiv ist. Da $\mathfrak{p}$ ein Primideal ist, ist $p$ eine Primzahl. Es ist dann leicht zu sehen, dass $p$ die Charakteristik des Körpers $K$ ist (Abschnitt I.4.2.2).

Gelte nun $\mathfrak{p} = \operatorname{Ker}(\mathbb{Z} \to K) = 0$, mit anderen Worten: Sei der Ringhomomorphismus $f\colon \mathbb{Z} \to K$ injektiv. Dann wird jede von Null verschiedene ganze Zahl auf eine Einheit in $K$ abgebildet und wir können $f$ fortsetzen zu einem Ringhomomorphismus

$$\mathbb{Q} \longrightarrow K, \quad \frac{a}{b} \mapsto \frac{f(a)}{f(b)}.$$

Dieser ist wieder injektiv, und sein Bild ist ein Teilkörper von $K$. Wir können also $\mathbb{Q}$ mit einem Teilkörper von $K$ identifizieren, genauer: Es gibt einen Isomorphismus von $\mathbb{Q}$ auf einen Teilkörper von $K$. □ Ergänzung 15.75

ERGÄNZUNG 15.79. Der Ring $\mathbb{Z}[i]$ ist euklidisch, also insbesondere faktoriell. Das kann man benutzen um zu beweisen, dass sich eine Primzahl $p > 2$ in $\mathbb{N}$ genau dann als Summe von zwei Quadraten schreiben lässt, wenn $p \equiv 1 \mod 4$ gilt. Siehe die Hausaufgaben auf den Übungsblättern 1, 2, 3.

Allgemein spielt die Ringtheorie eine sehr prominente Rolle in der elementaren und algebraischen Zahlentheorie, sowohl was die Untersuchung ähnlich konkreter (und einfacher) Fragen wie dieser angeht, als auch, was den weiteren konzeptionellen Aufbau der Theorie betrifft. □ Ergänzung 15.79

ERGÄNZUNG 15.80 (Der Satz von Mason und Stothers). Im Skript zur Linearen Algebra war kurz von der abc-Vermutung die Rede (Abschnitt I.3.5), die man als Vermutung über eine Eigenschaft des Rings $\mathbb{Z}$ der ganzen Zahlen verstehen sollte. Für den Polynomring $K[X]$ über einem Körper $K$ kann man eine analoge Aussage formulieren, deren Beweis interessanterweise gar nicht so schwierig ist. Dies ist der Satz von Mason und Stothers.

Um den Satz zu formulieren, definieren wir formal die »Ableitung« $f'$ eines Polynoms $f = \sum_{i=0}^{n} a_i X^i \in R[X]$ ($R$ ein kommutativer Ring) durch

$$f' = \sum_{i=1}^{n} i a_i X^{i-1},$$

also einfach durch Anwenden der üblichen Ableitungsregeln für Polynome. (Eine Interpretation wie über den reellen Zahlen, wo ein Grenzwertbegriff zur Verfügung steht, ist natürlich im allgemeinen Fall nicht möglich. Dennoch ist diese Definition öfters nützlich.) Man muss über allgemeinen Grundringen insofern ein bisschen aufpassen, als auch Polynome vom Grad $> 1$ als Ableitung das Nullpolynom haben können (zum Beispiel gilt das für $X^2 \in \mathbb{F}_2[X]$). Über einem Körper der Charakteristik 0, also einem Körper, der den Körper $\mathbb{Q}$ als Teilkörper enthält, tritt dieses Phänomen natürlich nicht auf.

Sei nun $K$ ein Körper. Das Radikal $\operatorname{rad}(f)$ eines Polynoms $f \in K[X]$ wird definiert als das Produkt aller normierten irreduziblen Polynome, die $f$ teilen. Es unterscheidet sich von $f$ also höchstens um den Leitkoeffizienten und dadurch, dass diese Teiler in der Primfaktorzerlegung von $f$ mit einem höheren Exponenten auftreten können. Zum Beispiel ist $\operatorname{rad}(X^n) = X$ für alle $n \geq 1$. Wenn $f$ vollständig in Linearfaktoren zerfällt (also zum Beispiel, wenn $K$ algebraisch abgeschlossen ist), dann ist $\deg(\operatorname{rad}(f))$ die Anzahl der verschiedenen Nullstellen von $f$ in $K$.

THEOREM 15.81 (Satz von Mason-Stothers). *Sei $K$ ein Körper und seien $a, b, c \in K[X] \setminus \{0\}$. Es gelte $\operatorname{ggT}(a, b) = 1$ und mindestens eines der Polynome $a', b', c'$ sei ungleich Null. Außerdem gelte*

$$a + b = c.$$

*Dann gilt*

$$\max(\deg(a), \deg(b), \deg(c)) \leq \deg(\operatorname{rad}(abc)) - 1.$$

Ein Beweis von Snyder wird auf der englischen Wikipedia-Seite[6] skizziert.

Als eine leichte Folgerung aus dem Theorem kann man zeigen, dass im Polynomring $K[X]$ über einem Körper der Charakteristik 0 das Analogon der Fermatschen Vermutung[7] gilt:

KOROLLAR 15.82. *Seien $K$ ein Körper der Charakteristik 0, $n \in \mathbb{N}$ und $x, y, z \in K[X]$ paarweise teilerfremde Polynome, von denen mindestens eines Grad $\geq 1$ hat und so dass*

$$x^n + y^n = z^n$$

*im Ring $K[X]$ gilt. Dann ist $n \leq 2$.*

[6] https://en.wikipedia.org/wiki/Mason%E2%80%93Stothers_theorem
[7] https://de.wikipedia.org/wiki/Gro%C3%9Fer_Fermatscher_Satz

BEWEIS. Da $x, y$ und $z$ paarweise teilerfremd sind, gilt $\operatorname{rad}(xyz) = \operatorname{rad}(x)\operatorname{rad}(y)\operatorname{rad}(z)$, und natürlich gilt $\operatorname{rad}(x) \mid x$, also $\deg(\operatorname{rad}(x)) \leq \deg(x)$, entsprechend für $y$ und $z$. Aus dem Satz von Mason und Stothers erhalten wir demnach

$$n \deg(x) = \deg(x^n) \leq \deg(x) + \deg(y) + \deg(z) - 1$$

und dieselbe Abschätzung auch für $n \deg(y)$ und $n \deg(z)$. Indem wir diese Ungleichungen addieren, sehen wir, dass

$$n(\deg(x) + \deg(y) + \deg(z)) \leq 3(\deg(x) + \deg(y) + \deg(z)) - 3$$

Da die Summe der Grade der drei Polynome als $> 0$ vorausgesetzt wurde, ist das nur für $n \leq 2$ möglich. □

*Zusatzfrage, die vermutlich nicht einfach ist.* Die Bedingung, dass $K$ Charakteristik 0 habe, ist hier nicht verzichtbar. Können Sie sehen, warum?

In der algebraischen Zahlentheorie und in der algebraischen Geometrie zeigt sich, dass die Ringe $\mathbb{Z}$ und $K[X]$ ($K$ ein Körper) viele Gemeinsamkeiten haben, und diese Analogie wird dort ausgebaut auf eine größere Klasse von Ringen (die nicht mehr notwendig Hauptidealringe, noch nicht einmal unbedingt faktoriell sind), die sogenannten Ganzheitsringe in Zahlkörpern einerseits und in Funktionenkörpern andererseits. Das ermöglicht es manchmal, zwischen eher zahlentheoretischen und eher geometrischen Fragestellungen und Methoden hin- und herzugehen und hat zu einer sehr engen Verzahnung der modernen algebraischen Zahlentheorie mit der algebraischen Geometrie geführt. □ Ergänzung 15.80

Und noch zwei »Platzhalter«, die ich hoffentlich später einmal mit mehr Inhalt füllen kann. Für den Moment gebe ich Ihnen nur Verweise auf andere Quellen.

ERGÄNZUNG 15.83. Bernstein-Polynome[8], siehe auch die englische Wikipedia[9]. Dies ist eine interessante Familie von Polynomen, die sowohl für theoretische Fragen als auch in der Praxis (Stichworte Computergrafik, Bezier-Kurven, Computer Aided Design) eine Rolle spielen. □ Ergänzung 15.83

ERGÄNZUNG 15.84 (Resultante und Diskriminante). Siehe zum Beispiel [**Bo-A**] 4.4. Die Diskriminante eines Polynoms (mit Koeffizienten in einem Körper $K$) ist ein allgemeiner Ausdruck in den Koeffizienten des Polynoms (eine »Formel«), die genau dann den Wert 0 hat, wenn das Polynom (in irgendeinem Erweiterungskörper von $K$) eine mehrfache Nullstelle hat.

Zum Beispiel ist die Diskriminante eines quadratischen Polynoms $aX^2 + bX + c$ gleich $b^2 - 4ac$ und Sie wissen (oder können es anhand der Lösungsformel für quadratische Gleichungen leicht nachprüfen), dass dieses Polynom genau dann eine doppelte Nullstelle hat, wenn $b^2 - 4ac = 0$ gilt.

Es ist interessant, dass es für Polynome beliebigen Grades möglich ist, anhand einer solchen Formel festzustellen, ob mehrfache Nullstellen vorliegen (in irgendeinem Erweiterungskörper von $K$), dass es aber andererseits für Polynome vom Grad $\geq 5$ keine allgemeine Formel für die Nullstellen selbst gibt. □ Ergänzung 15.84

[8] https://de.wikipedia.org/wiki/Bernsteinpolynom
[9] https://en.wikipedia.org/wiki/Bernstein_polynomial

KAPITEL 16

# Charakteristisches Polynom und Minimalpolynom

## 16.1. Das charakteristische Polynom

Sei $K$ ein Körper. Sei $V$ ein endlichdimensionaler $K$-Vektorraum und $f\colon V \to V$ ein Endomorphismus von $V$. Wir haben in der Linearen Algebra I den Begriff des Eigenwerts definiert und gesehen, dass $\lambda \in K$ genau dann ein Eigenwert von $f$ ist, wenn $\det(f - \lambda\ \mathrm{id}_V) = 0$ gilt, oder äquivalent, wenn $\det(\lambda\ \mathrm{id}_V - f) = 0$ gilt. Man kann also alle Eigenwerte von $f$ finden, indem man alle $\lambda$ findet, für die $\det(\lambda\ \mathrm{id}_V - f) = 0$ ist; das führt auf eine polynomiale Gleichung für $\lambda$, in der $\lambda^n$ und (in der Regel) kleinere Potenzen von $\lambda$ auftreten. Mit der neu eingeführten Sprache der Polynomringe und des Einsetzungshomomorphismus können wir die Theorie der Teilbarkeit in Polynomringen und der eindeutigen Primfaktorzerlegung hier mit einigem Nutzen anwenden, und wir machen daher die folgende Definition. (Wir bevorzugen jetzt die Version mit $\det(\lambda\ \mathrm{id}_V - f) = 0$, die vielleicht zunächst etwas unnatürlicher aussieht(?), aber den Vorteil hat, dass das im folgende definierte charakteristische Polynom von $f$ normiert ist.)

DEFINITION 16.1. (1) Sei $n \geq 0$ und $A \in M_n(K)$. Dann heißt das Polynom $\mathrm{charpol}_A(X) := \det(XE_n - A) \in K[X]$ das *charakteristische Polynom* der Matrix $A$.

(2) Sei $f\colon V \to V$ ein Endomorphismus des endlichdimensionalen $K$-Vektorraums $V$, $\mathscr{B}$ eine Basis von $V$, $A = M_{\mathscr{B}}^{\mathscr{B}}(f)$. Dann ist $\mathrm{charpol}_A(X)$ unabhängig von der Wahl der Basis $\mathscr{B}$ und heißt das *charakteristische Polynom* des Endomorphismus $f$. Wir bezeichnen dieses Polynom mit $\mathrm{charpol}_f \in K[X]$.

⊣

Hier ist $XE_n - A$ eine Matrix mit Einträgen im Polynomring $K[X]$, also ein Element von $M_n(K[X])$. Wie in Abschnitt 15.6 erklärt, ist die Determinante einer solchen Matrix durch die Leibniz-Formel definiert, wir können also das charakteristische Polynom der Matrix $A$ schreiben als

$$\mathrm{charpol}_A = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} (\delta_{i,\sigma(i)} X - a_{i,\sigma(i)}),$$

wobei wir

$$\delta_{i,j} = \begin{cases} 1 & \text{wenn } i = j \\ 0 & \text{wenn } i \neq j \end{cases} \qquad \text{(Kronecker-delta)}$$

setzen. Für die Definition des charakteristischen Polynoms kann man also auf die Diskussion in Abschnitt 15.6 verzichten. Um die Aussage über die Unabhängigkeit in Teil (2) zu beweisen, die aus dem nächsten Lemma folgt (bzw. dazu äquivalent ist), benutzen wir aber Satz 15.72.

Das Lemma besagt, das zueinander konjugierte Matrizen dasselbe charakteristische Polynom haben. Insbesondere ist das charakteristische Polynom für alle darstellenden Matrizen eines Endomorphismus dasselbe (natürlich muss »oben und unten« dieselbe Basis verwendet werden).

LEMMA 16.2. *Seien $K$ ein Körper, $n \in \mathbb{N}$, $A \in M_n(K)$ und $S \in GL_n(K)$. Dann gilt*

$$\operatorname{charpol}_A = \operatorname{charpol}_{SAS^{-1}}.$$

BEWEIS. Wir können $S$ und $S^{-1}$ als Elemente von $M_n(K[X])$ auffassen und haben dann nach Satz 15.72, dass

$$\det(XE_n - SAS^{-1}) = \det(S(XE_n - A)S^{-1}) = \det(S)\det(XE_n - A)\det(S^{-1}) = \det(XE_n - A).$$

Das ist die Behauptung des Lemmas. □

BEISPIEL 16.3. Wir berechnen das charakteristische Polynom in einigen konkreten Beispielen. Im Prinzip ist klar, was zu tun ist: Es ist eine Determinante auszurechnen, und dafür kann man die üblichen Verfahren benutzen.

(1) Sei

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \in M_3(\mathbb{Q}).$$

Es gilt

$$\begin{aligned} \operatorname{charpol}_A &= \det(XE_3 - A) \\ &= \det \begin{pmatrix} X-1 & 0 & -2 \\ -2 & X-1 & 0 \\ 0 & -1 & X-1 \end{pmatrix} = (X-1)^3 - 2 \cdot 2 \\ &= X^3 - 3X^2 + 3X - 5, \end{aligned}$$

wobei zur Berechnung der Determinante nach der ersten Zeile entwickelt wurde.

(2) Sei $K$ ein Körper und $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(K)$. Dann gilt

$$\operatorname{charpol}_A = \det \begin{pmatrix} X-a & -b \\ -c & X-d \end{pmatrix} = (X-a)(X-d) - bc = X^2 - (a+d)X + (ad-bc).$$

Der Absolutterm ist also $\det(A)$, der Koeffizient von $X$ ist $-\operatorname{Spur}(A)$ (siehe auch unten).

(3) Seien $K$ ein Körper, $n \in \mathbb{N}$ und sei $A = (a_{ij})_{i,j} \in M_n(K)$ eine *obere Dreiecksmatrix*. Dann ist auch $XE_n - A$ eine obere Dreiecksmatrix und folglich gilt

$$\operatorname{charpol}_A = (X - a_{11}) \cdot \cdots \cdot (X - a_{nn}).$$

◊

Alle Aussagen über das charakteristische Polynom lassen zwei Fassungen zu, eine für Matrizen und eine analoge für Endomorphismen eines endlichdimensionalen Vektorraums. Die Übersetzung zwischen den beiden Sichtweisen ist einfach, so dass wir im folgenden meist nur eine der beiden Versionen explizit ausschreiben – je nachdem, wie der Beweis natürlicher ist.

LEMMA 16.4. *Sei $A \in M_n(K)$. Dann gilt*

$$\operatorname{charpol}_A = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0,$$

*d.h.* $\operatorname{charpol}_A$ *ist normiert vom Grad $n$. Außerdem ist $a_0 = \det(-A) = (-1)^n \det A$.*

BEWEIS. Dass das charakteristische Polynom normiert vom Grad $n$ ist, folgt aus der Definition und der Leibniz-Formel. Dass wir ein normiertes Polynom erhalten, ist der Grund, warum wir mit $\det(XE_n - A)$ statt mit $\det(A - XE_n)$ arbeiten (aber es gibt auch Quellen, die es anders machen).

Außerdem gilt $a_0 = \operatorname{charpol}_A(0) = \det(0 \cdot E_n - A) = (-1)^n \det(A)$. Beim mittleren Gleichheitszeichen benutzen wir Lemma 15.71 für den Einsetzungshomomorphismus $K[X] \to K$, $X \mapsto 0$. □

Das folgende einfache Lemma ist mehrfach nützlich.

LEMMA 16.5. *Seien $K$ ein Körper, $V$ ein endlichdimensionaler $K$-Vektorraum und $f$ ein Endomorphismus von $V$. Sei $U \subseteq V$ ein Untervektorraum mit $f(U) \subseteq U$ und sei $W \subseteq V$ ein Komplementärraum zu $U$.*

*Sei $g := f_{|U}$ die Einschränkung von $f$ auf $U$, und sei $h$ die Verkettung*

$$W \to V \xrightarrow{f} V \to W,$$

*wobei links die Inklusion von $W$ nach $V$ und rechts die Projektion von $V = U \oplus W$ auf $W$ steht (also die Abbildung $U \oplus W \to W$, $u + w \mapsto w$ ($u \in U$, $w \in W$)).*

*Dann gilt*

$$\operatorname{charpol}_f = \operatorname{charpol}_g \cdot \operatorname{charpol}_h .$$

BEWEIS. Übung. □

Wir haben die Definition des charakteristischen Polynoms damit motiviert, dass seine Nullstellen, bzw. äquivalent die Nullstellen der zugehörigen Polynomfunktion gerade die Eigenwerte der zugehörigen Matrix bzw. des zugehörigen Endomorphismus sind. Das halten wir noch einmal im folgenden Satz fest.

SATZ 16.6. *Seien $K$ ein Körper, $V$ ein endlichdimensionaler $K$-Vektorraum und $f\colon V \to V$ ein Endomorphismus. Es bezeichne $\operatorname{charpol}_f$ das charakteristische Polynom von $f$. Ein Element $\lambda \in K$ ist genau dann eine Nullstelle von $\operatorname{charpol}_f$, wenn $\lambda$ ein Eigenwert von $f$ ist.*

Wir können aber den Satz noch präzisieren.

SATZ 16.7. *Seien $K$ ein Körper, $V$ ein endlichdimensionaler $K$-Vektorraum und $f\colon V \to V$ ein Endomorphismus. Es bezeichne $\chi := \operatorname{charpol}_f$ das charakteristische Polynom von $f$.*

(1) *Sei $\lambda \in K$. Es gilt $\operatorname{mult}_\lambda(\chi) > 0$ genau dann, wenn $\lambda$ ein Eigenwert von $f$ ist.*
*In diesem Fall gilt*
$$\dim V_\lambda(f) \leq \operatorname{mult}_\lambda(\chi).$$
*Man nennt $\dim V_\lambda(f)$ auch die* geometrische Vielfachheit *und $\operatorname{mult}_\lambda(\chi)$ die* algebraische Vielfachheit *des Eigenwerts $\lambda$.*

(2) *Der Endomorphismus $f$ ist genau dann diagonalisierbar, wenn $\operatorname{charpol}_f$ vollständig in Linearfaktoren zerfällt und für alle Eigenwerte $\lambda$ von $f$ die Gleichheit $\dim V_\lambda(f) = \operatorname{mult}_\lambda(\chi)$ gilt.*

BEWEIS. zu (1). Dass $\operatorname{mult}_\lambda(\chi) > 0$ gilt, ist dazu äquivalent, dass $\lambda$ eine Nullstelle von $\chi$ ist, also dass $\det(\lambda \operatorname{id} - f) = 0$ gilt. Wie oben besprochen, heißt das genau, dass $\lambda$ ein Eigenwert von $f$ ist.

Um die Abschätzung $\dim V_\lambda(f) \leq \operatorname{mult}_\lambda(\chi)$ zu zeigen, nutzen wir aus, dass wir $\operatorname{charpol}_f$ als das charakteristische Polynom der darstellenden Matrix von $f$ bezüglich einer Basis unserer Wahl berechnen können. Die Basis, die wir betrachten wollen, konstruieren wir, indem wir eine Basis von $V_\lambda(f)$ zu einer Basis $\mathscr{B}$ von $V$ ergänzen. Dann sind die ersten $r := \dim(V_\lambda(f))$ Vektoren in dieser Basis Eigenvektoren von $f$ zum Eigenwert $\lambda$. Die Matrix $M^{\mathscr{B}}_{\mathscr{B}}(f)$ hat also die Form $\begin{pmatrix} \lambda E_r & B \\ 0 & D \end{pmatrix}$ (als Blockmatrix geschrieben). Es gilt dann $\operatorname{charpol}_f =$

$\operatorname{charpol}_{\lambda E_r} \cdot \operatorname{charpol}_D = (X-\lambda)^r \operatorname{charpol}_D$ (diese Rechnung kann man als einen Spezialfall von Lemma 16.5 betrachten), also $\operatorname{mult}_\lambda(\chi) \geq r$.

zu (2). Dass $f$ diagonalisierbar ist, ist dazu äquivalent, dass die (direkte) Summe der Eigenräume von $A$ gleich $V$ ist, also dazu, dass die Summe der Dimensionen aller Eigenräume zu den verschiedenen Eigenwerten gleich $n$ ist. Nun ist $\deg(\chi) = n$, und die Summe der Vielfachheiten der Nullstellen von $\chi$ ist genau dann $n$, wenn $\chi$ vollständig in Linearfaktoren zerfällt. Das Kriterium folgt deswegen aus Teil (1). □

Die Bedingung, dass das charakteristische Polynom eines Endomorphismus (bzw. einer Matrix) vollständig in Linearfaktoren zerfällt, hat (auch unabhängig von der Frage, ob die geometrischen und algebraischen Vielfachheiten der Eigenwerte übereinstimmen) eine natürliche Interpretation. Dazu machen wir die folgende Definition.

DEFINITION 16.8. Eine Matrix $A \in M_n(K)$ heißt *trigonalisierbar*, wenn $A$ zu einer oberen Dreiecksmatrix konjugiert ist. Ein Endomorphismus von $V$ heißt trigonalisierbar, wenn eine Basis $\mathscr{B}$ von $V$ existiert, so dass die beschreibende Matrix $M^{\mathscr{B}}_{\mathscr{B}}(f)$ bezüglich dieser Basis eine obere Dreiecksmatrix ist. ⊣

SATZ 16.9. *Seien $K$ ein Körper und $V$ ein endlichdimensionaler $K$-Vektorraum. Ein Endomorphismus $f$ von $V$ ist genau dann trigonalisierbar, wenn sein charakteristisches Polynom vollständig in Linearfaktoren zerfällt.*

BEWEIS. Das charakteristische Polynom einer oberen Dreiecksmatrix zerfällt offenbar vollständig in Linearfaktoren (Beispiel 16.3 (3)), also gilt das auch für trigonalisierbare Endomorphismen.

Um die Umkehrung zu zeigen, führen wir Induktion nach der Dimension $n$ des Vektorraums $V$. Im Fall $n \leq 1$ ist jede $(n \times n)$-Matrix eine obere Dreiecksmatrix. Sei nun $n > 1$ und sei $f$ ein Endomorphismus, dessen charakteristisches Polynom vollständig in Linearfaktoren zerfällt. Dann besitzt das charakteristische Polynom eine Nullstelle $\lambda$, also hat $f$ einen Eigenvektor $v \in V \setminus \{0\}$.

Wir setzen $b_1 := v$ und ergänzen diesen Vektor (der ja $\neq 0$ ist, weil es sich um einen Eigenvektor handelt) zu einer Basis $\mathscr{B} = (b_1, \ldots, b_n)$. Aus Lemma 16.5, angewandt auf die Zerlegung $V = U \oplus W$ mit $U := \langle b_1 \rangle$ und $W = \langle b_2, \ldots, b_n \rangle$, folgt

$$\operatorname{charpol}_f = (X-\lambda) \cdot \operatorname{charpol}_h,$$

wobei $h\colon W \to W$ die in Lemma 16.5 beschriebene Abbildung ist.

Weil $\operatorname{charpol}_f$ vollständig in Linearfaktoren zerfällt, folgt aus der Eindeutigkeit der Primfaktorzerlegung im Ring $K[X]$, dass das auch für $\operatorname{charpol}_h$ gilt. Nach Induktionsvoraussetzung existiert also eine Basis $\mathscr{C} = (c_2, \ldots, c_n)$ von $W$, so dass $M^{\mathscr{C}}_{\mathscr{C}}(g)$ eine obere Dreiecksmatrix ist. Die Matrix, die $f$ bezüglich der Basis $(b_1, c_2, \ldots, c_n)$ darstellt, hat die Form

$$\begin{pmatrix} \lambda & * \\ 0 & M^{\mathscr{C}}_{\mathscr{C}}(h) \end{pmatrix}$$

und ist mithin eine obere Dreiecksmatrix. Also ist $f$ trigonalisierbar. □

**16.1.1. Die Spur einer Matrix.** Wir kommen noch einmal auf die Spur einer Matrix (oder eines Endomorphismus) zurück, siehe Abschnitt I.9.4. Für eine Matrix $A = (a_{ij})_{i,j} \in M_n(K)$ haben wir

$$\operatorname{Spur}(A) = \sum_{i=1}^{n} a_{ii} \in K$$

definiert. Die Spur von $A$ ist also einfach die Summe der Diagonaleinträge. Wir haben gezeigt (Korollar I.9.37), dass zueinander konjugierte Matrizen dieselbe Spur haben, so dass wir die Spur eines Endomorphismus $f$ als die Spur irgendeiner darstellenden Matrix von $f$ bezüglich einer Basis des zugrundeliegenden Vektorraums definieren können. Das Ergebnis ist unabhängig von der Wahl der Basis.

Mithilfe des charakteristischen Polynoms erhalten wir einen neuen Beweis, dass zueinander konjugierte Matrizen dieselbe Spur haben, denn es gilt:

LEMMA 16.10. (1) *Sei* $A \in M_n(K)$, *und schreibe* $\operatorname{charpol}_A = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$. *Dann gilt* $\operatorname{Spur}(A) = -a_{n-1}$.

(2) *Ist* $f$ *ein Endomorphismus eines* $n$*-dimensionalen Vektorraums* $V$ *mit* $\operatorname{charpol}_f = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$, *so gilt* $\operatorname{Spur}(f) = -a_{n-1}$.

BEWEIS. zu (1). Die Behauptung folgt leicht aus der Definition des charakteristischen Polynoms als Determinante und aus der Leibniz-Formel. Ein Summand der Leibnizformel, etwa zu einer Permutation $\sigma \in S_n$, kann nämlich nur dann einen Beitrag zum Koeffizienten von $X^{n-1}$ liefern, wenn in dem zugehörigen Produkt mindestens $n-1$ der Diagonaleinträge von $XE_n - A$ auftreten, also $\sigma(i) = i$ für alle bis auf höchstens ein $i$ in $\{1, \ldots, n\}$ gilt. Dann muss aber $\sigma = \mathrm{id}$ sein. Der zur Identität gehörige Summand ist $\prod_{i=1}^{n}(X - a_{ii})$, und der Koeffizient von $X^{n-1}$ in diesem Ausdruck ist $-\sum_{i=1}^{n} a_{ii}$.

Teil (2) folgt nun, indem wir den ersten Teil auf eine darstellende Matrix von $f$ anwenden. □

## 16.2. Das Minimalpolynom

Neben dem charakteristischen Polynom ordnet man jeder Matrix (bzw. jedem Endomorphismus) ein weiteres Polynom zu, das sogenannte Minimalpolynom. Wie wir sehen werden, enthalten diese beiden Polynome wesentliche Informationen über die zugrundeliegende Matrix, und insbesondere über ihre Eigenwerte und Eigenräume. Zum Beispiel werden wir am Ende dieses Kapitels beweisen, dass eine Matrix genau dann diagonalisierbar ist, wenn ihr Minimalpolynom vollständig in Linearfaktoren zerfällt und nur einfache Nullstellen hat.

Sei $K$ ein Körper und sei $n \in \mathbb{N}$. Sei $A \in M_n(K)$, und sei $\Phi\colon K[X] \to M_{n\times n}(K)$ der Ringhomomorphismus mit $\Phi(a) = aE_n$ für alle $a \in K$ und $\Phi(X) = A$ (eine Instanz des Einsetzungshomomorphismus, Satz 15.24), . Wir schreiben $K[A]$ für das Bild von $\Phi$ – dies ist ein kommutativer Unterring von $M_n(K)$, der $K$ enthält (und auch ein $K$-Vektorraum ist).

Weil $\Phi$ insbesondere ein Homomorphismus von $K$-Vektorräumen ist, der Vektorraum $K[X]$ nicht endlichdimensional, der Zielraum $M_n(K)$ jedoch endlichdimensional ist, kann $\Phi$ nicht injektiv sein. Der Kern von $\Phi$ ist also nicht das Nullideal. Es handelt sich um ein Hauptideal in $K[X]$, etwa $\operatorname{Ker}(\Phi) = (p)$, $p \neq 0$. Das Ideal $(p)$ ändert sich nicht, wenn wir $p$ mit einem Element aus $K^\times$ multiplizieren. Daher ist die folgende Definition sinnvoll.

DEFINITION 16.11. Sei wie oben $A \in M_n(K)$ und $\Phi\colon K[X] \to M_n(K)$, $X \mapsto A$. Das *Minimalpolynom* $\operatorname{minpol}_A$ von $A$ ist das eindeutig bestimmte normierte Polynom $p \in K[X]$ mit $\operatorname{Ker}\Phi = (p)$. ⊣

Etwas konkreter können wir das so formulieren: Für $p := \operatorname{minpol}_A$ gilt $p(A) = 0$, und alle Polynome $q \in K[X]$ mit $q(A) = 0$ werden von $p$ geteilt. Insbesondere haben alle $q \in K[X] \setminus \{0\}$ mit $q(A)$ Grad $\deg(q) \geq \deg \operatorname{minpol}_A$. Wir können also äquivalent sagen: Das Minimalpolynom $\operatorname{minpol}_A$ von $A$ ist das eindeutig bestimmte normierte Polynom $p$ kleinsten Grades, so dass $p(A) = 0$ gilt.

Wie üblich können wir eine analoge Definition für Endomorphismen endlichdimensionaler $K$-Vektorräume machen.

DEFINITION 16.12. Seien $K$ ein Körper, $V$ ein endlichdimensionaler Vektorraum über $K$ und $f \in \mathrm{End}_K(V)$. Sei $\Phi\colon K[X] \to \mathrm{End}_K(V)$ der Einsetzungshomomorphismus mit $X \mapsto f$.

Das eindeutig bestimmte normierte Polynom $p$, das das Ideal $\mathrm{Ker}(\Phi)$ erzeugt, heißt das *Minimalpolynom* des Endomorphismus $f$. ⊣

Die konkrete(re) Beschreibung für das Minimalpolynom einer Matrix lässt sich natürlich auf den Fall von Endomorphismen übertragen.

BEISPIEL 16.13. Sei $K$ ein Körper, $n \in \mathbb{N}$.

Ist $A = \mathrm{diag}(a_1, \ldots, a_n)$ eine Diagonalmatrix, so gilt für jedes Polynom $f \in K[X]$, dass $f(A) = \mathrm{diag}(f(a_1), \ldots, f(a_n))$. Schreiben wir $\{a_1, \ldots, a_n\} = \{\lambda_1, \ldots, \lambda_r\}$ mit *paarweise verschiedenen* $\lambda_1, \ldots, \lambda_r$ $(r \leq n)$, so gilt

$$\mathrm{minpol}_A = \prod_{i=1}^{r} (X - \lambda_i),$$

denn es ist nach der obigen Bemerkung klar, dass dieses Polynom die Matrix $A$ annulliert, aber keiner seiner echten Teiler diese Eigenschaft hat.

Ist speziell $A = aE_n$ ein Vielfaches der Einheitsmatrix, $a \in K^\times$, so gilt $\mathrm{minpol}_A = X - a$. Das Minimalpolynom der Nullmatrix ist das Polynom $X$. ◊

Anhand dieser Beispiele sieht man, dass jedenfalls alle Zahlen zwischen 1 und $n$ als Grad des Minimalpolynoms auftreten können. Weil $\dim_K(M_n(K)) = n^2$ ist, ist nicht schwer zu sehen, dass der Grad des Minimalpolynoms höchstens $n^2$ sein kann. Wir werden später (als Folgerung des Satzes von Cayley–Hamilton) zeigen, dass aber sogar immer $\deg(\mathrm{minpol}_A) \leq n$ gilt.

Die Begriffe des Minimalpolynoms für Matrizen und Endomorphismen sind in der offensichtlichen Art und Weise miteinander kompatibel. Das geht damit einher, dass zueinander konjugierte Matrizen dasselbe Minimalpolynom haben. Diese beiden Tatsachen halten wir im folgenden Lemma fest.

LEMMA 16.14. *Sei $K$ ein Körper.*

(1) *Sei $V$ ein endlichdimensionaler $K$-Vektorraum, $n = \dim V$ und sei $\mathscr{B}$ eine Basis von $V$. Ist $f$ ein Endomorphismus von $V$, so gilt*

$$\mathrm{minpol}_f = \mathrm{minpol}_{M_{\mathscr{B}}^{\mathscr{B}}(f)} \quad \in K[X].$$

(2) *Seien $n \in \mathbb{N}$, $A \in M_n(K)$, $S \in GL_n(K)$. Dann haben $A$ und $SAS^{-1}$ dasselbe Minimalpolynom.*

BEWEIS. zu (1). Es genügt zu zeigen, dass für ein Polynom $p \in K[X]$ genau dann $p(f) = 0$ gilt, wenn $p(M_{\mathscr{B}}^{\mathscr{B}}(f)) = 0$ ist. Das folgt direkt daraus, dass die Abbildung $M_{\mathscr{B}}^{\mathscr{B}}(-)\colon \mathrm{End}_K(V) \to M_n(K)$, $g \mapsto M_{\mathscr{B}}^{\mathscr{B}}(g)$, ein Ringisomorphismus ist.

Wir können die Situation in dem folgenden »kommutativen Diagramm« veranschaulichen (»kommutativ« heißt hier, dass die Verkettung $\Phi_A \circ M_{\mathscr{B}}^{\mathscr{B}}(-)$ mit $\Phi_f$ übereinstimmt).

$$\begin{array}{ccc} & K[X] & \\ {\scriptstyle \Phi_f}\swarrow & & \searrow{\scriptstyle \Phi_A} \\ \mathrm{End}_K(V) & \xrightarrow{M_{\mathscr{B}}^{\mathscr{B}}(-)} & M_n(K) \end{array}$$

Hier bezeichnet $\Phi_f$ den Einsetzungshomomorphismus, der durch $X \mapsto f$ bestimmt ist, und $\Phi_A$ denjenigen mit $X \mapsto A$.

Um Teil (2) zu beweisen, kann man Teil (1) anwenden (denn $A$ und $SAS^{-1}$ sind darstellende Matrizen des Endomorphismus $f_A : K^n \to K^n$ bezüglich unterschiedlicher Basen). Alternativ kann man ein analoges Argument für den Ringisomorphismus $M_n(K) \to M_n(K), B \mapsto SBS^{-1}$, durchführen. Dass diese Abbildung ein Ringisomorphismus ist, impliziert, dass $p(SAS^{-1}) = Sp(A)S^{-1}$ für jedes $p \in K[X]$ gilt. Insbesondere sind die Aussagen $p(A) = 0$ und $p(SAS^{-1}) = 0$ für jedes $p$ äquivalent. □

## 16.3. Der Satz von Cayley–Hamilton

In diesem Abschnitt beweisen wir den wichtigen *Satz von Cayley–Hamilton*. Der Satz ist benannt nach Arthur Cayley[1] (1821–1895), der als einer der ersten Mathematiker systematisch mit Matrizen gearbeitet hat, und William Rowan Hamilton[2] (1805–1865) (den wir im Zusammenhang mit den Quaternionen schon in der Linearen Algebra 1 erwähnt hatten). Sowohl Cayley als auch Hamilton haben aber nur Spezialfälle des Satzes bewiesen. Den ersten allgemeinen Beweis (jedenfalls über dem Körper $\mathbb{C}$) gab im Jahr 1878 Ferdinand Georg Frobenius[3] (1849–1917).

> ”
>
> As for everything else, so for a mathematical theory: beauty can be perceived but not explained.
>
> Arthur Cayley
> (angeblich) in: The Collected Mathematical Papers of Arthur Cayley (ed. 1895)
> (ich habe aber die 14 Bände mit jeweils
> mehreren hundert Seiten nicht alle durchgeschaut…)

Wir beginnen mit einigen Vorbereitungen für den Beweis. Seien $K$ ein Körper und $V$ ein $K$-Vektorraum.

Definition 16.15. Sei $f \in \mathrm{End}_K(V)$. Ein Untervektorraum $U \subseteq V$ heißt *$f$-invariant*, wenn $f(U) \subseteq U$ gilt. ⊣

Definition 16.16. Sei $f \in \mathrm{End}_K(V)$. Ein Untervektorraum $U \subseteq V$ heißt *$f$-zyklischer* Unterraum, falls $u \in U$ existiert mit $U = \langle u, f(u), f^2(u), \dots \rangle$. ⊣

Offenbar ist jeder $f$-zyklische Unterraum auch $f$-invariant. Ein $f$-invarianter Unterraum muss jedoch nicht $f$-zyklisch sein. (Suchen Sie hierfür ein Beispiel.)

Lemma 16.17. *Seien $K$ ein Körper und $V$ ein $K$-Vektorraum. Sei $U = \langle u, f(u), f^2(u), \dots \rangle \subseteq V$ ein endlichdimensionaler $f$-zyklischer Unterraum und sei $i = \dim U$. Dann ist $u, f(u), \dots, f^{i-1}(u)$ eine Basis von $U$.*

[1] https://en.wikipedia.org/wiki/Arthur_Cayley
[2] https://en.wikipedia.org/wiki/William_Rowan_Hamilton
[3] https://en.wikipedia.org/wiki/Ferdinand_Georg_Frobenius

**Über lineare Substitutionen und bilineare Formen**

Journal für die reine und angewandte Mathematik 84, 1–63 (1878)

In den Untersuchungen über die Transformation der quadratischen Formen in sich selbst hat man sich bisher auf die Betrachtung des allgemeinen Falles beschränkt, während die Ausnahmen, welche die Resultate in gewissen speciellen Fällen erfahren, nur für die ternären Formen erschöpfend behandelt worden sind (*Bachmann*, dieses Journal Bd. 76, S. 331; *Hermite*, dieses Journal Bd. 78, S. 325). Ich habe daher versucht, die Lücke zu ergänzen, die sich sowohl in dem Beweise der Formeln findet, welche die Herren *Cayley* (dieses Journal Bd. 32, S. 119) und *Hermite* (dieses Journal Bd. 47, S. 309) für die Coefficienten der Substitution gegeben haben, als auch in den Betrachtungen, welche Herr *Rosanes* (dieses Journal Bd. 80, S. 52) über den Charakter der Transformation angestellt

3. Nach Formel (2.) genügt jede Form $A$ einer gewissen Gleichung, und der Grad der Gleichung niedrigsten Grades $\psi(A) = 0$ ist nicht grösser als $n$. Ist $f(r)$ eine durch $\psi(r)$ theilbare ganze Function, $f(r) = \psi(r)\chi(r)$, so ist $f(A) = \psi(A)\chi(A) = 0$. Da die charakteristische Function $\varphi(r)$ durch $\psi(r)$ theilbar ist, so ist folglich stets $\varphi(A) = 0$. Sind $f(r)$ und $g(r)$ irgend zwei ganze Functionen von $r$, und ist $h(r)$ ihr grösster gemeinsamer Divisor, so lassen sich zwei ganze Functionen $F(r)$ und $G(r)$ so bestimmen, dass $f(r)G(r) - g(r)F(r) = h(r)$ ist. Daher ist auch $f(A)G(A) - g(A)F(A) = h(A)$. Genügt also $A$ den Gleichungen $f(A) = 0$ und $g(A) = 0$, so muss es auch die Gleichung $h(A) = 0$ befriedrigen.

ABBILDUNG 1. Zwei Ausschnitte aus der Arbeit *Über lineare Substitutionen und bilineare Formen*, Journal für die reine und angewandte Mathematik **84**, 1–63 (1878) von F. G. Frobenius. In dem zweiten Ausschnitt ist der »Satz von Cayley–Hamilton« markiert. Dass das Minimalpolynom, das im Artikel mit $\psi$ bezeichnet wird, das charakteristische Polynom (hier: $\varphi$) teilt, wurde vorher bewiesen. Aus F. G. Frobenius, *Gesammelte Abhandlungen I*, Hrsg. J.-P. Serre, Springer 1968

BEWEIS. Sei $j$ maximal mit der Eigenschaft, dass $u, f(u), \ldots, f^{j-1}(u)$ eine linear unabhängige Familie von Vektoren ist. Weil $U$ endliche Dimension hat, existiert ein solches $j$. Die Maximalität von $j$ impliziert, dass $a_0, \ldots, a_{j-1} \in K$ existieren mit

$$f^j(u) = \sum_{l=0}^{j-1} a_l f^l(u).$$

Folglich ist $\langle u, \dots, f^{j-1}(u)\rangle$ ein $f$-invarianter Unterraum, er enthält also alle Elemente der Form $f^n(u)$ und stimmt somit mit $U$ überein. Es folgt $j = i$, und daraus folgt die Behauptung. □

DEFINITION 16.18. Seien $K$ ein Körper, $n \in \mathbb{N}$. Sei $\chi = X^n + \sum_{i=0}^{n-1} a_i X^i \in K[X]$ ein normiertes Polynom vom Grad $n$. Dann heißt die Matrix

$$\begin{pmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & \vdots \\ & & \ddots & 0 & \vdots \\ & & & 1 & -a_{n-1} \end{pmatrix}$$

(wobei alle Einträge, die nicht hingeschrieben sind, $= 0$ sind) die *Begleitmatrix von* $\chi$. ⊣

BEMERKUNG 16.19. Ein Untervektorraum $U \subseteq V$ ist genau dann $f$-zyklisch, wenn $f(U) \subseteq U$ gilt und eine Basis von $U$ existiert, so dass die Matrix von $f_{|U}$ bezüglich dieser Basis die Form einer Begleitmatrix hat. ◊

LEMMA 16.20. *Sei $A \in M_n(K)$ die Begleitmatrix des normierten Polynoms $\chi$ (vom Grad $n$). Dann gilt* $\operatorname{charpol}_A = \chi$.

BEWEIS. Wir führen Induktion nach $n$. Für $n = 1$ ist die Sache klar. Im allgemeinen Fall ist die Determinante der Matrix

$$\begin{pmatrix} X & & & & a_0 \\ -1 & X & & & a_1 \\ & -1 & \ddots & & \vdots \\ & & \ddots & X & \vdots \\ & & & -1 & X + a_{n-1} \end{pmatrix}$$

zu berechnen. (Wir lassen die Nullen wieder der Übersichtlichkeit halber weg.) Durch Entwicklung nach der ersten Spalte erhalten wir als Determinante

$$X \cdot \det \begin{pmatrix} X & & & & a_1 \\ -1 & X & & & a_2 \\ & -1 & \ddots & & \vdots \\ & & \ddots & X & \vdots \\ & & & -1 & X + a_{n-1} \end{pmatrix} + \det \begin{pmatrix} 0 & & & a_0 \\ -1 & \ddots & & \vdots \\ & \ddots & 0 & \vdots \\ & & -1 & X + a_{n-1} \end{pmatrix}$$

Die Determinante im linken Summanden können wir nach Induktionsvoraussetzung schreiben, die Determinante im zweiten Summanden ist gleich $a_0$, wie man durch Entwicklung nach der ersten Zeile sieht. Wir haben also insgesamt

$$X \cdot (X^{n-1} + a_{n-1}X^{n-2} + \cdots + a_1) + a_0$$

und das ist gleich $\chi$, wie behauptet. □

Nun können wir den Satz von Cayley–Hamilton formulieren und beweisen.

SATZ 16.21 (Cayley–Hamilton). (1) *Ist $A \in M_n(K)$, so gilt* $\operatorname{charpol}_A(A) = 0 (\in M_{n\times n}(K))$.

(2) *Ist $f$ ein Endomorphismus des endlich-dimensionalen $K$-Vektorraums $V$, so gilt* $\operatorname{charpol}_f(f) = 0 (\in \operatorname{End}_K(V))$.

Jedenfalls für eine Diagonalmatrix $A$ ist die Aussage (von Teil (1)) klar. Diese einfache Beobachtung kann man sogar zu einem vollständigen Beweis machen, siehe Bemerkung 16.30.

Andererseits sei schon hier die Warnung formuliert, dass die folgende Gleichungskette

$$\operatorname{charpol}_A(A) = \det(A \cdot E_n - A) = \det(0) = 0$$

kein Beweis des Satzes ist, weil es sich nämlich gar nicht überall um Gleichungen handeln kann, denn links steht eine *Matrix in* $M_n(K)$, rechts aber ein *Element des Körpers* $K$. Siehe Bemerkung 16.22

BEWEIS. Es ist klar, dass die Aussagen (1) und (2) auseinander hervorgehen, es genügt daher, den zweiten Teil zu zeigen.

Sei also $f \in \operatorname{End}_K(V)$ und $\chi = \operatorname{charpol}_f$. Es genügt zu zeigen, dass $\chi(f)(v) = 0$ für alle $v \in V$ gilt, denn das bedeutet ja gerade, dass der Endomorphismus $\chi(f)$ die Nullabbildung ist.

Sei also $v \in V$. Wir betrachten den $f$-zyklischen Unterraum $U = \langle v, f(v), f^2(v), \dots \rangle$. Es gilt dann $f(U) \subseteq U$. Wir betrachten die Einschränkung $f_{|U}$ als Endomorphismus von $U$ und bezeichnen mit $\xi$ sein charakteristisches Polynom. Das charakteristische Polynom von $f$ ist nach Lemma 16.5 ein Vielfaches von $\xi$, etwa $\chi = \zeta \cdot \xi$. Aus $\xi(f)(v) = 0$ folgt also $\chi(f)(v) = \zeta(f)(\xi(f)(v)) = 0$.

Wir sehen so, dass es genügt, die Behauptung $\operatorname{charpol}_f(f)(v) = 0$ in dem speziellen Fall zu zeigen, dass $V$ ein $f$-zyklischer Vektorraum mit Basis $v, f(v), \dots, f^{n-1}(v)$ ist.

Die darstellende Matrix von $f$ bezüglich der Basis $v, f(v), \dots, f^{n-1}(v)$ von $V$ (Lemma 16.17) ist eine Begleitmatrix, genauer die Begleitmatrix des Polynoms $\chi$.

Dann ist $\chi(f)(v) = 0$ aber leicht nachzurechnen. Ist nämlich $\chi = X^n + \sum_{i=0}^{n-1} a_i X^i$, so lesen wir aus der letzten Spalte der Begleitmatrix ab, dass

$$f^n(v) = f(f^{n-1}(v)) = \sum_{i=0}^{n-1} -a_i f^i(v),$$

also

$$\chi(f)(v) = f^n(v) + \sum_{i=0}^{n-1} a_i f^i(v) = 0.$$

□

Es gibt viele andere Möglichkeiten, den Satz zu beweisen, selbst auf der englischen Wikipedia-Seite[4] werden mehrere skizziert.

BEMERKUNG 16.22. Es ist verlockend, die folgende »Rechnung« als einen Beweis des Satzes von Cayley–Hamilton anzusehen:

$$\det(XE_n - A)(A) = \det(AE_n - A) = \det(0) = 0.$$

Das Problem mit diesem »Beweis« (genauer mit dem ersten Gleichheitszeichen) ist, dass das Produkt $XE_n$ durch Einsetzen von $A$ für $X$ *nicht* das Matrizenprodukt $AE_n$ ergibt. In der Tat ist $XE_n$ die Matrix (in $M_n(K[X])$) auf deren Diagonale überall $X$ steht und deren Einträge außerhalb der Diagonalen gleich 0 sind. Setzen wir für alle $X$ nun die Matrix $A$ ein, so erhalten wir eine Matrix mit *Einträgen* in $M_n(K)$, nicht eine Matrix mit Einträgen in $K$ (wie $AE_n$ es ist).

Andere Wege zu sehen, dass man so nicht argumentieren kann, sind die folgenden:

(1) Im Satz von Cayley–Hamilton bedeutet $= 0$, dass der Ausdruck $\operatorname{charpol}_A(A)$ die *Nullmatrix* ist, aber $\det(AE_n - A)$ ist ein *Element des Grundkörpers* $K$!

[4] https://en.wikipedia.org/wiki/Cayley%E2%80%93Hamilton_theorem

(2) Analog zur Determinante können wir auch die Spur einer Matrix mit Einträgen in $K[X]$ definieren. Die Spur ist einfach die Summe aller Diagonaleinträge. Sei $A \in M_n(K)$ und $f = \operatorname{Spur}(XE_n - A) \in K[X]$. Dieselbe Methode würde auch zeigen, dass $f(A) = 0$ ist.
Es gilt aber $f(X) = \operatorname{Spur}(XE_n - A) = nX - \operatorname{Spur}(A)$, und es ist klar, dass im allgemeinen nicht $nA = \operatorname{Spur}(A)E_n$ gilt.

◊

Nach Definition des Minimalpolynoms können wir den Satz von Cayley–Hamilton äquivalent auch als Teilbarkeitsaussage formulieren. So erhalten wir auch die schon angekündigte Abschätzung für den Grad des Minimalpolynoms einer Matrix (bzw. eines Endomorphismus).

KOROLLAR 16.23. *Ist $A \in M_n(K)$, so gilt* $\operatorname{minpol}_A \mid \operatorname{charpol}_A$. *Insbesondere gilt* $\deg(\operatorname{minpol}_A) \leq n$.

Als weiteres Korollar erhalten wir, dass für eine Begleitmatrix charakteristisches Polynom und Minimalpolynom übereinstimmen. Insbesondere sehen wir, dass jedes normierte Polynom vom Grad $n \geq 0$ als charakteristisches Polynom und auch als Minimalpolynom einer $(n \times n)$-Matrix auftreten kann.

KOROLLAR 16.24. *Sei $A \in M_n(K)$ die Begleitmatrix des normierten Polynoms $\chi$ (vom Grad n). Dann gilt* $\operatorname{charpol}_A = \operatorname{minpol}_A = \chi$.

BEWEIS. Wegen des Satzes von Cayley–Hamilton ist $\operatorname{minpol}_A$ ein Teiler von $\operatorname{charpol}_A$, also genügt es zu zeigen, dass $\deg(\operatorname{minpol}_A) \geq n$ ist. Nun ist nach Definition des Begriffs Begleitmatrix $Ae_i = e_{i+1}$ für $i = 1, \ldots, n-1$, und wäre $p = \sum_{i=0}^{m} a_i X^i$ ein Polynom vom Grad $0 \leq m < n$ mit $p(A) = 0$, so wäre auch $p(A)e_1 = 0$, aber es ist

$$p(A)e_1 = a_0 e_1 + a_1 A e_1 + \cdots + a_m A^m e_1 = a_0 e_1 + \cdots + a_m e_{m+1}$$

und $e_1, \ldots, e_{m+1}$ ist eine linear unabhängige Familie. □

**16.3.1. Folgerungen aus dem Satz von Cayley–Hamilton.** Zunächst erlaubt der Satz von Cayley-Hamilton einen Zugang zur konkreten Berechnung des Minimalpolynoms einer Matrix.

BEMERKUNG 16.25 (Berechnung des Minimalpolynoms). Um das Minimalpolynom einer Matrix $A \in M_n(K)$ über einem Körper $K$ zu berechnen, kann man das charakteristische Polynom berechnen. Das erfolgt durch Berechnung der Determinante einer $(n \times n)$-Matrix in $M_n(K[X])$, was lästig sein kann, aber wofür uns mehrere Verfahren zur Verfügung stehen.

Danach sollte man die Zerlegung des charakteristischen Polynoms in irreduzible Polynome im faktoriellen Ring $K[X]$ bestimmen. Hierfür gibt es kein allgemeines Verfahren, aber in konkreten Fällen, insbesondere für nicht zu große Matrizen, ist das in der Regel möglich. (Konkreter: Übungsaufgaben sind so gewählt, dass das machbar ist.)

Danach kann man in alle Teiler des charakteristischen Polynoms die Matrix einsetzen und so den (eindeutig bestimmten) normierten Teiler kleinsten Grades finden, der die Matrix annulliert.

Beim Ausprobieren sollte man noch die Aussage von Satz 16.26 im Hinterkopf haben, der besagt, dass jeder irreduzible Teiler des charakteristischen Polynoms auch das Minimalpolynom teilt. Man muss also nur diejenigen irreduziblen Teiler von $\operatorname{charpol}_A$ untersuchen, die in der Primfaktorzerlegung mit Exponent $> 1$ auftreten, und schauen, ob der Exponent im Minimalpolynom kleiner ist.

Alternativ kann man natürlich das Minimalpolynom finden, indem man eine nicht-trivial Linearkombination der Matrizen $E_n, A, A^2, \ldots, A^d$ mit möglichst kleinem $d$ sucht. (Und der Satz von Cayley-Hamilton garantiert, dass es immer ein $d < n$ gibt, für das das möglich ist.) Das führt auf ein lineares Gleichungssystem, allerdings mit $n^2$ Gleichungen. ◊

Der folgende Satz zeigt eine noch engere Verbindung zwischen charakteristischem Polynom und Minimalpolynom eines Endomorphismus. Es wird uns aber im weiteren Verlauf der Vorlesung meistens genügen, die etwas schwächere Aussage des darauf folgenden Korollars zur Verfügung zu haben, für das wir einen kurzen direkten Beweis erklären. Sie können daher den Beweis des Satzes, wenn Sie möchten, zunächst überspringen.

SATZ 16.26. *Sei $f \in \mathrm{End}_K(V)$, und sei $p \in K[X]$ ein irreduzibles Polynom. Dann sind äquivalent:*

(i) *$p$ ist ein Teiler von* $\mathrm{charpol}_f$,

(ii) *$p$ ist ein Teiler von* $\mathrm{minpol}_f$.

BEWEIS. (i) $\Rightarrow$ (ii). Wir führen Induktion nach $\dim V$. Ist $\dim V \leq 1$, so ist notwendigerweise $\mathrm{charpol}_f = \mathrm{minpol}_f$. Sei nun $\dim V > 1$. Sei $v \in V \setminus \{0\}$ und sei wieder $U = \langle v, f(v), f^2(v), \ldots \rangle$ der $f$-zyklische Untervektorraum, der von den Vektoren $f^i(v)$ erzeugt wird. Sei $g = f_{|U} \in \mathrm{End}_K(U)$ die Einschränkung von $f$.

Sei $W \subseteq V$ ein Komplementärraum von $U$, und sei $\pi\colon V \to W$ die Projektion auf $W$ (d.h. für $v = u + w \in V$ mit $u \in U$, $w \in W$ gelte $\pi(v) = \pi(u + w) = w$). Sei $h \in \mathrm{End}_K(W)$ der Endomorphismus von $W$, der durch $h(w) = \pi(f(w))$ gegeben ist.

Wir sind dann in der Situation von Lemma 16.5, es gilt folglich $\mathrm{charpol}_f = \mathrm{charpol}_g \mathrm{charpol}_h$.

Weil $p$ irreduzibel ist, und damit ein Primelement im Ring $K[X]$, folgt aus unserer Voraussetzung, dass $p \mid \mathrm{charpol}_g$ oder $p \mid \mathrm{charpol}_h$. Im ersten Fall folgt direkt der Satz: Weil $U$ ein $f$-zyklischer Untervektorraum ist, ist nämlich $\mathrm{charpol}_g = \mathrm{minpol}_g$, und weil $\mathrm{minpol}_f(g) = 0$ ist, gilt $\mathrm{minpol}_g \mid \mathrm{minpol}_f$.

Wenn $p \mid \mathrm{charpol}_h$ gilt, dann folgt aus der Induktionsvoraussetzung, dass $p \mid \mathrm{minpol}_h$, und wieder gilt $\mathrm{minpol}_f(h) = 0$, also $\mathrm{minpol}_h \mid \mathrm{minpol}_f$.

Die Implikation (ii) $\Rightarrow$ (i) folgt direkt aus dem Satz von Cayley–Hamilton, der besagt, dass $\mathrm{minpol}_f$ ein Teiler von $\mathrm{charpol}_f$ ist.

*Alternativer Beweis.* Eine ganz andere Möglichkeit, die Richtung (i) $\Rightarrow$ (ii) zu beweisen, liefert das folgende Lemma. Da der Ring $K[X]$ faktoriell ist, ist klar, dass aus dessen Aussage die Implikation (i) $\Rightarrow$ (ii) folgt.

LEMMA 16.27. *Seien $K$ ein Körper, $n \in \mathbb{N}$ und $A \in M_n(K)$. Dann gilt*

$$\mathrm{charpol}_A \mid (\mathrm{minpol}_A)^n.$$

BEWEIS. Der Beweis, den wir geben, ist kurz und auch nicht schwierig, aber insofern »trickreich«, als nicht offensichtlich ist, wie man auf dieses Argument kommen würde.

*Vorbemerkung.* Sei $p \in K[X]$ irgendein Polynom. Wir betrachten den Polynomring $K[X, Y]$ in zwei Unbestimmten $X$ und $Y$. Wenn wir in $p = p(X)$ für $X$ die neue Unbestimmte $Y$ einsetzen, erhalten wir $p(Y) \in K[X, Y]$. Dann gilt im Ring $K[X, Y]$, dass $(X - Y) \mid p(X) - p(Y)$. In der Tat, im Fall $p = X^i$ haben wir

$$X^i - Y^i = (X - Y)(X^{i-1} + X^{i-2}Y + \cdots + XY^{i-2} + Y^{i-1}),$$

wie man unmittelbar nachrechnet. Daraus folgt leicht der allgemeine Fall.

Sei nun zur Abkürzung $\mu = \operatorname{minpol}_A$. Wie in der Vorbemerkung schreiben wir $\mu(X) - \mu(Y) = (X - Y) \cdot p(X, Y)$ für ein Polynom $p(X, Y) \in K[X, Y]$. Wir nutzen diese Umschreibung unten in der Form, dass wir für $X$ die Matrix $XE_n \in M_n(K[X])$ und für $Y$ die Matrix $A \in M_n(K) \subseteq M_n(K[X])$ einsetzen, wir haben dann also

$$\mu(XE_n) - \mu(A) = (XE_n - A)\, B \in M_n(K[X]),$$

wobei $B := p(XE_n, A) \in M_n(K[X])$ ist. (Es genügt uns, dass die obige Gleichung für irgendeine Matrix $B \in M_n(K[X])$ gilt, wir müssen nichts weiter über $B$ wissen.)

Wir können nun wie folgt rechnen:

$$\mu^n = \det(\mu \cdot E_n) = \det(\mu(XE_n) - \mu(A)) = \det((XE_n - A)B) = \operatorname{charpol}_A \cdot \det(B),$$

wobei wir den Produktsatz für die Determinante von Matrizen in $M_n(K[X])$ benutzt haben.

Also ist $\mu^n$ ein Vielfaches von $\operatorname{charpol}_A$, und das ist genau, was wir zeigen wollten. □

□

Korollar 16.28. *Seien $K$ ein Körper, $n \in \mathbb{N}$, $A \in M_n(K)$ und $\lambda \in K$. Dann sind äquivalent:*

(i) *$\lambda$ ist ein Eigenwert von $A$,*

(ii) *$\lambda$ ist eine Nullstelle von* $\operatorname{charpol}_A$,

(iii) *$\lambda$ ist eine Nullstelle von* $\operatorname{minpol}_A$.

Beweis. Die Äquivalenz von (i) und (ii) haben wir bereits bewiesen (Satz 16.7). Die Äquivalenz von (ii) und (iii) ist eine direkte Folgerung aus dem vorherigen Satz, denn $\lambda$ ist genau dann Nullstelle eines Polynoms $p$, wenn $p$ durch das (irreduzible) Polynom $X - \lambda$ teilbar ist. Es ist aber auch leicht, das Korollar direkt zu beweisen.

Dass jede Nullstelle vom Minimalpolynom auch eine Nullstelle des charakteristischen Polynoms ist, folgt aus dem Satz von Cayley–Hamilton.

Sei nun $\lambda \in K$ ein Eigenwert von $A$ und $v \in V$ ein Eigenvektor zum Eigenwert $\lambda$. Es gilt dann $A^i v = \lambda^i v$, und daraus folgt leicht, dass

$$p(A)(v) = p(\lambda)v \quad \text{für alle } p \in K[X]$$

ist.

Insbesondere sehen wir

$$\operatorname{minpol}_A(\lambda)v = \operatorname{minpol}_A(A)v = 0,$$

und da $v$ als Eigenvektor nicht 0 ist, folgt $\operatorname{minpol}_A(\lambda) = 0$. □

Wir können außerdem die Eigenschaften *trigonalisierbar* und *diagonalisierbar* nun in einfacher Weise anhand des Minimalpolynoms charakterisieren. Wir formulieren dieses Ergebnis für Endomorphismen, aber wie immer gilt natürlich die analoge Formulierung für Matrizen.

Korollar 16.29. *Seien $K$ ein Körper und $V$ ein endlichdimensionaler $K$-Vektorraum. Sei $f \in \operatorname{End}_K(V)$. Dann gilt:*

(1) *Der Endomorphismus $f$ ist genau dann trigonalisierbar, wenn* $\operatorname{minpol}_f$ *vollständig in Linearfaktoren zerfällt.*

(2) *Der Endomorphismus $f$ ist genau dann diagonalisierbar, wenn $\operatorname{minpol}_f$ vollständig in Linearfaktoren zerfällt und nur einfache Nullstellen besitzt.*

Wir werden Teil (2) in etwas größerer Allgemeinheit noch einmal im Kapitel über die Jordansche Normalform beweisen (Korollar 17.16); wenn Sie in Eile sind, können Sie den Beweis an dieser Stelle überspringen. Aber vielleicht ist es gerade eine gute Vorbereitung, den Beweis für die hier betrachtete Aussage als Vorbereitung für die spätere Verallgemeinerung durchzugehen. Jedenfalls sollten Sie sich die Aussage des obigen Satzes merken, sie ist oft nützlich.

BEWEIS. Teil (1) folgt aus Satz 16.9 und Satz 16.26, denn letzterer impliziert, dass $\operatorname{minpol}_f$ genau dann vollständig in Linearfaktoren zerfällt, wenn das für $\operatorname{charpol}_f$ gilt.

zu (2). Es ist auch klar, dass für einen diagonalisierbaren Endomorphismus das Minimalpolynom vollständig in Linearfaktoren zerfällt und nur einfache Nullstellen hat. Denn wir können $f$ dann bezüglich einer geeigneten Basis durch eine Diagonalmatrix darstellen und deren Minimalpolynom kann man direkt ablesen. (Vergleiche Beispiel 16.13.)

Nun sei $f$ ein Endomorphismus, dessen Minimalpolynom das Produkt von paarweise verschiedenen Linearfaktoren ist. Wir führen Induktion nach $\dim(V)$, wobei der Fall $\dim(V) \leq 1$ klar ist, da dann jeder Endomorphismus diagonalisierbar ist. Seien $\lambda_1, \ldots, \lambda_r \in K$ die paarweise verschiedenen Nullstellen von $\operatorname{minpol}_f$. Nach Satz 16.26 sind das auch genau die Nullstellen von $\operatorname{charpol}_f$, also die paarweise verschiedenen Eigenwerte von $f$.

Wir schreiben $\operatorname{minpol}_f = (X - \lambda_1)p$ für ein Polynom $p$, das nach Voraussetzung ebenfalls vollständig in Linearfaktoren zerfällt und nur einfache Nullstellen hat, und für das $p(\lambda_1) \neq 0$ gilt. Es sei $U := \operatorname{Ker}(p(f))$. Dann gilt $f(U) \subseteq U$. Wie üblich bezeichnen wir mit $V_{\lambda_1}$ den Eigenraum von $f$ zum Eigenwert $\lambda_1$.

*Behauptung.* Es gilt $V = V_{\lambda_1} \oplus U$.

*Begründung.* Wir zeigen zuerst, dass $V_{\lambda_1} \cap U = 0$ ist. Ist $f(v) = \lambda_1 v$, so folgt $p(f)(v) = p(\lambda_1)v \neq 0$, es sei denn $v = 0$ (denn $p(\lambda_1) \neq 0$, wie oben bemerkt).

Es bleibt zu zeigen, dass $V_{\lambda_1} + U = V$ ist. Weil $X - \lambda_1$ irreduzibel und kein Teiler von $p$ ist, ist 1 ein ggT von $X - \lambda_1$ und $p$ im Hauptidealring $K[X]$, wir können folglich das konstante Polynom $1 \in K[X]$ in der Form $(X - \lambda_1)g + ph = 1$ ausdrücken (für geeignete $g, h \in K[X]$).

Damit sehen wir $v = (f - \lambda_1 \operatorname{id}_V)(g(f)(v)) + p(f)(h(f)(v))$, und dies ist ein Element von $U + V_{\lambda_1}$, weil $0 = \operatorname{minpol}_f(f) = p(f) \circ (f - \lambda_1 \operatorname{id}_V) = (f - \lambda_1 \operatorname{id}_V) \circ p(f)$ gilt.

Nun folgt nach Induktionsvoraussetzung, dass $f_{|U}$ diagonalisierbar ist. Jedenfalls gilt $\dim(U) < \dim(V)$. Außerdem ist $p(f_{|U}) = 0 \in \operatorname{End}_K(U)$, wie direkt aus der Definition von $U$ als $\operatorname{Ker}(p(f))$ folgt. Also gilt $\operatorname{minpol}_{f_{|U}} \mid p$ und deshalb zerfällt $\operatorname{minpol}_{f_{|U}}$ vollständig in Linearfaktoren und hat nur einfache Nullstellen. (Es ist auch nicht schwer zu sehen, dass $\operatorname{minpol}_{f_{|U}} = p$ gilt.)

Es ist andererseits klar, dass $f(V_{\lambda_1}) \subseteq V_{\lambda_1}$ gilt und dass $f_{V_{\lambda_1}}$ diagonalisierbar ist. Es folgt, dass $f$ diagonalisierbar ist. □

## 16.4. Ergänzungen*

BEMERKUNG 16.30. In dieser Bemerkung wird ein anderer Beweis des Satzes von Cayley-Hamilton skizziert, in dem der Satz über den komplexen Zahlen durch ein »Stetigkeitsargument« aus dem Fall von Diagonalmatrizen abgeleitet wird. Um das Argument durchzuführen, werden allerdings Grundkenntnisse der Analysis und Topologie benötigt. Hat man diese Vorkenntnisse zur Verfügung, erhält man so ein schlagendes Argument für den Satz, und

dieses Beweisprinzip der Reduktion auf den Fall von Diagonalmatrizen lässt sich auch an anderer Stelle einsetzen. Mithilfe der sogenannten Zariski-Topologie (nach Oskar Zariski), die in der algebraischen Geometrie eine fundamentale Rolle spielt, lässt sich das Argument auch über einem beliebigen Grundkörper durchführen.

Wie oben bemerkt, ist die Aussage des Satzes von Cayley–Hamilton für Diagonalmatrizen offensichtlich. Weil zueinander konjugierte Matrizen dasselbe charakteristische Polynom und Minimalpolynom haben, folgt der Satz (in der Form, dass das Minimalpolynom das charakteristische Polynom teile) damit für alle diagonalisierbaren Matrizen.

Sei nun zunächst $K = \mathbb{C}$ der Körper der komplexen Zahlen. Wir können dann von stetigen Abbildungen $\mathbb{C}^m \to \mathbb{C}^m$ sprechen und den Satz von Cayley–Hamilton mit dem folgenden »topologischen« Argument beweisen. Die Abbildung

$$M_n(\mathbb{C}) \to M_n(\mathbb{C}), \quad A \mapsto \operatorname{charpol}_A(A),$$

ist eine stetige Abbildung, denn die Einträge der Matrix $\operatorname{charpol}_A(A)$ lassen sich als polynomiale Ausdrücke in den Einträgen von $A$ schreiben, und Polynomfunktionen sind stetig.

Nun gilt für jede stetige Abbildung, dass des Urbild einer abgeschlossenen Teilmenge des Wertebereichs ebenfalls abgeschlossen ist. (Das ist sogar äquivalent zur Stetigkeit.) Angewandt auf die abgeschlossene Teilmenge $\{0\} \subseteq M_n(\mathbb{C})$ sehen wir damit, dass die Teilmenge von $M_n(\mathbb{C})$, die aus allen Matrizen $A$ mit $\operatorname{charpol}_A(A) = 0$ besteht, abgeschlossen ist.

Damit genügt es, die folgende Aussage zu zeigen: Jede abgeschlossene Teilmenge von $M_n(\mathbb{C})$, die alle diagonalisierbaren Matrizen enthält, stimmt mit $M_n(\mathbb{C})$ überein. Mit anderen Worten müssen wir begründen, dass in jedem Ball mit Radius $\varepsilon > 0$ um eine beliebige Matrix stets eine diagonalisierbare Matrix liegt.

Dafür benutzen wir, dass eine Matrix, deren charakteristisches Polynom in $n$ verschiedene Linearfaktoren zerfällt, jedenfalls diagonalisierbar ist. Das folgt – ohne den Satz von Cayley–Hamilton benutzen zu müssen – aus den obigen Ergebnissen. Denn das Minimalpolynom muss dann auch in $n$ verschiedene Linearfaktoren zerfallen, es hat also nur einfache Nullstellen.

Die Bedingung, dass das charakteristische Polynom in $n$ verschiedene Linearfaktoren zerfalle, bedeutet, dass es nur einfache Nullstellen hat (denn über dem algebraisch abgeschlossenen Körper $\mathbb{C}$ zerfällt es jedenfalls vollständig in Linearfaktoren), also dass die Diskriminante $\Delta_{\operatorname{charpol}_A} \in \mathbb{C}$ dieses Polynoms von 0 verschieden ist (Bemerkung 15.84). Die Menge der nicht-diagonalisierbaren Matrizen ist also enthalten in der Menge

$$\{A \in M_n(\mathbb{C});\ \Delta_{\operatorname{charpol}_A} = 0\}.$$

Nun ist auch $\Delta_{\operatorname{charpol}_A}$ ein polynomialer Ausdruck in den Koeffizienten von $A$, und die Nullstellenmenge eines Polynoms $\neq 0$ (in mehreren Variablen, in diesem Fall in den $n^2$ Variablen, die zu den Einträgen der Matrix $A \in M_n(\mathbb{C})$ korrespondieren) kann keinen offenen Ball enthalten. (Dies kann man durch Induktion nach Anzahl der Unbestimmten zeigen.)

Den Fall des Körpers $K = \mathbb{R}$ der reellen Zahlen kann man ähnlich behandeln, wenn man benutzt, dass das charakteristische Polynom eine Matrix $A \in M_n(\mathbb{R})$ davon unabhängig ist, ob man $A$ als Element von $M_n(\mathbb{R})$ oder von $M_n(\mathbb{C})$ betrachtet. ◊

Bemerkung 16.31. Wir können jetzt Bemerkung I.10.18 noch präzisieren: Ist $K$ ein Körper der Charakteristik 0, d.h. dass der eindeutig bestimmte Ringhomomorphismus $\mathbb{Z} \to K$ injektiv ist, und sind $A, B \in M_n(K)$, so sind äquivalent:

(i) Für alle $i \geq 1$ gilt $\operatorname{Spur}(A^i) = \operatorname{Spur}(B^i)$.

(ii) Es gilt $\operatorname{charpol}_A = \operatorname{charpol}_B$.

Insbesondere haben also in dieser Situation $A$ und $B$ dieselben Eigenwerte, und ihre algebraischen Vielfachheiten, also die Vielfachheiten als Nullstelle des charakteristischen Polynoms, stimmen ebenfalls überein. ◊

ERGÄNZUNG 16.32 (Der Fundamentalsatz der Algebra). Von H. Derksen wurde ein Beweis des Fundamentalsatzes der Algebra gegeben, der bis auf die beiden unten angegebenen Fakten (1) und (2) nur lineare Algebra benötigt. Allerdings sind die Beweise, die mit Methoden von fortgeschrittenen Vorlesungen (speziell der Funktionentheorie einerseits und der Algebra andererseits) gegeben werden können, letztlich erhellender, weil die Struktur der Situation insgesamt klarer wird.

THEOREM 16.33 (Fundamentalsatz der Algebra). *Ist $f \in \mathbb{C}[X]$ ein Polynom vom Grad $> 1$, dann besitzt $f$ eine Nullstelle in $\mathbb{C}$.*

Die beiden »analytischen« Eigenschaften, die in Derksens Beweis benötigt werden, sind

(1) Jedes Polynom in $\mathbb{R}[X]$ von ungeradem Grad besitzt eine Nullstelle in $\mathbb{R}$.

(2) Jedes quadratische Polynom in $\mathbb{C}[X]$ hat eine Nullstelle in $\mathbb{C}$.

Den ersten Punkt erhält man aus dem Zwischenwertsatz und der Betrachtung des Grenzwerts der gegebenen Polynomfunktion für $x \to \pm\infty$. Der zweite Punkt folgt (mit einer Methode zur Lösung quadratischer Gleichungen nach Wahl) daraus, dass jede komplexe Zahl eine Quadratwurzel besitzt.

Der Beweis beruht auf einer trickreichen Formulierung, die es erlaubt, an mehreren Stellen mit vollständiger Induktion zu arbeiten.

Siehe https://math.berkeley.edu/~ribet/110/f03/derksen.pdf. □ Ergänzung 16.32

KAPITEL 17

# Die Jordansche Normalform

Der Satz über die Jordansche Normalform besagt, dass jede trigonalisierbare Matrix konjugiert ist zu einer oberen Dreiecksmatrix *einer besonders einfachen Form*, die zudem im wesentlichen eindeutig bestimmt ist, und als die Jordansche Normalform der gegebenen Matrix bezeichnet wird. Sie ist benannt nach dem französischen Mathematiker Camille Jordan[1] (1838 – 1922).

## 17.1. Aussage und Eindeutigkeit

Matrizen in Jordanscher Normalform sind Block-Diagonalmatrizen, und die Blöcke auf der Diagonale sind besonders einfache obere Dreiecksmatrizen, die *Jordan-Blöcke* heißen und folgendermaßen definiert sind.

DEFINITION 17.1. Seien $K$ ein Körper, $\lambda \in K$ und $r \geq 1$. Dann heißt die Matrix

$$J_{r,\lambda} = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \in M_r(K)$$

der *Jordan-Block* der Größe $r \times r$ mit Diagonaleintrag $\lambda$.

Es sind also alle Diagonaleinträge der Matrix gleich $\lambda$, die Einträge auf der Nebendiagonalen direkt oberhalb der Diagonalen sind $= 1$, und alle anderen Einträge der Matrix sind $= 0$. ⊣

Da es sich bei dem Jordan-Block $J_{r,\lambda}$ um eine obere Dreiecksmatrix handelt, ist klar, dass $\lambda$ der einzige Eigenwert von $J_{r,\lambda}$ ist.

Eine besondere Rolle spielen später die Jordan-Blöcke $J_{r,0}$ mit Diagonaleintrag 0. Wie man leicht nachrechnet (Sie sollten das tun!), gilt $J_{r,0}^r = 0$ und $J_{r,0}^i \neq 0$ für $0 \leq i < r$. Alternativ lässt sich das leicht begründen, indem man den zu $J_{r,0}$ gehörigen Endomorphismus von $K^r$ betrachtet.

Damit können wir definieren, was wir unter einer Matrix in Jordanscher Normalform verstehen wollen. Wir benutzen die Schreibweise $\operatorname{diag}(A_1, \ldots, A_r)$ um eine Block-Diagonalmatrix zu bezeichnen.

DEFINITION 17.2. Seien $K$ ein Körper und $n \in \mathbb{N}$. Wir sagen, eine Matrix $A \in M_n(K)$ habe *Jordansche Normalform (JNF)*, falls $r_1, \ldots, r_k \geq 1$ und $\lambda_1, \ldots \lambda_k \in K$ existieren, so dass

$$A = \operatorname{diag}(J_{r_1,\lambda_1}, \ldots, J_{r_k,\lambda_k})$$

ist. ⊣

[1] https://de.wikipedia.org/wiki/Camille_Jordan

Die $\lambda_i$ müssen hier nicht paarweise verschieden sein, sondern derselbe Eigenwert kann in mehreren Blöcken auftreten, und es kann auch mehrere Blöcke derselben Größe zum selben oder zu unterschiedlichen Eigenwerten geben. Zum Beispiel hat jede Diagonalmatrix Jordansche Normalform – dann haben alle Blöcke die Größe $1 \times 1$.

Der Satz über die Jordansche Normalform besagt, dass jede trigonalisierbare Matrix konjugiert ist zu einer Matrix in Jordanscher Normalform, und dass letztere bis auf die Reihenfolge der Blöcke eindeutig bestimmt ist.

THEOREM 17.3 (Jordansche Normalform für Matrizen). *Seien $K$ ein Körper und $n \in \mathbb{N}$. Sei $A \in M_n(K)$ eine trigonalisierbare Matrix. Dann existieren $S \in GL_n(K)$ und $r_1, \ldots, r_k \geq 1$, $\lambda_1, \ldots \lambda_k \in K$, so dass*

$$SAS^{-1} = \operatorname{diag}(J_{r_1,\lambda_1}, \ldots, J_{r_k,\lambda_k})$$

*ist. Dabei ist die Zahl $k$ eindeutig bestimmt (also unabhängig von $S$) und die Paare $(r_1, \lambda_1), \ldots, (r_k, \lambda_k)$ sind eindeutig bestimmt bis auf ihre Reihenfolge.*

Wie wir in Satz 16.9 gesehen haben, ist die Bedingung, dass $A$ trigonalisierbar sei, dazu äquivalent, dass das charakteristische Polynom von $A$ vollständig in Linearfaktoren zerfällt.

Es ist klar, dass die Reihenfolge, in der die Blöcke in der Matrix auftreten, nicht eindeutig bestimmt sind: Wenn sich zwei Block-Diagonalmatrizen $A$ und $B$ nur hinsichtlich der Reihenfolge unterscheiden, in der die Blöcke auf der Diagonalen stehen, aber die Blöcke ansonsten übereinstimmen, dann existiert eine Permutationsmatrix $P$ mit $B = PAP^{-1}$.

Mit dem Beweis dieses Theorems werden wir den überwiegenden Teil dieses Kapitels verbringen. Wir beginnen damit, einige Konsequenzen des Theorems zu beleuchten.

SATZ 17.4. *Seien $K$ ein Körper, $n \in \mathbb{N}$ und sei*

$$A = \operatorname{diag}(J_{r_1,\lambda_1}, \ldots, J_{r_k,\lambda_k}) \in M_n(K)$$

*eine Matrix in Jordanscher Normalform über $K$.*

(1) *Es gilt*

$$\operatorname{charpol}_A = \prod_{i=1}^{k} (X - \lambda_i)^{r_i}.$$

(2) *Wenn $\mu_1, \ldots, \mu_s$ die paarweise verschiedenen Eigenwerte von $A$ und $m_i$ die Größe des größten Jordan-Blocks zu $\mu_i$ bezeichnen, dann ist*

$$\operatorname{minpol}_A = \prod_{i=1}^{s} (X - \mu_i)^{m_i}.$$

BEWEIS. zu (1). Dies ist leicht zu sehen, da die einzelnen Jordan-Blöcke und damit auch jede Matrix in Jordanscher Normalform obere Dreiecksmatrizen sind.

zu (2). Weil für $r > 0$ gilt, dass $J_{r,0}^r = 0$ ist, ist

$$\prod_{i=1}^{s} (A - \mu_i E_n)^{m_i} = 0,$$

also gilt $\operatorname{minpol}_A \mid \prod_{i=1}^{s} (X - \mu_i)^{m_i}$.

Weil $J_{r,0}^{r-1} \neq 0$ ist, und für $\lambda \neq 0$ alle Potenzen von $J_{r,\lambda}$ von Null verschieden sind, folgt, dass kein echter Teiler dieses Produkts die Matrix $A$ annulliert, und das impliziert die behauptete Gleichheit. □

Wie üblich haben wir eine analoge Fassung für Endomorphismen endlichdimensionaler Vektorräume.

THEOREM 17.5 (Jordansche Normalform für Endomorphismen). *Seien $K$ ein Körper und $V$ ein endlichdimensionaler $K$-Vektorraum. Sei $f$ ein trigonalisierbarer Endomorphismus von $V$.*

*Dann existieren eine Basis $\mathscr{B}$ von $V$ und $r_1, \ldots, r_k \geq 1$, $\lambda_1, \ldots \lambda_k \in K$, so dass*

$$M_{\mathscr{B}}^{\mathscr{B}}(f) = \operatorname{diag}(J_{r_1,\lambda_1}, \ldots, J_{r_k,\lambda_k})$$

*ist. Dabei ist die Zahl $k$ eindeutig bestimmt (also unabhängig von der Wahl von $\mathscr{B}$) und die Paare $(r_1, \lambda_1), \ldots, (r_k, \lambda_k)$ sind eindeutig bestimmt bis auf ihre Reihenfolge.*

Es wird nicht behauptet, dass die Basis $\mathscr{B}$ im Satz eindeutig bestimmt sei (und schon das Beispiel der Identitätsabbildung $\mathrm{id}_V$ zeigt, dass es im allgemeinen viele Möglichkeiten gibt, eine solche Basis zu wählen). Eine solche »Jordanbasis« $\mathscr{B}$ zu berechnen ist (möglich, aber meistens) eine ziemlich aufwändige Rechnung. Siehe Ergänzung 17.24.

Der Beweis des Satzes über die Jordansche Normalform ist nicht einfach. Um die einzelnen Schritte zu verstehen, ist es vielleicht nützlich, sich zunächst klarzumachen, dass die behaupteten Aussagen für eine Matrix, die schon Jordansche Normalform hat, »offensichtlich« sind. In diesem Sinne arbeiten wir uns schrittweise vor und beweisen, dass jede trigonalisierbare gewisse Eigenschaften hat, die wir an einer Matrix in Jordanscher Normalform direkt ablesen können.

Etwas konkreter suchen wir (für einen gegebenen trigonalisierbaren Endomorphismus $f$ eines endlichdimensionalen $K$-Vektorraums $V$) eine Basis $\mathscr{B}$, so dass $M_{\mathscr{B}}^{\mathscr{B}}(f)$ in Jordanscher Normalform ist.

- Wenn wir die Basis $\mathscr{B}$ entsprechend der Darstellung von $M_{\mathscr{B}}^{\mathscr{B}}(f)$ als Block-Diagonalmatrix »zerlegen«, entspricht dem eine Zerlegung von $V$ als direkte Summe $f$-invarianter Unterräume. Unser erstes Ziel wird sein, für gegebenes $V$ und $f$ die Existenz einer (im allgemeinen etwas gröberen) Zerlegung $V = \bigoplus_i \tilde{V}_i$ zu zeigen, in der die verschiedenen Eigenwerte von $f$ isoliert sind. Die einzelnen $\tilde{V}_i$ sollen also $f$-invariante Unterräume sein, so dass wir für die Einschränkung $f_{|\tilde{V}_i}$ eine darstellende Matrix finden, die eine obere Dreiecksmatrix ist und auf deren Diagonale überall derselbe Wert steht.

  Diese Zerlegung ist die Zerlegung in »verallgemeinerte Eigenräume«, siehe Abschnitt 17.2.
- Nach diesem ersten Schritt ist es leicht, das Problem auf den Fall eines nilpotenten Endomorphismus (Definition 17.17) zu reduzieren, d.h. wir werden zeigen, dass es genügt, den Fall zu behandeln, dass $f^m = 0$ für ein $m \in \mathbb{N}$ ist.

  Das ist damit gleichbedeutend, dass $f$ durch eine obere Dreiecksmatrix beschrieben werden kann, auf deren Diagonale überall Nullen stehen.

  Es geht dann darum zu zeigen, dass man eine obere Dreiecksmatrix dieser Form immer konjugieren kann zu einer oberen Dreiecksmatrix, die überall Nullen hat mit den Einträgen direkt oberhalb der Diagonale als einziger Ausnahme. Dort dürfen Nullen oder Einsen stehen. Mit anderen Worten: Es ist dann zu zeigen, dass eine Basis $b_1, \ldots, b_n$ von $V$ existiert, so dass jedes $b_i$ entweder auf $b_{i-1}$ oder auf 0 abgebildet wird. Dies ist eine relativ konkrete Fragestellung, die wir in Abschnitt 17.3 behandeln werden.

Die Jordansche Normalform ist ein mächtiges Werkzeug der linearen Algebra. Zum Beispiel erhalten wir aus dem Satz über die Jordansche Normalform zusammen mit Satz 17.4 einen neuen Beweis von Korollar 16.29 im trigonalisierbaren Fall:

KOROLLAR 17.6. *Seien $K$ ein Körper und $f$ ein trigonalisierbarer Endomorphismus eines endlichdimensionalen $K$-Vektorraums. Dann gilt: Der Endomorphismus $f$ ist genau dann diagonalisierbar, wenn sein Minimalpolynom nur einfache Nullstellen hat.*

Dementsprechend ist die Jordansche Normalform auch an vielen Stellen wichtig, wo Methoden der linearen Algebra zur Anwendung kommen. Ein konkretes Beispiel ist die Theorie der linearen Differentialgleichungen mit konstanten Koeffizienten, siehe Abschnitt 17.7.2 für einige weitere Bemerkungen und Verweise dazu.

ERGÄNZUNG 17.7. Man kann zeigen, dass zu jedem Körper $K$ ein algebraisch abgeschlossener Erweiterungskörper $\overline{K}$ existiert. Über diesem ist dann jede Matrix aus $M_n(K)$ trigonalisierbar, besitzt also eine Jordansche Normalform. In dieser werden natürlich im allgemeinen Einträge aus $\overline{K} \setminus K$ auftreten. Dennoch kann das sinnvoll sein, um Aussagen über Matrizen (oder Endomorphismen) über $K$ zu beweisen.

Für $\mathbb{Q}$ und $\mathbb{R}$ kennen wir ja (wenn wir den Fundamentalsatz der Algebra verwenden) einen solchen Erweiterungskörper, nämlich den Körper der komplexen Zahlen. □ Ergänzung 17.7

**17.1.1. Die duale Partition einer Partition.** In diesem Abschnitt führen wir den Begriff der Partition einer natürlichen Zahl ein. Das ist ein einfacher und rein kombinatorischer Begriff, der nützlich ist, um die Eindeutigkeit der Jordanschen Normalform zu zeigen.

DEFINITION 17.8. Ein Tupel $r_1 \geq r_2 \geq r_3 \geq \cdots$ natürlicher Zahlen heißt *Partition* von $n \in \mathbb{N}$, falls $n = \sum_{i \geq 1} r_i$ ist. (Insbesondere dürfen nur endlich viele $r_i \neq 0$ sein.) ⊣

Zu jeder Partition kann man die sogenannte duale Partition bilden.

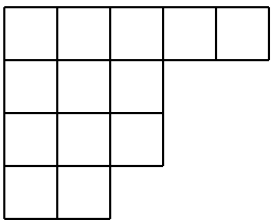
DEFINITION 17.9. Sei $r_1 \geq r_2 \geq r_3 \geq \cdots$ eine Partition von $n$. Dann ist auch $s_1 \geq s_2 \geq \cdots$ mit

$$s_i = \#\{j;\ r_j \geq i\}$$

eine Partition von $n$. Sie wird als die *zu* $(r_i)_i$ *duale Partition* bezeichnet. ⊣

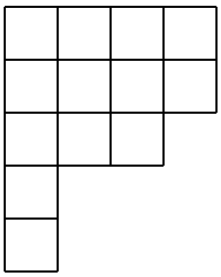
LEMMA 17.10. *Sei $r_1 \geq r_2 \geq r_3 \geq \cdots$ eine Partition von $n$, $s_1 \geq s_2 \geq \cdots$ ihre duale Partition. Dann ist $r_1 \geq r_2 \geq r_3 \geq \cdots$ die duale Partition von $(s_i)_i$.*

BEWEIS. Das ist eine einfache kombinatorische Überlegung, die wir, statt einen Beweis zu geben, nur an dem folgenden konkreten Beispiel illustrieren. □



Die Partition $13 = 5 + 3 + 3 + 2$ kann man durch das nebenstehende Diagramm veranschaulichen. In der ersten Reihe stehen 5 Kästchen, in der zweiten Reihe 3 Kästchen, usw. Insgesamt handelt es sich um 13 Kästchen, und in jeder Reihe sind höchstens so viele Kästchen wie in der Reihe darüber.
Die duale Partition entspricht dann der Partition derselben Zahl 13, die durch das an der Diagonale von links oben nach rechts unten gespiegelte Diagramm beschrieben wird.

Das Diagramm rechts beschreibt die Partition $13 = 4+4+3+1+1$. Das ist genau die zur obigen Partition duale Partition.



In der Kombinatorik hat der Begriff der Partition eine große Bedeutung. Das Problem, für die Anzahl der Partitionen einer gegebenen Zahl $n$ einen geschlossenen Ausdruck anzugeben, ist auch in der Zahlentheorie von einem gewissen Interesse. In der Theorie der Jordanschen Normalform benutzen wir den Begriff allerdings »nur« als ein relativ simples – wenngleich nützliches – Hilfsmittel.

**17.1.2. Eindeutigkeit der Jordanschen Normalform.** Sei $A$ eine Matrix in Jordanscher Normalform. Das charakteristische Polynom von $A$ bestimmt die Diagonaleinträge zusammen mit ihrer Vielfachheit, insbesondere ändern sich diese Daten nicht, wenn $A$ durch eine konjugierte Matrix ersetzt wird. Die Größe der Jordan-Blöcke lässt sich wie folgt beschreiben.

LEMMA 17.11. *Sei $\lambda$ einer der Eigenwerte von $A$, und seien $r_1 \geq r_2 \geq \cdots$ die Größen der Jordan-Blöcke mit Diagonaleintrag $\lambda$. Sei $s_1 \geq s_2 \geq \cdots$ die zu $(r_i)_i$ duale Partition. Dann gilt*

$$s_i = \dim \operatorname{Ker}((A - \lambda E_n)^i) - \dim \operatorname{Ker}((A - \lambda E_n)^{i-1}).$$

BEWEIS. Weil $A$ Jordansche Normalform hat, hat auch $A - \lambda E_n$ Jordansche Normalform. Die Diagonaleinträge sind genau in denjenigen Blöcken gleich 0, die zu Jordan-Blöcken zum Eigenwert $\lambda$ in der Matrix $A$ korrespondieren. Jordan-Blöcke mit einem Diagonaleintrag $\neq 0$ sind invertierbare Matrizen, diese liefern also keinen Beitrag zum Kern.

Andererseits gilt $\operatorname{rg}(J^i_{r,0}) = r - i$ für $i \leq r$ und $\operatorname{rg}(J^i_{r,0}) = 0$ für $i > r$. Das heißt

$$\dim \operatorname{Ker}(J^i_{r,0}) = i \text{ für } i \leq r, \quad \text{und} \quad \dim \operatorname{Ker}(J^i_{r,0}) = r \text{ für } i > r.$$

Damit sehen wir

$$\dim \operatorname{Ker}(J^i_{r,0}) - \dim \operatorname{Ker}(J^{i-1}_{r,0}) = \begin{cases} 1 & \text{falls } i \leq r, \\ 0 & \text{falls } i > r. \end{cases}$$

Folglich ist

$$\dim \operatorname{Ker}((A - \lambda E_n)^i) - \dim \operatorname{Ker}((A - \lambda E_n)^{i-1})$$

die Anzahl der Jordan-Blöcke in $A$ zum Eigenwert $\lambda$, die mindestens die Größe $i$ haben, also mit der Notation in der Aussage des Lemmas die Anzahl der $j \geq 1$, so dass $r_j \geq i$ gilt. Das ist die Definition von $s_i$ im Sinne der dualen Partition. □

Die Zahlen $\dim \operatorname{Ker}(A - \lambda E_n)^i$ ändern sich nicht, wenn man $A$ durch eine zu $A$ konjugierte Matrix ersetzt. Dies beweist, dass die Größen $r_i$ der Jordan-Blöcke in der Jordanschen Normalform einer trigonalisierbaren Matrix eindeutig bestimmt sind, da sie die duale Partition der Partition $(s_i)_i$ wie im Lemma bilden.

Wie immer können wir die Aussage auf trigonalisierbare Endomorphismen eines endlichdimensionalen Vektorraums übertragen: Die kombinatorischen Informationen der Jordanschen Normalform, also die Anzahl und Größe der Jordanblöcke zu den einzelnen Eigenwerten sind eindeutig bestimmt. Es gibt aber in aller Regel viele verschiedene Basen, so dass die darstellende Matrix Jordanform hat. Auch die zu den einzelnen Blöcken auf der Diagonale korrespondierenden Unterräume des zugrundeliegenden Vektorraums sind nicht eindeutig bestimmt. Immerhin ist für jeden Eigenwert $\lambda$ die Summe *aller* Unterräume zu den Jordanblöcken mit Eigenwert $\lambda$ eindeutig bestimmt, wie wir im nächsten Abschnitt sehen werden. Dieser Unterraum ist der sogenannte verallgemeinerte Eigenraum.

## 17.2. Zerlegung in verallgemeinerte Eigenräume

Unser Ziel ist nun, für einen trigonalisierbaren Endomorphismus $f$ eines endlichdimensionalen $K$-Vektorraums $V$ eine Zerlegung von $V$ zu finden, die die Zerlegung in Eigenräume, die wir im diagonalisierbaren Fall haben, verallgemeinert. Wir wollen also die verschiedenen Eigenwerte von $f$ »trennen« und dann die Unterräume (auf denen $f$ nur einen einzigen Eigenwert hat) einzeln behandeln.

Wir wissen, dass die direkte Summe der Eigenräume von $f$ nur dann gleich $V$ ist, wenn $f$ diagonalisierbar ist. Andernfalls müssen wir geeignete größere Untervektorräume von $V$ betrachten als die Eigenräume, und zwar definieren wir zu einem Eigenwert $\lambda \in K$ von $f$ den »verallgemeinerten Eigenraum«.

DEFINITION 17.12. Seien $K$ ein Körper, $V$ ein endlichdimensionaler Vektorraum über $K$ und sei $f \in \mathrm{End}_K(V)$. Sei $\lambda \in K$ ein Eigenwert von $f$. Der Untervektorraum

$$\tilde{V}_\lambda := \tilde{V}_\lambda(f) = \bigcup_{i \geq 0} \mathrm{Ker}(f - \lambda \,\mathrm{id})^i \tag{1}$$

heißt der *verallgemeinerte Eigenraum* (oder *Hauptraum*) von $f$ zum Eigenwert $\lambda$. ⊣

Wir sehen insbesondere, dass der Eigenraum von $f$ zum Eigenwert $\lambda$, das ist $\mathrm{Ker}(f - \lambda \,\mathrm{id})$, im verallgemeinerten Eigenraum enthalten ist. Weil $V$ endlichdimensional ist, ist klar, dass in der aufsteigenden Kette

$$V_\lambda = \mathrm{Ker}(f - \lambda \,\mathrm{id}) \subseteq \mathrm{Ker}(f - \lambda \,\mathrm{id})^2 \subseteq \cdots \subseteq \tilde{V}_\lambda$$

höchstens endlich viele Inklusionen echte Teilmengen sein können. Es gibt also ein $m \in \mathbb{N}$ mit $\tilde{V}_\lambda = \mathrm{Ker}(f - \lambda \,\mathrm{id})^m$. (Wir werden später sehen, dass diese Gleichheit immer schon für $m = \mathrm{mult}_\lambda(\mathrm{minpol}_f)$ richtig ist.)

Der wesentliche Punkt, um die gesuchte Zerlegung zu beweisen, ist das folgende Ergebnis, für das wir nicht vorauszusetzen brauchen, dass $f$ trigonalisierbar ist. Es liefert zu jeder Zerlegung des Minimalpolynoms eines Endomorphismus $f\colon V \to V$ in zueinander teilerfremde Faktoren eine Zerlegung von $V$ in $f$-invariante Unterräume, so dass das Minimalpolynom der Einschränkungen auf die beiden Summanden der jeweilige vorgegebene Faktor ist.

SATZ 17.13. *Seien $K$ ein Körper, $V$ ein endlichdimensionaler $K$-Vektorraum und $f\colon V \to V$ ein Endomorphismus.*

*Sei* $\mathrm{minpol}_f = \zeta \cdot \xi$ *eine Zerlegung in zueinander teilerfremde normierte Polynome* $\zeta, \xi \in K[X]$.

*Dann sind* $U := \mathrm{Ker}(\zeta(f))$ *und* $W := \mathrm{Ker}(\xi(f))$ *invariante Untervektorräume von $V$. Weiter gilt:*

(1) $U = \mathrm{Im}(\xi(f)), \quad W = \mathrm{Im}(\zeta(f))$,

(2) $V = U \oplus W$,

(3) $\mathrm{minpol}_{f_{|U}} = \zeta, \quad \mathrm{minpol}_{f_{|W}} = \xi$.

BEWEIS. Weil $f \circ \zeta(f) = \zeta(f) \circ f$ gilt, ist $U$ ein $f$-invarianter Unterraum. Analog gilt das für $W$.

Wir zeigen nun, dass $U = \mathrm{Im}(\xi(f))$ gilt. Weil $\zeta$ und $\xi$ teilerfremd sind, können wir Polynome $p, q \in K[X]$ mit $p\zeta + q\xi = 1$ finden. Setzen wir in diese Gleichheit $f$ ein, so erhalten wir

$$p(f) \circ \zeta(f) + q(f) \circ \xi(f) = \mathrm{id}_V .$$

Sei nun $u \in U$, das heißt $\zeta(f)(u) = 0$. Dann ist

$$u = p(f)(\zeta(f)(u)) + \xi(f)(q(f)(u)) = \xi(f)(q(f)(u)) \in \mathrm{Im}(\xi(f)).$$

Für die andere Inklusion sei $u \in \mathrm{Im}(\xi(f))$, etwa $u = \xi(f)(v)$. Dann folgt $\zeta(f)(\xi(f)(v)) = \mathrm{minpol}_f(f)(v) = 0$, also $u \in U$.

Entsprechend haben wir $W = \mathrm{Im}(\zeta(f))$, und aus der Dimensionsformel für den Endomorphismus $\zeta(f)$ von $V$ folgt, dass $\dim U + \dim W = \dim V$ ist. Für Teil (2) genügt es folglich, $U \cap W = 0$ zu zeigen.

Sei also $v \in \mathrm{Ker}(\zeta(f)) \cap \mathrm{Ker}(\xi(f))$. Wir haben dann

$$v = p(f)(\zeta(f)(u)) + q(f)(\xi(f)(u)) = 0.$$

Es bleibt Teil (3) zu zeigen. Sicher ist $\zeta(f_{|U})$ die Nullabbildung, denn $U$ wurde ja als der Kern von $\zeta(f)$ definiert. Das bedeutet $\mathrm{minpol}_{f_{|U}} \mid \zeta$. Entsprechend sehen wir $\mathrm{minpol}_{f_{|W}} \mid \xi$.

Außerdem folgt aus der Zerlegung $V = U \oplus W$ (weil $U$ und $W$ invariant unter $f$ sind), dass $\operatorname{minpol}_f \mid \operatorname{minpol}_{f_{|U}} \operatorname{minpol}_{f_{|W}}$ ist: Wir können $f$ entsprechend dieser Zerlegung durch eine Block-Diagonalmatrix darstellen, und $\operatorname{minpol}_{f_{|U}}$ bzw. $\operatorname{minpol}_{f_{|W}}$ annullieren den zu $U$ bzw. $W$ korrespondierenden Block. Also annulliert das Produkt die gesamte Matrix und damit den Endomorphismus $f$, wird also von $\operatorname{minpol}_f$ geteilt.

Wir erhalten damit eine Kette

$$\operatorname{minpol}_f \mid \operatorname{minpol}_{f_{|U}} \cdot \operatorname{minpol}_{f_{|W}} \mid \zeta \cdot \xi = \operatorname{minpol}_f$$

von Teilbarkeitsbeziehungen. Weil links und rechts dasselbe Polynom stehen und alle auftretenden Polynome normiert sind, muss in dieser Kette überall Gleichheit gelten. Wir haben gesehen, dass $\operatorname{minpol}_{f_{|U}} \mid \zeta$ und $\operatorname{minpol}_{f_{|W}} \mid \xi$ gilt; die Gleichheit $\operatorname{minpol}_{f_{|U}} \cdot \operatorname{minpol}_{f_{|W}} = \zeta \cdot \xi$ impliziert daher (wiederum, weil alle Polynome hier normiert sind), dass $\operatorname{minpol}_{f_{|U}} = \zeta$ und $\operatorname{minpol}_{f_{|W}} = \xi$ ist. □

Als nächstes ergänzen wir den Satz um die folgende Präzisierung in dem speziellen Fall, dass $\zeta$ die Potenz eines irreduziblen Polynoms ist.

Satz 17.14. *Seien $K$ ein Körper, $V$ ein endlichdimensionaler $K$-Vektorraum und $f$ ein Endomorphismus von $V$. Sei $\pi \in K[X]$ ein irreduzibles Polynom, das ein Teiler von $\operatorname{minpol}_f$ ist. Wir schreiben $\operatorname{minpol}_f = \pi^m \cdot \xi$ mit $\pi \nmid \xi$. Das bedeutet, dass $m$ die größte natürliche Zahl ist, so dass $\pi^m \mid \operatorname{minpol}_f$ gilt.*

*Es ist dann $\pi$ auch ein Teiler von $\operatorname{charpol}_f$ und wir schreiben $\operatorname{charpol}_f = \pi^{m'} \eta$ mit $\pi \nmid \eta$.*

(1) *Es gilt*

$$\bigcup_{i \geq 1} \operatorname{Ker}(\pi^i(f)) = \operatorname{Ker}(\pi^m(f)) \quad =: U.$$

(2) *Das charakteristische Polynom von $f_{|U}$ ist $\pi^{m'}$, das von $f_{|W}$ ist $\eta$.*

Beweis. Wir wenden Satz 17.13 auf $\zeta := \pi^m$ und $\xi$ an und erhalten insbesondere $\operatorname{Ker}(\pi^m(f)) = \operatorname{Im}(\xi(f))$. Weil für jedes $i \geq 0$ die Elemente $\pi^i$ und $\xi$ teilerfremd sind, zeigt dasselbe Argument wie im Beweis von Satz 17.13, dass $\operatorname{Ker}(\pi^i(f)) \subseteq \operatorname{Im}(\xi(f))$ gilt. Damit folgt Teil (1).

Für Teil (2) benutzen wir, dass die irreduziblen Teiler von Minimalpolynom und charakteristischem Polynom eines Endomorphismus übereinstimmen (also ist $\operatorname{charpol}_{f_{|U}}$ eine Potenz von $\pi$ und $\pi \nmid \operatorname{charpol}_{f_W}$). Außerdem gilt

$$\operatorname{charpol}_f = \operatorname{charpol}_{f_{|U}} \cdot \operatorname{charpol}_{f_{|W}}.$$

Zusammen folgt die Behauptung von Teil (2). □

Für die Jordansche Normalform brauchen wir diese Sätze nur im trigonalisierbaren Fall anzuwenden und Sie sollten sich ihre Aussagen und Beweise (zumindest im ersten Durchgang) mindestens in diesem speziellen Fall klarmachen. Dann hat $\pi$ die Form $X - \lambda$ für ein $\lambda \in K$ und es ist $m = \operatorname{mult}_\lambda(\operatorname{minpol}_f)$, $m' = \operatorname{mult}_\lambda(\operatorname{charpol}_f)$. Wir formulieren für diesen Fall das Ergebnis noch einmal explizit als das folgende Korollar, das eine direkte Übersetzung der beiden obigen Sätze im hier betrachteten Spezialfall ist.

Korollar 17.15. *Seien $K$ ein Körper, $V$ ein endlichdimensionaler $K$-Vektorraum und $f$ ein Endomorphismus von $V$. Sei $\lambda \in K$ ein Eigenwert von $f$ und sei*

$$\tilde{V}_\lambda = \bigcup_{i \geq 1} \operatorname{Ker}((f - \lambda \operatorname{id}_V)^i)$$

*der* verallgemeinerte Eigenraum *von $f$ zum Eigenwert $\lambda$. Wir setzen $m := \mathrm{mult}_\lambda(\mathrm{minpol}_f)$ und schreiben* $\mathrm{minpol}_f = (X-\lambda)^m \cdot \xi$ *(mit $\xi \in K[X]$, $\xi(\lambda) \neq 0$). Dann gilt*

(1) $\tilde{V}_\lambda = \mathrm{Ker}((f-\lambda\,\mathrm{id}_V)^m)$,

(2) $V = \tilde{V}_\lambda \oplus W$, *wobei* $W = \mathrm{Ker}(\xi(f)) = \mathrm{Im}((f-\lambda\,\mathrm{id}_V)^m)$.

(3) *Das Minimalpolynom der Einschränkung von $f$ auf $\tilde{V}_\lambda$ ist $(X-\lambda)^m$. Das charakteristische Polynom dieser Einschränkung ist $(X-\lambda)^{m'}$ mit $m' = \mathrm{mult}_\lambda(\mathrm{charpol}_f)$. Insbesondere ist $\lambda$ der einzige Eigenwert von $f_{|\tilde{V}_\lambda}$ und* $\dim(\tilde{V}_\lambda) = m'$.

(4) *Der Untervektorraum $W$ aus Teil (2) ist $f$-invariant und* $\mathrm{minpol}_{f_{|W}} = \xi$.

Weil $\tilde{V}_\lambda$ jedenfalls den Eigenraum $V_\lambda = \mathrm{Ker}(f-\lambda\,\mathrm{id}_V) \neq 0$ enthält, sehen wir mit Teil (1), dass $m \geq 1$ gelten muss.

Indem wir im trigonalisierbaren Fall, wo charakteristisches Polynom und Minimalpolynom vollständig in Linearfaktoren zerfallen, das Korollar wiederholt auf alle Eigenwerte anwenden, können wir den Unterraum $W$ weiter zerlegen und erhalten induktiv die oben schon angekündigte Zerlegung in verallgemeinerte Eigenräume.

Korollar 17.16 (Zerlegung in verallgemeinerte Eigenräume). *Seien $K$ ein Körper, $V$ ein endlichdimensionaler $K$-Vektorraum und $f$ ein trigonalisierbarer Endomorphismus von $V$. Seien $\lambda_1, \ldots, \lambda_r \in K$ die paarweise verschiedenen Eigenwerte von $K$ und sei für $i = 1, \ldots, r$ mit*

$$\tilde{V}_{\lambda_i} = \bigcup_{j\geq 1} \mathrm{Ker}((f-\lambda_i\,\mathrm{id}_V)^j) = \mathrm{Ker}((f-\lambda_i\,\mathrm{id}_V)^{\mathrm{mult}_{\lambda_i}(\mathrm{minpol}_f)})$$

*der verallgemeinerte Eigenraum von $f$ zum Eigenwert $\lambda_i$ bezeichnet.*

*Dann gilt*

$$V = \tilde{V}_{\lambda_1} \oplus \cdots \oplus \tilde{V}_{\lambda_r}.$$

Insbesondere sehen wir erneut, dass $V$ die direkte Summe der (gewöhnlichen) Eigenräume ist, wenn das Minimalpolynom vollständig in Linearfaktoren zerfällt und nur einfache Nullstellen hat. Wir haben also Korollar 16.29 (2) erneut bewiesen.

Mit diesem Korollar haben wir den ersten Zwischenschritt zum Beweis der Existenz der Jordanschen Normalform für den Endomorphismus $f$ erreicht, denn wir haben den Vektorraum in eine direkte Summe von $f$-invarianten Untervektorräumen zerlegt, die den einzelnen Eigenwerten von $f$ »zugeordnet« sind. Genauer gesagt ist der einzige Eigenwert, den die Einschränkung von $f$ auf $\tilde{V}_{\lambda_i}$ hat, gerade das Element $\lambda_i \in K$, denn dies ist die einzige Nullstelle des Minimalpolynoms von $f_{|\tilde{V}_{\lambda_i}}$.

Überlegen Sie sich, dass das Korollar (im trigonalisierbaren Fall) leicht aus dem Satz über die Jordansche Normalform folgen würde (das ist an dieser Stelle natürlich nur ein Plausibilitätstest, weil wir das obigen Korollar als einen Baustein im Beweis der Existenz der Jordanschen Normalform benutzen möchten.)

## 17.3. Die Jordansche Normalform für nilpotente Endomorphismen

Ist $f$ ein trigonalisierbarer Endomorphismus eines Vektorraums $V$, der nur einen einzigen Eigenwert $\lambda$ hat, dann ist $f-\lambda\,\mathrm{id}_V$ trigonalisierbar mit dem einzigen Eigenwert 0. Den letzteren Fall wollen wir in diesem Abschnitt genauer untersuchen. Unter Ausnutzung der Zerlegung in verallgemeinerte Eigenräume werden wir danach den Beweis des allgemeinen Satzes über die Jordansche Normalform leicht abschließen können.

DEFINITION 17.17. Seien $K$ ein Körper und $V$ ein endlichdimensionaler $K$-Vektorraum, $n = \dim(V)$. Ein Endomorphismus $f \in \mathrm{End}_K(V)$ heißt *nilpotent*, falls die folgenden äquivalenten Bedingungen erfüllt sind:

(i) Es existiert $i \geq 0$, so dass $f^i = 0$.

(ii) $f^n = 0$,

(iii) $\mathrm{minpol}_f \mid X^n$.

(iv) $\mathrm{charpol}_f = X^n$

(v) Es gibt eine Basis $\mathscr{B}$ von $V$, so dass $M^{\mathscr{B}}_{\mathscr{B}}(f)$ eine obere Dreiecksmatrix ist, deren Diagonaleinträge alle $= 0$ sind.

⊣

BEWEIS DER ÄQUIVALENZ. Mit den Sätzen aus dem vorherigen Kapitel sind alle Implikationen leicht zu zeigen. Versuchen Sie es erstmal selbst, bevor Sie den Beweis hier lesen!

Aus dem Satz von Cayley-Hamilton in der Form von Korollar 16.23 – $\mathrm{minpol}_f \mid \mathrm{charpol}_f$ – folgt (iv) $\Rightarrow$ (iii). Die Implikationen (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) sind offensichtlich. Weil das Polynom $X^n$ vollständig in Linearfaktoren zerfällt, folgt aus (iv), dass $f$ trigonalisierbar ist, und damit (v), denn 0 ist die einzige Nullstelle von $X^n$. Dass andererseits (iv) aus (v) folgt, ist klar.

Nun bleibt noch zu begründen, dass (i) $\Rightarrow$ (iv) gilt. Wenn $f^i = 0$ ist, dann muss $\mathrm{minpol}_f$ ein Teiler von $X^i$ sein. Das einzige irreduzible Polynom, das $X^i$ teilt, ist $X$, und weil charakteristisches Polynom und Minimalpolynom von denselben irreduziblen Polynomen geteilt werden (Satz 16.26), muss $\mathrm{charpol}_f$ ebenfalls eine Potenz von $X$ sein. Weil $\deg(\mathrm{charpol}_f) = n$ ist, gilt (iv).

Alternativ kann man Satz 16.26 mit dem folgenden Argument vermeiden, das direkt die Implikation (i) $\Rightarrow$ (v) zeigt. Sei etwa $f^m = 0$. Wir betrachen die Kette

$$0 \subset \mathrm{Ker}(f) \subseteq \mathrm{Ker}(f^2) \subseteq \cdots \subseteq \mathrm{Ker}(f^{m-1}) \subseteq \mathrm{Ker}(f^m) = V.$$

Wir wählen nacheinander Komplementärräume zu diesen Inklusionen, es sei also $U_1$ ein Komplement von $\mathrm{Ker}(f)$ in $\mathrm{Ker}(f^2)$, es sei $U_2$ ein Komplement von $\mathrm{Ker}(f^2) = \mathrm{Ker}(f) \oplus U_1$ in $\mathrm{Ker}(f^3)$, usw., und schließlich $U_{m-1}$ ein Komplement von $\mathrm{Ker}(f^{m-1}) = \mathrm{Ker}(f) \oplus U_1 \oplus \cdots \oplus U_{m-2}$ in $\mathrm{Ker}(f^m) = V$. Schreiben wir noch $U_0 := \mathrm{Ker}(f)$, so erhalten wir eine Zerlegung

$$V = U_0 \oplus U_1 \oplus \cdots \oplus U_{m-1}$$

mit der Eigenschaft, dass für alle $i$ gilt, dass

$$f(U_i) \subseteq U_0 \oplus \cdots \oplus U_{i-1},$$

denn offenbar gilt $f(\mathrm{Ker}(f^{i+1})) \subseteq \mathrm{Ker}(f^i)$. Wählen wir Basen für $U_0, U_1, \ldots, U_{m-1}$ und setzen diese zu einer Basis von $V$ zusammen, so erhalten wir eine Basis $\mathscr{B}$ von $V$, so dass $M^{\mathscr{B}}_{\mathscr{B}}(f)$ eine Block-obere-Dreiecksmatrix ist, deren Diagonalblöcke gleich Null sind. Insbesondere ist $M^{\mathscr{B}}_{\mathscr{B}}(f)$ eine obere Dreiecksmatrix mit Nullen auf der Diagonale. □

Vergleiche auch Satz I.6.56, wo wir die Folgerung (i) $\Rightarrow$ (ii) mit einem ähnlichen Argument wie am Ende des Beweises gezeigt haben. Mithilfe des Satzes von Cayley-Hamilton haben wir nun einen neuen Beweis erhalten.

Analog definieren wir den Begriff der nilpotenten Matrix; dort ist natürlich eine entsprechende Charakterisierung durch zueinander äquivalente Bedingungen möglich.

DEFINITION 17.18. Seien $K$ ein Körper und $n \in \mathbb{N}$. Eine Matrix $A \in M_n(K)$ heißt *nilpotent*, falls die folgenden äquivalenten Bedingungen erfüllt sind:

(i) Es existiert $i$, so dass $A^i = 0$.

(ii) $A^n = 0$,

(iii) $\text{minpol}_A \mid X^n$.

(iv) $\text{charpol}_A = X^n$

(v) Die Matrix $A$ ist konjugiert zu einer oberen Dreiecksmatrix, deren Diagonaleinträge alle $= 0$ sind.

⊣

LEMMA 17.19. *Seien $K$ ein Körper und $V$ ein endlichdimensionaler $K$-Vektorraum. Seien $f_1, f_2 \in \text{End}_K(V)$ Endomorphismen mit $f_1 \circ f_2 = f_2 \circ f_1$.*

(1) *Sind $f_1$ und $f_2$ diagonalisierbar, so existiert eine Basis $\mathscr{B}$ von $V$, so dass sowohl $M^{\mathscr{B}}_{\mathscr{B}}(f_1)$ als auch $M^{\mathscr{B}}_{\mathscr{B}}(f_2)$ Diagonalmatrizen sind. Wir sagen, $f_1$ und $f_2$ seien* simultan diagonalisierbar.

(2) *Sind $f_1$ und $f_2$ diagonalisierbar, so ist $f_1 + f_2$ diagonalisierbar.*

(3) *Sind $f_1$ und $f_2$ nilpotent, so ist $f_1 + f_2$ nilpotent.*

BEWEIS. zu (1). Übung (Hausaufgabe 3.4). Teil (2) folgt leicht aus Teil (1), weil die Summe von Diagonalmatrizen offenbar eine Diagonalmatrix ist.

Für Teil (3) skizzieren wie zwei Beweismöglichkeiten. Eine Möglichkeit ist, den *binomischen Lehrsatz* zu verwenden, und zwar im kommutativen Ring

$$K[f_1, f_2] = \left\{ \sum_{i,j \geq 0} a_{ij} f_1^i f_2^j \,;\, a_{ij} \in K, \text{ nur endlich viele } a_{ij} \neq 0 \right\} \qquad (\subseteq \text{End}_K(V)).$$

LEMMA 17.20 (Binomischer Lehrsatz). *Sei $R$ ein kommutativer Ring und seien $x, y \in R$ und $n \in \mathbb{N}$. Dann gilt*

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}.$$

Sind dann $n_1, n_2 \in \mathbb{N}$ mit $f_i^{n_i} = 0, i = 1, 2$, so liefert der binomische Lehrsatz eine Darstellung von $(f_1 + f_2)^{n_1+n_2}$ als Summe, in der in jedem Summanden entweder der Exponent von $f_1$ mindestens $n_1$, oder der Exponent von $f_2$ mindestens $n_2$ ist. Also ist jeder Summand $= 0$.

*Alternativer Beweis.* Eine andere Möglichkeit ist, die Aussage von Hausaufgabe 4.3 zu benutzen, dass miteiander kommutierende trigonalisierbare Endomorphismen simultan trigonalisierbar sind. Weil $f_1 \circ f_2 = f_2 \circ f_1$ gilt und weil jeder nilpotente Endomorphismus trigonalisierbar ist, folgt also, dass bezüglich einer geeigneten Basis sowohl $f_1$ als auch $f_2$ durch eine obere Dreiecksmatrix dargestellt werden. Weil beide nur den Eigenwert 0 haben, sind alle Diagonaleinträge gleich 0, und folglich gilt das auch für die Summe dieser beiden Matrizen. Es folgt, dass $f_1 + f_2$ ebenfalls nilpotent ist. □

Wir benutzen unten die folgende präzisere Fassung von Lemma 16.17 für den Fall eines nilpotenten Endomorphismus.

LEMMA 17.21. *Sei $f \in \text{End}_K(V)$ ein nilpotenter Endomorphismus und sei $U = \langle u, f(u), \dots \rangle$ ein zyklischer Unterraum. Dann ist* $\dim U = \min\{m; f^m(u) = 0\}$. *Ist $u' \in U \setminus f(U)$, so gilt*

$$U = \langle u', f(u'), f^2(u'), \dots \rangle.$$

BEWEIS. Sei $d$ definiert als $\min\{m; f^m(u) = 0\}$. Wir zeigen, dass $u, f(u), ..., f^{d-1}(u)$ eine linear unabhängige Familie ist. Weil sie (wegen $f^d(u) = 0$) offenbar den Raum $U$ erzeugt, folgt daraus $\dim(U) = d$. Angenommen, es gäbe eine nicht-triviale Linearkombination

$$a_0 u + a_1 f(u) + \cdots a_{d-1} f^{d-1}(u) = 0.$$

Sei $i$ minimal mit $a_i \neq 0$. Wir wenden $f^{d-i-1}$ an und erhalten

$$a_i f^{d-1}(u) = 0,$$

und das ist ein Widerspruch.

(Alternativ kann man auch Lemma 16.17 verwenden.)

Wir sehen (zum Beispiel an der Form der darstellenden Matrix oder durch eine direkte Betrachtung der gewählten Basis), dass $U = \langle u \rangle \oplus \operatorname{Im}(f)$ gilt. Jeder Vektor $u' \in U \setminus f(U)$ lässt sich also schreiben als $\alpha u + f(v)$ mit $\alpha \in K^\times$, $v \in U$. Es folgt $f^{d-1}(u') = f^{d-1}(u) + f^d(v) = f^{d-1}(u) \neq 0$, und mit dem ersten Teil, nun angewandt auf $\langle u', f(u'), \dots \rangle$, dass $\dim\langle u', f(u'), \dots \rangle = d$ gilt. Also ist dieser Unterraum, wie behauptet, gleich $U$.

(Wenn man die Basis stattdessen in der Reihenfolge $\mathscr{B} = (f^{d-1}(u), ..., u)$ schreibt, hat die Begleitmatrix die Form eines Jordan-Blocks: $M^{\mathscr{B}}_{\mathscr{B}}(f) = J_{d,0}$.) □

Nach diesen Vorbereitungen können wir die Existenz der Jordanschen Normalform für nilpotente Endomorphismen beweisen.

SATZ 17.22 (Normalform für nilpotente Endomorphismen/Matrizen). *Es sei $f$ ein nilpotenter Endomorphismus von $V$. Dann existieren eine Basis $\mathscr{B}$ von $V$ und natürliche Zahlen $r_1 \geq \cdots \geq r_k \geq 1$, so dass*

$$M^{\mathscr{B}}_{\mathscr{B}}(f) = \operatorname{diag}(J_{r_1,0}, ..., J_{r_k,0}).$$

Eine entsprechende Aussage gilt für nilpotente Matrizen in $M_n(K)$.

Wegen der Eindeutigkeitsaussage im Satz über die Jordansche Normalform, die wir bereits bewiesen haben, sind die $r_i$ eindeutig bestimmt. Wir haben auch bereits gesehen, dass die Anzahl der Jordan-Blöcke gleich der Dimension des Eigenraums von $f$ zum Eigenwert 0, also gleich $\dim \operatorname{Ker} f$ sein muss.

BEWEIS. Wir haben im Beweis des vorherigen Lemmas bemerkt, dass für jeden $f$-zyklischen Unterraum die darstellende Matrix einer geeigneten Basis die Form eines Jordan-Blocks hat. Deshalb ist die Aussage des Satzes äquivalent dazu, dass $V$ die direkte Summe von $f$-zyklischen Unterräumen ist.

Wir zeigen das durch Induktion nach $n = \dim V$. Für $n = 1$ ist nichts zu zeigen, sei also nun $n > 1$. Sei $U \subseteq V$ ein Unterraum der Dimension $n - 1$ mit $\operatorname{Im} f \subseteq U$. Ein solcher Unterraum existiert, weil $f$ nilpotent ist und daher nicht surjektiv sein kann. Es gilt dann $f(U) \subseteq U$ und wir können die Induktionsvoraussetzung auf $f_{|U}$ anwenden. Wir erhalten so eine Zerlegung $U = U_1 \oplus \cdots U_l$ als direkte Summe $f$-zyklischer Unterräume.

Sei nun $v \in V \setminus U$. Wir schreiben

$$f(v) = \sum_{i=1}^{l} u_i, \qquad \text{mit } u_i \in U_i \, i = 1, ..., l.$$

Für die $i$, für die $u_i \in f(U_i)$ liegt, sagen wir $u_i = f(u_i')$, $u_i' \in U_i$, ersetzen wir nun $v$ durch $v - u_i'$, und ersetzen $u_i$ durch 0. Die obige Gleichung ist dann immer noch richtig. Wir erhalten am Ende einen Vektor $v \in V \setminus U$ mit einer Darstellung

$$f(v) = \sum_{i=1}^{l} u_i, \qquad \text{mit } u_i \in U_i \, i = 1, ..., l,$$

so dass für alle $i$ gilt: $u_i = 0$ oder $u_i \notin f(U_i)$.

*1. Fall:* $f(v) = 0$. Dann ist $\langle v \rangle$ ein $f$-zyklischer Untervektorraum und folglich

$$V = \langle v \rangle \oplus U_1 \oplus \cdots \oplus U_\ell$$

eine Zerlegung von $V$ in $f$-zyklische Unterräume und wir sind fertig.

*2. Fall:* $f(v) \neq 0$. Sei $m$ minimal mit der Eigenschaft, dass $f^{m+1}(v) = f^m(f(v)) = 0$ gilt. Weil die Untervektorräume $U_i$ eine direkte Summe bilden, gilt dann auch $f^m(u_i) = 0$ für alle $i$. Aber für mindestens eines der $u_i$ muss $f^{m-1}(u_i) \neq 0$ sein. Nach Umnummerieren der $U_i$ (und entsprechend der $u_i$) können wir annehmen, dass $m$ auch minimal ist mit $f^m(u_1) = 0$. Wegen $f(v) \neq 0$ ist dann $u_1 \neq 0$, nach unserer Vorüberlegung also $u_1 \notin f(U_1)$. Wir wenden nun Lemma 17.21 an. Weil $U_1$ ein $f$-zyklischer Unterraum ist, folgt $U_1 = \langle u_1, f(u_1), \ldots, f^{m-1}(u_1) \rangle$ und $\dim U_1 = m$. Andererseits hat $W := \langle v, f(v), \ldots, f^m(v) \rangle$ die Dimension $m+1$.

*Behauptung.* $V = W \oplus U_2 \oplus \cdots \oplus U_\ell$.

*Begründung.* Da $\dim V = \dim W + \dim \sum_{i>1} U_i$ ist, genügt es zu zeigen, dass

$$W \cap (U_2 \oplus \cdots \oplus U_\ell) = 0.$$

Nehmen wir also an, dass $a_j \in K$ sind mit

$$\sum_{j=0}^{m} a_j f^j(v) \in U_2 \oplus \cdots \oplus U_\ell.$$

Weil $v \notin U$ aber $\operatorname{Im}(f) \subseteq U$ gilt, muss $a_0 = 0$ sein. Indem wir wieder die Darstellung $f(v) = \sum u_i$ hernehmen, können wir die Summe umschreiben als

$$\sum_{j=0}^{m} a_j f^j(v) = \sum_{j=1}^{m} a_j f^j(v) = \sum_{i=1}^{l} \sum_{j=0}^{m-1} a_{j+1} f^j(u_i).$$

Jetzt nutzen wir noch aus, dass die ganze Summe in $U_2 \oplus \cdots \oplus U_\ell$ liegt, und jeder einzelne Summand zum Index $i$ in $U_i$ enthalten ist, und erhalten

$$\sum_{j=0}^{m-1} a_{j+1} f^j(u_1) \in U_1 \cap (U_2 \oplus \cdots \oplus U_\ell) = 0,$$

und das impliziert $a_1 = \cdots = a_m = 0$, weil $u_1, \ldots, f^{m-1}(u_1)$ linear unabhängig sind. □

## 17.4. Beweis des Satzes über die Jordansche Normalform

Beweis von Theorem 17.3. Die Eindeutigkeitsaussage haben wir bereits in Abschnitt 17.1.2 bewiesen. Ist $f \in \operatorname{End}_K(V)$ gegeben, so zerlegen wir $V = \bigoplus_{i=1}^{r} \tilde{V}_{\lambda_i}$ in die direkte Summe der verallgemeinerten Eigenräume zu den paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r$ von $f$, siehe Korollar 17.16.

Wir wählen mit Hilfe von Satz 17.22 Basen der $\tilde{V}_{\lambda_i}$, so dass der nilpotente Endomorphismus $f_{|\tilde{V}_i} - \lambda_i \operatorname{id}_{\tilde{V}_i}$ von $\tilde{V}_i$ durch eine Matrix in Jordanscher Normalform beschrieben wird. Indem wir alle diese Basen zusammensetzen, erhalten wir eine Basis von $V$, bezüglich derer $f$ durch eine Matrix in Jordanscher Normalform beschrieben wird. □

Bemerkung 17.23 (Berechnung der Jordanschen Normalform einer Matrix/eines Endomorphismus). Um die Jordansche Normalform eines trigonalisierbaren Endomorphismus (bzw. einer Matrix) zu finden, genügt es, die Dimensionen $\dim \operatorname{Ker}(f - \lambda \operatorname{id})^i$ für alle Eigenwerte $\lambda$ und alle $i$ zwischen 1 und $\operatorname{mult}_\lambda(\operatorname{minpol}_f)$ zu berechnen, was man mit (mehrfacher …) Anwendung des Gauß-Algorithmus erledigen kann. Daraus findet man, wie der Eindeutigkeitsbeweis zeigt, die Jordansche Normalform. Oft kann man einen Großteil dieser

Berechnungen sparen, wenn man zunächst das charakteristische Polynom und das Minimalpolynom berechnet und in Linearfaktoren zerlegt, weil das gewisse Einschränkungen an die Jordansche Normalform mit sich bringt, siehe Satz 17.4. ◊

ERGÄNZUNG 17.24 (Berechnung einer Jordanbasis). Sei $f$ ein trigonalisierbarer Endomorphismus eines endlichdimensionalen $K$-Vektorraums $V$. Eine Basis $\mathscr{B}$ von $V$ zu finden, so dass $M_{\mathscr{B}}^{\mathscr{B}}(f)$ Jordansche Normalform hat (eine sogenannte *Jordanbasis*), ist in der Regel wesentlich aufwändiger, als diese Jordansche Normalform zu berechnen. (Wir haben ja bereits gesehen, dass es im allgemeinen Fall eine umfangreiche Rechnung erfordert, um für zueinander konjugierte Matrizen $A$ und $B$ eine invertierbare Matrix $S$ mit $B = SAS^{-1}$ zu finden.)

Als erstes berechnet man die verallgemeinerten Eigenräume von $f$. Auch das kann schon rechenintensiv sein, aber im Prinzip ist klar, wie vorzugehen ist. Danach kann man sich auf den Fall beschränken, dass $V$ ein einziger verallgemeinerter Eigenraum ist, etwa zum Eigenwert $\lambda$. Ersetzt man $f$ durch $f - \lambda \operatorname{id}_V$, so hat man die Aufgabe auf den Fall eines nilpotenten Endomorphismus reduziert.

Im Prinzip könnte man wie im Beweis von Satz 17.22 vorgehen, um die gesuchte Basis zu konstruieren. Um die Berechnung einigermaßen effizient auszuführen, ist es aber besser, die Sache etwas systematischer anzugehen. Das klärt die Situation vielleicht auch zusätzlich auf (allerdings ist die zusätzlich erforderliche »Buchhaltung« etwas lästig, weswegen wir für Satz 17.22 einen kürzeren Beweis gewählt haben).

Es sei also $f\colon V \to V$ nilpotent, etwa $f^m = 0, f^{m-1} \neq 0$. Wir betrachten die folgende Kette von Untervektorräumen von $V$:

$$0 \subseteq \operatorname{Ker}(f) \subseteq \operatorname{Ker}(f^2) \subseteq \cdots \subseteq \operatorname{Ker}(f^{m-1}) \subseteq \operatorname{Ker}(f^m) = V.$$

Wir wählen nun nacheinander

- ein Komplement $U_{m-1}$ von $\operatorname{Ker}(f^{m-1})$ in $\operatorname{Ker}(f^m) = V$,
- ein Komplement $U_{m-2}$ von $f(U_{m-1}) \oplus \operatorname{Ker}(f^{m-2})$ in $\operatorname{Ker}(f^{m-1})$,
- ein Komplement $U_{m-3}$ von $f^2(U_{m-1}) \oplus f(U_{m-2}) \oplus \operatorname{Ker}(f^{m-3})$ in $\operatorname{Ker}(f^{m-2})$,
- ...
- ein Komplement $U_0$ von $f^{m-1}(U_{m-1}) \oplus f^{m-2}(U_{m-2}) \oplus \cdots \oplus f(U_1)$ in $\operatorname{Ker}(f)$.

Für alle $i = 0, \ldots, m-1$ sei $u_1^{(i)}, \ldots, u_{d_i}^{(i)}$ eine Basis von $U_i$. Dann bilden die Vektoren

$$f^j(u_k^{(i)}), \quad i = 0, \ldots, m-1,\ k = 1, \ldots, d_i,\ j = 0, \ldots, i,$$

eine Basis $\mathscr{B}$ von $V$, so dass $M_{\mathscr{B}}^{\mathscr{B}}(f)$ Jordansche Normalform hat. Dabei ordnen wir die Basisvektoren so an, dass für jedes $i$ und $k$ die Vektoren $u_k^{(i)}, f(u_k^{(i)}), \ldots, f^i(u_k^{(i)})$ direkt hintereinander stehen. Sortiert man noch nach $i$, so kann man zusätzlich erreichen, dass die Jordan-Blöcke der Größe nach geordnet sind.

Dass diese Familie von Vektoren eine Basis bildet, ist mit Blick auf die Konstruktion der $U_i$ ohne größere Schwierigkeiten einzusehen.

Dass die Summe $f^j(U_{m-1}) + f^{j-1}(U_{m-2}) + \cdots + f(U_{m-j}) + \operatorname{Ker}(f^{m-j-1})$ in der obigen Konstruktion in $\operatorname{Ker}(f^{m-j})$ enthalten ist, folgt aus $f^m = 0$ und der Konstruktion der $U_i$. Es bleibt aber noch zu begründen, dass diese Summe

$$f^j(U_{m-1}) + f^{j-1}(U_{m-2}) + \cdots + f(U_{m-j}) + \operatorname{Ker}(f^{m-j-1})$$

in jedem der obigen Schritte eine *direkte* Summe ist. Dafür wollen wir zum Abschluss ein Argument skizzieren. Wir führen Induktion nach $j$. Für $j = 1$ ist die Sache klar. Nehmen wir nun an, dass

$$f^j(u_{m-1}) + \cdots + f(u_{m-j}) + v = 0$$

ist mit $j > 1$, $u_i \in U_i$, $v \in \operatorname{Ker}(f^{m-j-1})$. Wir wollen zeigen, dass alle einzelnen Summanden $= 0$ sind. Durch Anwenden von $f^{m-j-1}$ erhalten wir

$$f^{m-1}(u_{m-1}) + \cdots + f^{m-j}(u_{m-j}) = 0.$$

Wenn wir zeigen können, dass daraus $u_i = 0$ für alle $i = m-j, \ldots, m-1$ folgt, dann sind wir fertig.

Wir schreiben die obige Gleichung um als

$$f^{m-j}(f^{j-1}(u_{m-1}) + \cdots + u_{m-j}) = f^{m-1}(u_{m-1}) + \cdots + f^{m-j}(u_{m-j}) = 0.$$

Nach Induktionsvoraussetzung ist $f^{j-1}(U_{m-1}) \oplus f^{j-2}(U_{m-2}) \oplus \cdots \oplus U_{m-j} \oplus \operatorname{Ker}(f^{m-j})$ eine direkte Summe. Dass das Element $f^{j-1}(u_{m-1}) + \cdots + u_{m-j}$ in $\operatorname{Ker}(f^{m-j})$ liegt, wie wir hier sehen, impliziert also $f^{j-1}(u_{m-1}) + \cdots + u_{m-j} = 0$, und damit $f^{j-1}(u_{m-1}) = \cdots = u_{m-j} = 0$.

Nun gilt $U_i \cap \operatorname{Ker}(f^i) = 0$ nach Konstruktion von $U_i$, und daher erhalten wir schließlich $u_{m-1} = \cdots = u_{m-j} = 0$.

□ Ergänzung 17.24

## 17.5. Die Jordan-Zerlegung

Oft ist es ausreichend, anstelle der genauen Jordanschen Normalform die sogenannte Jordan-Zerlegung zur Verfügung zu haben, die es erlaubt, eine trigonalisierbare Matrix als die Summe einer diagonalisierbaren und einer nilpotenten Matrix zu schreiben, die zudem miteinander kommutieren. Besonders nützlich ist die Aussage des folgenden Satzes wegen der Eindeutigkeit dieser Zerlegung.

SATZ 17.25 (Jordan-Zerlegung). *Seien $K$ ein Körper und $V$ ein endlichdimensionaler $K$-Vektorraum. Sei $f \in \operatorname{End}(V)$ ein trigonalisierbarer Endomorphismus.*

*Dann existieren eindeutig bestimmte Endomorphismen $D$ und $N$ von $V$ mit den folgenden Eigenschaften: $D$ ist diagonalisierbar, $N$ ist nilpotent,*

$$f = D + N, \qquad \text{und} \qquad DN = ND.$$

*Ferner existieren Polynome $p_d, p_n \in K[X]$ mit Absolutterm 0, so dass $D = p_d(f)$, $N = p_n(f)$.*

BEWEIS. Seien $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von $f$. Sei $V = \bigoplus_{i=1}^r \tilde{V}_{\lambda_i}$ die Zerlegung in verallgemeinerte Eigenräume und $\operatorname{charpol}_f = \prod_{i=1}^r (X - \lambda_i)^{n_i}$. Mit dem Chinesischen Restsatz, Satz 15.61, finden wir ein Polynom $p_d$, so dass

$$p_d \equiv \lambda_i \mod (X - \lambda_i)^{n_i},\ i = 1, \ldots, s, \quad p_d \equiv 0 \mod X.$$

Man beachte, dass die letzte Bedingung aus den vorherigen folgt, falls 0 ein Eigenwert von $f$ ist, und dass ansonsten $X$ mit allen $(X - \lambda_i)^{n_i}$ teilerfremd ist.

Dann gilt $p_d(f)_{|\tilde{V}_{\lambda_i}} = \lambda_i \operatorname{id}$ für alle $i$, also ist $D := p_d(f)$ diagonalisierbar. Andererseits sei $p_n := X - p_d$ und $N := p_n(f)$. Dann hat $N_{|\tilde{V}_{\lambda_i}}$ nur den Eigenwert 0, ist daher nilpotent, also ist $N$ nilpotent. Offenbar gilt $DN = ND$, da sich $D$ und $N$ als Polynome in $f$ ausdrücken lassen.

*Eindeutigkeit.* Sei $f = D + N$ die soeben konstruierte Zerlegung und $f = D' + N'$ eine weitere. Wir zeigen $D = D'$, $N = N'$. Auch wenn wir nicht voraussetzen, dass sich $D'$ und $N'$ als Polynome in $f$ schreiben lassen, gilt das, wie wir gesehen haben, für $D$ und $N$ und es folgt, dass $f, D, N, D', N'$ alle miteinander kommutieren. Insbesondere ist in der Gleichung

$$D - D' = N' - N$$

die linke Seite diagonalisierbar und die rechte Seite nilpotent. Es folgt $D - D' = 0 = N' - N$, wie gewünscht. □

BEMERKUNG 17.26. Um den Satz über die Jordan-Zerlegung ohne die Eindeutigkeitsaussage und ohne die Aussage, dass sich $D$ und $N$ als Polynome in $f$ ausdrücken lassen, zu beweisen, kann man elementarer vorgehen: Man definiere $D$ als die eindeutig bestimmte Abbildung mit $D_{|\tilde{V}_{\lambda_i}} = \lambda_i \operatorname{id}_{\tilde{V}_{\lambda_i}}$ und setze $N = f - D$. Es lässt sich dann leicht prüfen, dass $D$ diagonalisierbar und $N$ nilpotent ist und dass $DN = ND$ gilt.

Alternativ kann man natürlich auch benutzen, dass $A$ zu einer Matrix $B$ in Jordanscher Normalform konjugiert ist. Wie sieht die Jordan-Zerlegung für $B$ aus? ◊

Ein entsprechendes Ergebnis hat man natürlich für trigonalisierbare Matrizen. Eine andere Variante, die manchmal nützlich ist, ist die multiplikative Jordan-Zerlegung.

ERGÄNZUNG 17.27 (Die multiplikative Jordan-Zerlegung). Sei $K$ ein Körper und $V$ ein endlichdimensionaler Vektorraum über $K$. Ist $f\colon V \to V$ ein trigonalisierbarer Automorphismus von $V$, dann existieren eindeutig bestimmte Automorphismen $U$ und $D$ von $V$, so dass gilt:

(a) $D$ ist diagonalisierbar,

(b) $U$ ist trigonalisierbar mit 1 als einzigem Eigenwert,

(c) $A = U \circ D = D \circ U$.

BEWEIS. Übung. (Das Ergebnis lässt sich leicht aus der additiven Jordan-Zerlegung folgern.) □

□ Ergänzung 17.27

## 17.6. Die rationale Normalform *

Wenn der Grundkörper $K$ nicht algebraisch abgeschlossen ist, dann ist nicht jede quadratische Matrix über $K$ trigonalisierbar. In diesem Fall ist es nützlich, andere »Normalformen« als die Jordansche Normalform zu betrachten. Es ist klar, dass diese im allgemeinen keine Dreiecksform haben können. Es ist aber immer möglich, eine gegebene Matrix zu einer Block-Diagonalmatrix zu konjugieren, deren Blöcke Begleitmatrizen sind. Wir wollen das in diesem Abschnitt ein bisschen präzisieren, aber nicht beweisen.

Wir sprechen wieder über Endomorphismen statt über Matrizen. Sei also $V$ ein endlichdimensionaler $K$-Vektorraum und $f\colon V \to V$ ein Endomorphismus. Sei $\mu = \operatorname{minpol}_f$, $\chi = \operatorname{charpol}_f$ und seien

$$\mu = p_1^{m_1} \cdot \dots \cdot p_r^{m_r}$$

und

$$\chi = p_1^{n_1} \cdot \dots \cdot p_r^{n_r}$$

die Zerlegungen in irreduzible Polynome in $K[X]$, d.h. es seien $p_1, \dots, p_r$ normiert, irreduzibel und paarweise verschieden und $1 \leq m_i \leq n_i$ für alle $i$. (Wir benutzen hier, dass ein irreduzibles Polynom genau dann $\mu$ teilt, wenn es $\chi$ teilt.

Mit Satz 17.13 und Satz 17.14 erhalten wir eine Zerlegung von $V$ als direkte Summe $V = \bigoplus_{i=1}^{r} V_i$ von $f$-invarianten Unterräumen, so dass für alle $i$ die Einschränkung von $f$ auf $V_i$ Minimalpolynom $p_i^{m_i}$ und charakteristisches Polynom $p_i^{n_i}$ hat. Insbesondere gilt $\dim V_i = \deg(p_i^{n_i}) = n_i \deg(p_i)$. Um diese Zerlegung zu erhalten, ist es nicht erforderlich, weitere Wahlen zu treffen, sie ist durch $f$ eindeutig bestimmt.

Indem wir die einzelnen Summanden dieser Zerlegung einzeln behandeln, können wir im folgenden annehmen, dass $\mathrm{minpol}_f$ und $\mathrm{charpol}_f$ Potenzen eines einzigen irreduziblen Polynoms $p$ sind.

Die wesentliche Arbeit beim Beweis der Existenz der unten angegebenen »rationalen Normalform« besteht darin, den folgenden Satz zu zeigen:

SATZ 17.28. *Der Vektorraum $V$ lässt sich als eine direkte Summe von $f$-zyklischen Untervektorräumen schreiben.*

Insgesamt kann man dann das folgende Theorem beweisen, das eine Normalform für Endomorphismen angibt, ohne dass man die Trigonalisierbarkeit annehmen muss.

THEOREM 17.29 (Rationale Normalform). *Seien $K$ ein Körper und $V$ ein endlichdimensionaler $K$-Vektorraum. Sei $f \in \mathrm{End}_K(V)$, und sei*

$$\mathrm{charpol}_f = \prod_{i=1}^{s} p_i^{n_i}$$

*die Zerlegung in ein Produkt irreduzibler normierter Polynome ($p_i \in K[X]$ paarweise verschieden). Dann existieren für jedes $i \in \{1, \ldots, s\}$ natürliche Zahlen $r_{i,1} \geq r_{i,2} \geq \cdots$ mit $\sum_j r_{i,j} = n_i$ und eine Basis $\mathscr{B}$ von $V$, so dass*

$$M_{\mathscr{B}}^{\mathscr{B}}(f) = \mathrm{diag}(A_1, \ldots, A_s)$$

*eine Diagonal-Blockmatrix ist, und für jedes $i$ die Matrix $A_i \in M_{N_i}$, $N_i := n_i \deg p_i$, selbst eine Diagonal-Blockmatrix ist, die zusammengesetzt ist aus den Begleitmatrizen der Polynome $p_i^{r_{i,1}}, p_i^{r_{i,2}}, \ldots$. Dabei sind die $p_i$ als die irreduziblen Teiler von $\mathrm{charpol}_f$ bis auf ihre Reihenfolge eindeutig und die Zahlen $r_{i,j}$ eindeutig bestimmt.*

*Für alle $i$ ist $p_i$ ein Teiler von $\mathrm{minpol}_f$, und $p_i^{r_{i,1}}$ ist die maximale Potenz von $p_i$, die $\mathrm{minpol}_f$ teilt.*

Siehe zum Beispiel [**Bo**] Kapitel 6.5, Theorem 18 für einen konzeptionellen Beweis, der allerdings einen weiteren Ausbau der Ringtheorie erfordert (siehe auch Abschnitt 18.7) oder [**Zi**] Kapitel 7.4.

## 17.7. Ergänzungen *

**17.7.1. Die Jordansche Normalform über $\mathbb{R}$.** Ist $A \in M_n(\mathbb{R})$ trigonalisierbar, dann besagt der Satz über die Jordansche Normalform, dass $A$ konjugiert ist zu einer Matrix $B \in M_n(\mathbb{R})$, die Jordansche Normalform hat.

Es ist eine naheliegende Frage, ob es eine »einfache Normalform« für beliebige Matrizen aus $M_n(\mathbb{R})$ gibt. In der Tat kann man mit der folgenden Definition eine solche Normalform beschreiben:

DEFINITION 17.30. Für $a, b \in \mathbb{R}$, $b \neq 0$, und $r \in \mathbb{N}_{>0}$ setzen wir

$$M_{a,b} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

und definieren die (nach einem zu den Jordanblöcken analogen Prinzip gebildete) Blockmatrix

$$J_{r,a,b} = \begin{pmatrix} M_{a,b} & E_2 & & & \\ & M_{a,b} & E_2 & & \\ & & \ddots & \ddots & \\ & & & M_{a,b} & E_2 \\ & & & & M_{a,b} \end{pmatrix} \in M_{2r}(\mathbb{R}).$$

⊣

THEOREM 17.31 (Jordansche Normalform über $\mathbb{R}$). *Sei $A \in M_n(\mathbb{R})$ eine quadratische Matrix über dem Körper der reellen Zahlen. Dann ist $A$ konjugiert zu einer Block-Diagonalmatrix, deren Blöcke entweder gewöhnliche Jordanblöcke oder Blöcke der Form $J_{r,a,b}$ (mit $r \geq 1$, $a, b \in \mathbb{R}$, $b \neq 0$) sind. Diese Normalform ist eindeutig bestimmt bis auf die Reihenfolge der Blöcke.*

Die Blöcke $J_{r,a,b}$ korrespondieren zu den irreduziblen Polynomen vom Grad 2, die das charakteristische Polynom (und das Minimalpolynom) von $A$ teilen. Es gilt ein ähnlicher Zusammenhang zwischen den Größen der Blöcke und den Vielfachheiten, mit denen diese Polynome in Minimalpolynom bzw. charakteristischem Polynom auftreten.

Zum Beweis kann man - grob skizziert – folgendermaßen vorgehen: Jedenfalls kann man die Matrix $A \in M_n(\mathbb{R})$ als Element von $M_n(\mathbb{C})$ betrachten, und eine Matrix $S \in GL_n(\mathbb{C})$ finden, für die $SAS^{-1}$ Jordansche Normalform hat (allerdings in $M_n(\mathbb{C})$, es werden also, wenn $A$ über $\mathbb{R}$ nicht trigonalisierbar ist, auch komplexe Zahlen als Einträge auftreten). Weil das charakteristische Polynom Koeffizienten in $\mathbb{R}$ hat, sind seine Nullstellen entweder reell, oder tritt für eine Nullstelle $\lambda \in \mathbb{C} \setminus \mathbb{R}$ die komplex konjugierte Zahl $\overline{\lambda}$ mit derselben Vielfachheit als Nullstelle auf. Man zeigt, dass auch die Größen der Jordanblöcke zu $\lambda$ bzw. zu $\overline{\lambda}$ übereinstimmen. (Das folgt aus der Eindeutigkeitsaussage über die Jordansche Normalform über $\mathbb{C}$.) Man kann dann zeigen, dass man je einen Jordanblock der Größe $r$ zu $\lambda$ und $\overline{\lambda}$ »zusammenfassen« kann zu einem Block der Form $J_{r,a,b}$.

Siehe zum Beispiel [**Kl**] Kapitel 5.6 für weitere Details.

**17.7.2. Lineare Differentialgleichungen mit konstanten Koeffizienten.** Die Jordansche Normalform ist nüztlich in der Theorie der linearen Differentialgleichungen mit konstanten Koeffizienten. Damit kann man die Methoden, die wir in Ergänzung I.10.28 im diagonalisierbaren Fall skizziert haben, auf den allgemeinen Fall übertragen (über $\mathbb{C}$ direkt, und über $\mathbb{R}$ mit Hilfe der Jordanschen Normalform über $\mathbb{R}$, Theorem 17.31.

Siehe [**Kl**] Kapitel 5.7. Siehe auch [**Wa**] Kapitel 1.

KAPITEL 18

# Konstruktionen von Vektorräumen

Wir kennen schon einige Möglichkeiten, um aus gegebenen Vektorräumen »neue« zu konstruieren, unter anderem die direkte Summe und das Produkt von Vektorräumen (Abschnitt I.6.6) und die Räume $\mathrm{Hom}_K(V, W)$ von linearen Abbildungen zwischen Vektorräumen.

In diesem Kapitel kommen wir zuerst noch einmal kurz auf Summe und Produkt zu sprechen, und betrachten dann einige weitere Konstruktionen von Vektorräumen:

- den Quotientenvektorraum $V/U$ eines Vektorraums $V$ nach einem Untervektorraum $U$,
- das Tensorprodukt von Vektorräumen,
- die äußeren Potenzen eines Vektorraums.

Die Quotientenkonstruktion ist eine Methode, die nicht nur für Vektorräume sondern auch für Mengen, Gruppen und Ringe in ähnlicher Weise durchführbar ist, und speziell im Kontext von Gruppen und von Ringen noch eine wesentlich größere Bedeutung hat, als für Vektorräume. Siehe die Abschnitte 18.3 und 18.4 für kurze Einführungen.

Um die Gemeinsamkeiten zwischen den verschiedenen Quotientenkonstruktionen deutlich zu machen (und das Fundament für weitere Verallgemeinerungen auf noch kompliziertere Objekte zu legen), diskutieren wir die Charakterisierung des Quotienten durch seine »universelle Eigenschaft«. Das klingt zuerst ein bisschen kompliziert, ist aber ein sehr mächtiges Konzept, das zum Beispiel in der Algebra und der algebraischen Geometrie von Bedeutung ist.

Es wird oft nützlich sein, die gegebenen Objekte und Abbildungen in einem sogenannten »kommutativen Diagramm« darzustellen.

DEFINITION 18.1. Ein *Diagramm* von Abbildungen ist gegeben durch eine Menge von Objekten und eine Menge von Abbildungen dazwischen.

Wir sprechen von einem *kommutativen Diagramm*, wenn für je zwei Objekte in dem Diagramm alle Verkettungen entlang verschiedener Wege vom ersten zum zweiten Objekt dieselbe Abbildung ergeben. ⊣

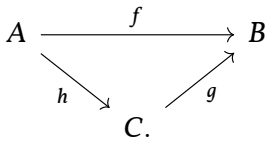
Die Definition lässt sich am einfachsten anhand einiger Beispiele erklären.

BEISPIEL 18.2. (1) Gegeben sei ein Diagramm

$$\begin{array}{ccc} A & \xrightarrow{g} & B \\ \downarrow{\scriptstyle t} & & \downarrow{\scriptstyle s} \\ C & \xrightarrow{f} & D. \end{array}$$

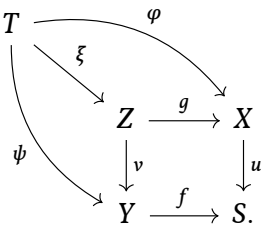
Das Diagramm ist genau dann kommutativ, wenn $f \circ t = s \circ g$ gilt.

(2) Gegeben sei ein Diagramm




Das Diagramm ist genau dann kommutativ, wenn $f = g \circ h$ gilt.

(3) Gegeben sei ein Diagramm




Das Diagramm ist genau dann kommutativ, wenn $\psi = v \circ \xi$, $\varphi = g \circ \xi$ und $f \circ v = u \circ g$ gilt. Die anderen Bedingungen, zum Beispiel $u \circ \varphi = f \circ v \circ \xi$, folgen daraus.

◊

## 18.1. Produkt, direkte Summe von VR

**18.1.1. Die universelle Eigenschaft des Produkts.** Sei $K$ ein Körper. Sei $I$ eine Menge ("Indexmenge"), und sei für jedes $i \in I$ ein Vektorraum $V_i$ gegeben. Wir haben in Abschnitt I.6.6 das Produkt und die direkte Summe der Familie $V_i$ definiert, und zwar ist

$$\prod_{i\in I} V_i = \{(v_i)_{i\in I};\ v_i \in V_i\}$$

als Menge das gewöhnliche kartesische Produkt, und die Vektorraumstruktur ist durch komponentenweise Addition und Skalarmultiplikation definiert. Die direkte Summe

$$\bigoplus_{i\in I} V_i = \left\{ (v_i)_{i\in I} \in \prod_{i\in I} V_i;\ v_i = 0 \text{ für alle bis auf höchstens endlich viele } i \in I \right\} \subseteq \prod_{i\in I} V_i$$

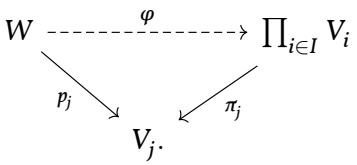
ist der Untervektorraum derjenigen Elemente, in denen nur endlich viele Einträge $\neq 0$ sind. Ist $I$ endlich, dann stimmen direkte Summe und direktes Produkt überein.

Ist $I = \{1, \ldots n\}$, so schreiben wir auch $\prod_{i=1}^{n} V_i$ oder $V_1 \times \cdots \times V_n$ statt $\prod_{i\in I} V_i$.

Das Produkt erfüllt die folgende sogenannte »universelle Eigenschaft«.

Satz 18.3 (Universelle Eigenschaft des Produkts). *Mit den obigen Notationen sei* $V := \prod_{i\in I} V_i$. *Die* Projektionen $\pi_j \colon V \to V_j$, $(v_i)_i \mapsto v_j$, *sind Vektorraumhomomorphismen.*

*Sei $W$ ein Vektorraum zusammen mit Homomorphismen $p_j \colon W \to V_j$. Dann gibt es genau einen Homomorphismus $\varphi \colon W \to V$, so dass für alle $j \in I$ gilt: $p_j = \pi_j \circ \varphi$.*




Beweis. Wir definieren $\varphi$ durch

$$\varphi(w) = (p_i(w))_{i\in I}.$$

Es ist leicht zu sehen, dass diese Abbildung die gewünschten Eigenschaften hat, und dass es keine andere Möglichkeit gibt, eine solche Abbildung zu definieren. □

In Teil (2) des Satzes nennt man den Vektorraum $W$ (zusammen mit den Homomorphismen $p_j$) auch das *Testobjekt* für die universelle Eigenschaft. Es ist wichtig, dass hier *jeder* Vektorraum als Testobjekt verwendet werden darf.

Der Beweis des Satzes ist so simpel, dass sich die Frage stellt, warum der Satz überhaupt nützlich ist. Uns dient der Satz hier vor allem der Illustration, wie eine universelle Eigenschaft eine *Charakterisierung* der entsprechenden Konstruktion liefert. Das formulieren wir folgendermaßen.

SATZ 18.4. *Seien $K$ ein Körper, $I$ eine Menge, und für $i \in I$ sei ein $K$-Vektorraum $V_i$ gegeben. Sei $P$ ein $K$-Vektorraum zusammen mit Vektorraum-Homomorphismen $\psi_j \colon P \to V_j$, so dass gilt:*

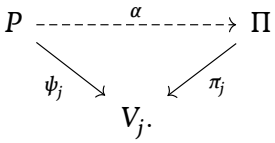
*Für jeden $K$-Vektorraum $W$ (»Testobjekt«) zusammen mit Homomorphismen $p_j \colon W \to V_j$ gibt es genau einen Homomorphismus $\varphi \colon W \to P$, so dass für alle $j \in I$ gilt: $p_j = \psi_j \circ \varphi$.*

*Dann gibt es einen* eindeutig bestimmten *Isomorphismus $\alpha \colon P \xrightarrow{\sim} \prod_{i\in I} V_i$, so dass $\psi_j = \pi_j \circ \alpha$ für alle $j$ gilt. (Hier bezeichnet wieder $\pi_j \colon \prod_i V_i \to V_j$ die Projektion.)*

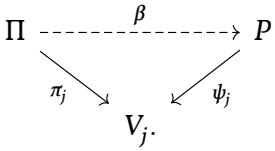
BEWEIS. Weil wir schon gesehen haben, dass das Produkt $\Pi := \prod_{i\in I} V_i$ und $P$ dieselbe universelle Eigenschaft erfüllen, lässt sich der Satz durch ein *rein formales Argument* in den folgenden vier Schritten beweisen.

Wir bezeichnen die Projektion $\Pi = \prod_{i\in I} V_i \to V_j$ wie oben mit $\pi_j$.

*Schritt 1: Konstruktion eines Homomorphismus $\alpha \colon P \to \Pi$.* Wir wenden die universelle Eigenschaft von $\Pi$ an (mit Testobjekt $P$). Mit $P$ und den Abbildungen $\psi_j \colon P \to V_j$ haben wir ein Testobjekt, das die Voraussetzungen erfüllt, und wir erhalten einen eindeutig bestimmten Homomorphismus $\alpha \colon P \to \Pi$, so dass $\psi_j = \pi_j \circ \alpha$ für alle $j \in I$ gilt.




*Schritt 2: Konstruktion eines Homomorphismus $\beta \colon \Pi \to P$.* Symmetrisch dazu wenden wir jetzt die universelle Eigenschaft von $P$ an. Mit $\Pi$ und den Abbildungen $\pi_j \colon \Pi \to V_j$ haben wir ein Testobjekt, das die Voraussetzungen erfüllt, und wir erhalten einen eindeutig bestimmten Homomorphismus $\beta \colon \Pi \to P$, so dass $\pi_j = \psi_j \circ \beta$ für alle $j \in I$ gilt.

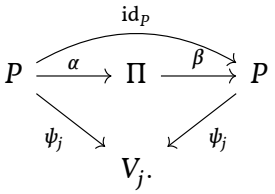



*Schritt 3: $\beta \circ \alpha = \mathrm{id}_P$.* Wir betrachten nun die beiden Homomorphismen $\mathrm{id}_P \colon P \to P$ und $\beta \circ \alpha \colon P \to P$. Es gilt

(a) $\psi_j = \psi_j \circ \mathrm{id}_P$, und

(b) $\psi_j = \pi_j \circ \alpha = \psi_j \circ (\beta \circ \alpha)$.

Wegen der Eindeutigkeitsaussage in der universellen Eigenschaft von $P$ (mit Testobjekt $P$) folgt, dass $\beta \circ \alpha = \mathrm{id}_P$ ist.




*Schritt 4: $\alpha \circ \beta = \mathrm{id}_\Pi$.* Das Argument verläuft genau symmetrisch zu Schritt 3.

Da $\alpha$ und $\beta$ zueinander invers sind, handelt es sich um Isomorphismen. Die Eindeutigkeit in den Schritten 1 und 2 folgt direkt aus der Eindeutigkeitsaussage der universellen Eigenschaft (auch ohne schon zu wissen, dass $\alpha$ und $\beta$ Isomorphismen sind). □

BEMERKUNG 18.5 (Nutzen der Charakterisierung durch eine universelle Eigenschaft). Die Charakterisierung durch eine universelle Eigenschaft erlaubt es, gewisse Konzepte – wie das Produkt – rein in Termen von Objekten und zugehörigen Abbildungen auszudrücken. Das funktioniert nicht nur für Vektorräume und Vektorraum-Homomorphismen, sondern immer, wenn wir eine »vernünftige« Klasse von Objekten und dazugehörigen Abbildungen (»Homomorphismen«, oder oft auch einfach »Morphismen«) an der Hand haben. (Der

»richtige« Begriff, um diese Situation zu formalisieren ist der Begriff der *Kategorie*, siehe Ergänzung 18.8.1.)

Zum Beispiel kann man so den Begriff des Produkts auch charakterisieren für

- Mengen und Abbildungen,
- Gruppen und Gruppenhomomorphismen,
- Ringe und Ringhomomorphismen,

und auch in vielen anderen Situationen.

So elegant die Definition eines Begriffs über die universelle Eigenschaft ist (wenn man sich erst einmal daran gewöhnt hat), hat sie doch einen Haken: Zwar bekommt man die Eindeutigkeit bis auf eindeutig bestimmten Isomorphismus »geschenkt«, aber ob so ein Objekt überhaupt existiert, lässt sich aus der Definition nicht ablesen. In der Tat ist es leicht, Beispiele von »Situationen« zu geben, wo ein Objekt, das die obige universelle Eigenschaft des Produkts hat, nicht existiert! Und ähnlich ist es für die anderen Beispiele aus der Liste unten, die man durch universelle Eigenschaften charakterisieren kann: Die Existenz muss jedesmal noch extra bewiesen werden.

Zum Beispiel gibt es keinen Körper $K$, der die universelle Eigenschaft des Produkts von $\mathbb{Q}$ und $\mathbb{F}_2$ erfüllt (mit Ringhomomorphismen als Abbildungen). Man kann zwar den Produktring $\mathbb{Q} \times \mathbb{F}_2$ betrachten, aber das ist (warum?) kein Körper. Dass es so einen Körper gar nicht geben kann, sieht man daran, dass es schon keinen Körper $K$ gibt, für den es sowohl einen Ringhomomorphismus $K \to \mathbb{Q}$ als auch einen Ringhomomorphismus $K \to \mathbb{F}_2$ gibt.

Bei Begriffen, die wir ohnehin durch eine konkrete Konstruktion definiert haben, ist das natürlich kein Problem. Aber zum Beispiel beim Tensorprodukt ist der Beweis, dass ein Objekt mit der gesuchten universellen Eigenschaft überhaupt existiert, etwas »lästig«. Immerhin ist das Gute, dass man die explizite Konstruktion, wenn die Existenz des gesuchten Objektes einmal gezeigt ist, in vielen Fällen nie wieder braucht, weil man alle Eigenschaften des Objekts mit der universellen Eigenschaft begründen kann.

Dasselbe Prinzip (aber eben mit anderen »universellen Eigenschaften«) lässt sich auf viele Konstruktionen anwenden, zum Beispiel kann man auch die folgenden Konstruktionen durch universelle Eigenschaften charakterisieren:

- die direkte Summe von Vektorräumen, siehe unten,
- den sogenannten Quotienten eines Vektorraums nach einem Unterraum (oder einer Gruppe nach einem Normalteiler oder eines Rings nach einem Ideal …), siehe die Abschnitte 18.2, 18.3, 18.4,
- den Kern eines (Vektorraum-)Homomorphismus,
- das Bild eines (Vektorraum-)Homomorphismus,
- den Polynomring über einem kommutativen Ring,
- das Tensorprodukt von Vektorräumen und die äußeren Potenzen eines Vektorraums, siehe die Abschnitte 18.5, 18.6.

◊

Bemerkung 18.6 (Analogien zur universellen Eigenschaft). Vielleicht ist es hilfreich, noch einmal an die folgenden Konstruktionen/Definitionen zu erinnern, die (in gewissem Maße) der Charakterisierung durch eine universelle Eigenschaft ähneln:

(1) Seien $R$ ein Integritätsring und $a, b \in R$. Ein Element $d$ heißt *ggT* von $a$ und $b$, wenn gilt:
   (a) $d \mid a, d \mid b$,
   (b) für jedes Element $d'$ mit $d' \mid a$ und $d' \mid b$ gilt $d' \mid d$.

Wenn Sie hier $x \mid y$ gedanklich als »es existiert $x \to y$« interpretieren, und voraussetzen, dass zwischen zwei »Objekten« immer höchstens eine Abbildung (»ein Pfeil«) existiert, dann liest sich die obige Definition ganz ähnlich wie die universelle Eigenschaft des Produkts von zwei Objekten $a$ und $b$.

Die Ähnlichkeit erstreckt sich auch dahin, dass aus der Definition nicht die Existenz eines ggT folgt, und dass ein ggT eindeutig bestimmt ist *bis auf Multiplikation mit einer (eindeutig bestimmten) Einheit von $R^\times$*.

(2) Sei $V$ ein Vektorraum und sei $M \subseteq V$ eine Teilmenge. Ein Untervektorraum $U \subseteq V$ heißt *von $M$ erzeugter Untervektorraum*, wenn gilt:

(a) $M \subseteq U$,

(b) für jeden Untervektorraum $U' \subseteq V$ mit $M \subseteq U'$ gilt $U \subseteq U'$.

Hier spielt $\subseteq$ die Rolle der Abbildungen und wir bekommen für jedes »Testobjekt« $U'$ eine »Abbildung« von $U$ nach $U'$. Diese Definition ähnelt daher der universellen Eigenschaft der direkten Summe, die wir in Abschnitt 18.1.2 anschauen wollen.

◊

Bemerkung 18.7. Man kann die universelle Eigenschaft des Produkts auch folgendermaßen umformulieren: Seien wie oben $K$-Vektoräume $V_i$, $i \in I$, gegeben, und seien $\pi_j \colon \prod_i V_i \to V_j$ die Projektionen. Für jeden $K$-Vektorraum $W$ ist der Homomorphismus

$$\operatorname{Hom}_K\left(W, \prod_{i\in I} V_i\right) \to \prod_{i\in I} \operatorname{Hom}_K(W, V_i), \quad \psi \mapsto (\pi_i \circ \psi)_{i\in I}$$

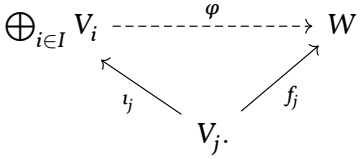
bijektiv. ◊

**18.1.2. Die universelle Eigenschaft des Koprodukts.** Die direkte Summe kann man in ähnlicher Weise durch eine universelle Eigenschaft charakterisieren.

Satz 18.8 (Universelle Eigenschaft der direkten Summe). *Seien $K$ ein Körper, $I$ eine Menge und sei für jedes $i \in I$ ein Vektorraum $V_i$ gegeben.*

(1) *Mit den obigen Notationen sei $V := \bigoplus i \in I V_i$. Die Inklusionen $\iota_i \colon V_i \to V$, $v \mapsto (\dots, 0, v, 0, \dots)$ ($v$ steht an der Stelle $I$) sind Homomorphismen.*

(2) *Sei $W$ ein Vektorraum zusammen mit Homomorphismen $f_i \colon V_i \to W$. Dann gibt es genau einen Homomorphismus $\varphi \colon \bigoplus_{i\in I} V_i \to W$, so dass für alle $i \in I$ gilt: $f_i = \varphi \circ \iota_i$.*




(3) *Sei $V'$ ein Vektorraum zusammen mit Homomorphismen $\iota_i' \colon V_i \to V'$, der auch die Eigenschaft in (2) hat. Dann gibt es einen eindeutig bestimmten Isomorphismus $\varphi \colon V \to V'$, so dass für alle $i$: $\iota_i' = \varphi \circ \iota_i$.*

Beweis. Der Beweis der Teile (1) und (2) ist einfach. Die Abbildung $\varphi$ in Teil (2) definiert man durch

$$\varphi((v_i)_{i\in I}) = \sum_{i\in I} f_i(v_i).$$

Weil höchstens endlich viele $v_i$ von Null verschieden sind, hat die Summe auf der rechten Seite nur endlich viele Summanden $\neq 0$. Für Teil (3) kann man ganz analog zu Satz 18.4

vorgehen. Man konstruiert in den ersten beiden Schritten Homomorphismen $V \to V'$ und $V' \to V$ mit der Existenzaussage der universellen Eigenschaften, und benutzt dann die Eindeutigkeitsaussage, um zu beweisen, dass beide Verkettungen mit der jeweiligen Identitätsabbildung übereinstimmen. □

Überlegen Sie sich, dass die direkte Summe aber (wenn $I$ unendlich ist und unendlich viele $V_i \neq 0$ sind) *nicht* die universelle Eigenschaft des Produkts erfüllt, und dass ebenso das Produkt in diesem Fall *nicht* die universelle Eigenschaft der direkten Summe erfüllt.

Zwischen Produkt und direkter Summe von Vektorräumen gibt es eine formale Analogie: Man erhält die universelle Eigenschaft der direkten Summe aus derjenigen des Produkts, indem man bei allen Abbildungen (allen »Pfeilen«) die Richtung umdreht:


Deshalb nennt man die direkte Summe manchmal auch das *Koprodukt* der Familie $(V_i)_{i\in I}$, besonders dann, wenn man über die universelle Eigenschaft spricht. Für das Koprodukt verwendet man auch das Symbol $\coprod_{i\in I}$.

Bemerkung 18.9 (Koprodukt von Mengen). Das gewöhnliche kartesische Produkt von Mengen erfüllt für Mengen und Abbildungen zwischen Mengen dieselbe universelle Eigenschaft wie das Produkt von $K$-Vektorräumen (und auch wie das Produkt von Gruppen und das Produkt von Ringen). Beim Koprodukt ist die Sache interessanter. Es gibt nämlich für Mengen $X$, $Y$ keine »natürliche« Abbildung $X \to X \times Y$, weil es – anders als im Fall von Vektorräumen mit dem Nullvektor – kein »ausgezeichnetes« Element von $Y$ gibt. Deshalb lässt sich ein Koprodukt $X \coprod Y$ von Mengen $X$ und $Y$ (also eine Menge, die die universelle Eigenschaft des Koprodukts für die Familie $X$, $Y$ erfüllt) nicht als Teilmenge des Produkts $X \times Y$ konstruieren.

Man kann aber Koprodukte von Mengen auf eine andere Art und Weise konstruieren, und zwar hat die *disjunkte Vereinigung* von zwei Mengen (bzw. allgemeiner von einer Familie $(X_i)_{i\in I}$ von Mengen) die richtige universelle Eigenschaft. Unter der disjunkten Vereinigung einer Familie $X_i$, $i \in I$ verstehen wir eine Menge $\coprod_{i\in I} X_i$ mit injektiven Abbildungen $\iota_i \colon X_i \to \coprod_{i\in I} X_i$, so dass $\coprod_{i\in I} X_i$ die Vereinigung aller $\iota_i(X_i)$ ist, und so dass $\iota_i(X_i) \cap \iota_j(X_j) = \emptyset$ für alle $i \neq j$ ist. Man bildet sozusagen die Vereinigung in einer Art und Weise, dass die Elemente der einzelnen $X_i$ jedenfalls voneinander getrennt bleiben. (Formal kann man das als die Vereinigung der Mengen $\{i\} \times X_i$ konstruieren, die Abbildung $\iota_i$ ist dann durch $x \mapsto (i, x_i)$ gegeben. Es ist nicht schwer nachzuprüfen, dass diese Konstruktion eine Menge (zusammen mit den Abbildungen $\iota_i$) liefert, die die universelle Eigenschaft des Koprodukts erfüllt.

Mithilfe dieser Konstruktion kann man dann wieder die universelle Eigenschaft des Koprodukts von Vektorräumen umformulieren: Ist $V_i$, $i \in I$, eine Familie von $K$-Vektorräumen, so ist für jeden Vektorraum $W$ die Abbildung

$$\operatorname{Hom}_K\left(\bigoplus_{i\in I} V_i, W\right) \to \coprod_{i\in I} \operatorname{Hom}_K(V_i, W), \quad \varphi \mapsto (\varphi \circ \iota_i)_{i\in I}$$

bijektiv. Vergleiche Satz I.7.15. ◊

## 18.2. Der Quotientenvektorraum

Wir kommen nun zur Konstruktion des Quotientenvektorraums. Dieser Konstruktion liegt die folgende Idee zugrunde: Sind $K$ ein Körper, $V$ ein $K$-Vektorraum und $U \subseteq V$ ein Untervektorraum, so möchten wir einen neuen Vektorraum $V/U$ zusammen mit einem surjektiven Homomorphismus $\pi\colon V \to V/U$ konstruieren, der $U$ als Kern hat. Die Konstruktion soll dabei nicht von irgendwelchen Wahlen abhängen (insbesondere wollen wir nicht benutzen, dass $U$ ein Komplement in $V$ besitzt – eine Tatsache, die den Basisergänzungssatz erfordert und die wir in der Linearen Algebra I deshalb auch nur für endlich erzeugte $V$ bewiesen haben).

BEMERKUNG 18.10. Zur Einstimmung und Motivation stellen wir zwei Vorüberlegungen an.

(1) Seien $K$ ein Körper, $V$ ein $K$-Vektorraum und $U \subseteq V$ ein Untervektorraum. Nehmen wir erstmal an, dass ein surjektiver Vektorraum-Homomorphismus $p\colon V \to W$ mit Kern $U$ gegeben ist. Wie sehen die Fasern von $p$ aus?

Jedenfalls gilt $p^{-1}(0) = \operatorname{Ker}(p) = U$. Allgemeiner gilt für $v, v' \in V$, dass sie genau dann in derselben Faser liegen, wenn $p(v) = p(v')$, oder äquivalent, wenn $v - v' \in U$ gilt.

Wählen wir zu $w \in W$ also irgendein $v \in V$ mit $p(v) = w$, so erhalten wir

$$p^{-1}(w) = v + U := \{v + u;\ u \in U\}.$$

Die Schreibweise $v + U$ haben wir schon in den ersten Wochen der Vorlesung *Lineare Algebra I* eingeführt, um die Lösungsmengen von inhomogenen linearen Gleichungssystemen zu beschreiben. Teilmengen von $V$ dieser Form entstehen einfach, indem man $U$ »verschiebt«, d.h. zu allen Elementen von $U$ denselben Vektor $v$ addiert. Und ist $v'$ irgendein Element aus $v + U$, so gilt $v + U = v' + U$. Mit anderen Worten haben wir

$$v + U = v' + U \quad \Longleftrightarrow \quad v - v' \in U.$$

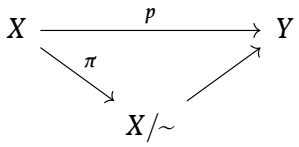
Diese Überlegungen können wir als Fahrplan benutzen, um in einer Situation, wo nur $V$ und $U$, aber nicht $W$ und $p$ gegeben sind, einen surjektiven Homomorphismus mit Kern $U$ zu konstruieren.

(2) Die zweite Vorbereitung ist eine kurze Erinnerung an den Begriff der Äquivalenzrelation. Ist $X$ eine Menge und $\sim$ eine Äquivalenzrelation, dann bezeichnen wir $[x]$ die Äquivalenzklasse von $x \in X$ und mit $X/\!\sim$ die Menge der Äquivalenzklassen. Dann ist $\pi\colon X \to X/\!\sim,\ x \mapsto [x]$ eine surjektive Abbildung, und $\pi(x) = \pi(x')$ gilt genau dann, wenn $x \sim x'$ ist.

Ist umgekehrt $p\colon X \to Y$ eine surjektive Abbildung, so ist

$$x \sim x' \quad :\Longleftrightarrow \quad p(x) = p(x')$$

eine Äquivalenzrelation auf $X$, und es gibt eine eindeutig bestimmte Abbildung $X/\!\sim \to Y$, so dass das Diagramm




kommutativ ist. Außerdem ist die Abbildung $X/\!\sim \to Y$ bijektiv.

$\Diamond$

Nach diesen vorbereitenden Überlegungen können wir den Quotientenvektorraum konstruieren. Sei $K$ ein Körper. Es seien $V$ ein $K$-Vektorraum und $U \subseteq V$ ein Untervektorraum. Wir definieren auf $V$ die folgende Äquivalenzrelation:

$$v \sim v' \quad :\Longleftrightarrow \quad v - v' \in U.$$

Es ist fast offensichtlich, dass es sich um eine Äquivalenzrelation handelt: $v \sim v$ gilt, weil $U$ als Untervektorraum die $0$ enthält, für $v \sim v'$ gilt auch $v' \sim v$, weil $U$ mit jedem Element auch sein Negatives enthält, und die Transitivität folgt (für $v \sim v'$, $v' \sim v''$) aus

$$v'' - v = (v'' - v') + (v' - v) \in U,$$

weil $U$ abgeschlossen ist unter der Addition. Wir bezeichnen die Menge der Äquivalenzklassen mit $V/U := V/\sim$.

Die Äquivalenzklasse von $v \in V$ bezüglich dieser Äquivalenzrelation ist die Menge

$$v + U = \{v + u;\ u \in U\}.$$

Man nennt die Äquivalenzklassen die *Nebenklassen (von $U$ in $V$)*. Die Elemente der Äquivalenzklassen nennen wir auch *Repräsentanten* oder *Vertreter* der Äquivalenzklasse oder der Nebenklasse. Denn für jedes $v' \in v + U$ gilt $v + U = v' + U$ (denn zwei Nebenklassen sind – wie allgemein zwei Äquivalenzklassen bezüglich einer Äquivalenzrelation – entweder disjunkt oder gleich).

Als nächstes definieren wir auf $V/U$ die Struktur eines $K$-Vektorraums, d.h. wir definieren eine Addition und eine Skalarmultiplikation, so dass die Vektorraumaxiome erfüllt sind.

*Addition.* Wir würden für $v, v' \in V$ gerne die Definition

$$(v + U) + (v' + U) := (v + v') + U$$

machen. Wir müssen aber begründen, dass dies überhaupt *wohldefiniert* ist, weil die Nebenklassen $v + U$ und $v' + U$ sich auch anders darstellen lassen:

$$v + U = w + U \text{ wenn } v - w \in U, \quad v' + U = w' + U \text{ wenn } v' - w' \in U.$$

Dass die obige Definition tatsächlich sinnvoll ist, folgt daraus, dass wir dasselbe Ergebnis erhalten, wenn wir statt $v$ und $v'$ die Vektoren $w$ und $w'$ verwenden, um die Nebenklassen darzustellen:

$$v - w, v' - w' \in U \quad \Rightarrow \quad (v + v') - (w + w') \in U \quad \Rightarrow \quad (v + v') + U = (w + w') + U.$$

Weil die Zuordnungsvorschrift davon unabhängig ist, welche Repräsentanten der Nebenklassen wir verwenden, erhalten wir eine wohldefinierte Abbildung

$$+\colon V/U \times V/U \longrightarrow V/U\,, \quad (v + U) + (v' + U) = (v + v') + U.$$

Analog definieren wir eine Skalarmultiplikation. Die Vorschrift

$$\alpha(v + U) := (\alpha v + U), \qquad \text{für } v \in V,\ \alpha \in K$$

ist wohldefiniert, denn im Fall $v + U = v' + U$ gilt $v - v' \in U$, also auch $\alpha v - \alpha v' = \alpha(v - v') \in U$ und damit $\alpha v + U = \alpha v' + U$. Wir erhalten also eine Abbildung

$$K \times V/U \longrightarrow V/U\,, \quad \alpha \cdot (v + U) = (\alpha v) + U.$$

Es ist dann leicht nachzuprüfen, dass alle Vektorraumaxiome erfüllt sind. Weil die Abbildung $\pi\colon V \to V/U$, $\pi(v) = v + U$, die jeden Vektor $v$ auf seine Äquivalenzklasse abbildet, mit den Verknüpfungen verträglich ist, d.h.

$$\pi(v + v') = \pi(v) + \pi(v'), \quad \pi(\alpha v) = \alpha\pi(v), \quad \text{für alle } v, v' \in V,\ \alpha \in K,$$

kann man das als eine rein formale Angelegenheit erledigen. Zum Beispiel wie folgt für das Assoziativgesetz:

$$((v_1 + U) + (v_2 + U)) + (v_3 + U) = (\pi(v_1) + \pi(v_2)) + \pi(v_3) = \pi(v_1 + v_2) + \pi(v_3) = \pi(v_1 + v_2 + v_3),$$

und genauso gilt

$$(v_1 + U) + ((v_2 + U) + (v_3 + U)) = \pi(v_1 + v_2 + v_3).$$

Der Nullvektor in $V/U$ ist $0 + U = U$. Das Negative von $v + U$ ist $-v + U$.

Zudem ist dann klar, dass $\pi$ ein surjektiver Vektorraum-Homomorphismus ist. Der Kern des Homomorphismus $\pi\colon V \to V/U$ ist $\operatorname{Ker}\pi = U$, denn $\pi(v) = 0 + U$ ist gleichbedeutend mit $v + U = 0 + U$, also mit $v \in U$.

DEFINITION 18.11. Der oben konstruierte $K$-Vektorraum $V/U$ heißt der *Quotient des Vektorraums $V$ nach dem Untervektorraum $U$*.

Den surjektiven Homomorphismus $\pi\colon V \to V/U$ nennen wir die *kanonische Projektion* auf den Quotienten (oder manchmal die *Quotientenabbildung*). ⊣

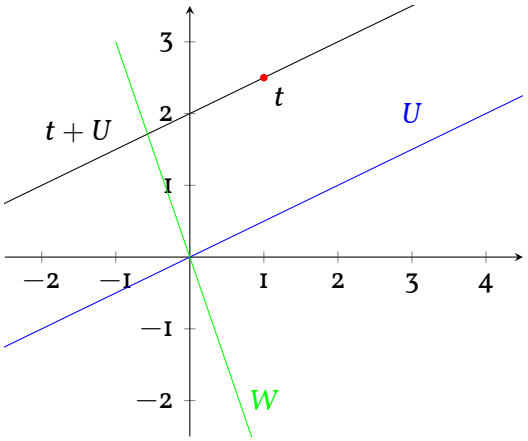
BEISPIEL 18.12. (1) Für $U = \{0\}$ ist die kanonische Projektion $V \to V/U$ ein Isomorphismus, wir können also $V \,/\, 0$ mit $V$ identifizieren.

(2) Für $U = V$ ist $V/U$ der Nullvektorraum, und für $U \subsetneq V$ gilt $V/U \neq 0$.

(3) Seien $U, W \subseteq V$ Komplementärräume, d.h. es gelte $V = U \oplus W$. Dann ist die Verkettung

$$f\colon W \to V \to V/U$$

der Inklusion von $W$ in $V$ mit der kanonischen Projektion $\pi$ ein Isomorphismus. (Einerseits ist $\operatorname{Ker}(f) = \operatorname{Ker}(\pi) \cap W = U \cap W = 0$, andererseits gilt für $v = u + w \in V$ (mit $u \in U, w \in W$) $f(w) = w + U = v + U$, und daraus folgt die Surjektivität.)




Im hier gezeigten Beispiel ist $V = \mathbb{R}^2$ und $U$ eindimensional. Die Nebenklassen sind die zu $U$ parallelen Geraden. Für jeden Komplementärraum $W$ von $U$ in $\mathbb{R}^2$ (also für jede Ursprungsgerade $W \neq U$) gilt, dass jede Nebenklasse die Gerade $W$ in *genau einem* Punkt schneidet. Das besagt genau, dass die Abbildung

$$W \to \mathbb{R}^2 \to \mathbb{R}^2/U$$

bijektiv ist.

◊

Wie der folgende Satz zeigt, lässt sich auch der Quotientenvektorraum durch eine universelle Eigenschaft beschreiben. Wie im Fall von Produkt und Koprodukt charakterisiert die universelle Eigenschaft den Quotientenvektorraum (zusammen mit der kanonischen Projektion) eindeutig bis auf eindeutigen Isomorphismus. Zusammen mit der Präzisierung in Teil (2) wird der Satz oft als Homomorphiesatz bezeichnet.

SATZ 18.13. *Seien $K$ ein Körper, $V$ ein Vektorraum über $K$ und $U \subseteq V$ ein Untervektorraum. Sei $W$ ein $K$-Vektorraum und $p\colon V \to W$ ein Homomorphismus.*

(1) *(Universelle Eigenschaft des Quotienten) Wenn $U \subseteq \operatorname{Ker} p$ gilt, dann existiert ein eindeutig bestimmter Homomorphismus $f\colon V/U \to W$ mit $f \circ \pi = p$.*

(2) *Existiert $f$ mit $f \circ \pi = p$, so folgt $U \subseteq \operatorname{Ker} p$. Sind $p$ mit $U \subseteq \operatorname{Ker} p$ und $f$ wie in (1), so gilt:* $\operatorname{Im} f = \operatorname{Im} p$. *Die Abbildung $f$ ist genau dann injektiv wenn $U = \operatorname{Ker} p$, genauer gilt stets* $\operatorname{Ker} f = \operatorname{Ker}(p) / U$.

BEWEIS. zu (1). Da $\pi$ surjektiv ist, gibt es wegen der Bedingung $f \circ \pi = p$ höchstens eine Möglichkeit, die Abbildung $f$ zu definieren: Es muss

$$f(v + U) = p(v)$$

gelten. Zu beweisen ist hier aber (als erstem Schritt), dass diese Vorschrift wohldefiniert ist! Für $v, v' \in V$ mit $v + U = v' + U$ gilt $v - v' \in U \subseteq \operatorname{Ker}(p)$, also $p(v - v') = 0$, d.h. tatsächlich $p(v) = p(v')$. Wir können also $f(v + U) = p(v)$ definieren, weil der Wert $p(v)$ nicht von der Wahl des Repräsentanten der Nebenklasse $v + U$ abhängt.

Dass $f$ linear ist, ist dann leicht nachzuprüfen, zum Beispiel gilt

$$f((v + U) + (v' + U)) = f((v + v') + U) = p(v + v') = p(v) + p(v') = f(v + U) + f(v' + U).$$

Die Verträglichkeit mit der Skalarmultiplikation kann man anhand einer ähnlichen Rechnung einsehen.

zu (2). Wenn andererseits $f\colon V/U \to W$ mit $p = f \circ \pi$ existiert, dann gilt $\operatorname{Ker}(f) \subseteq \operatorname{Ker}(\pi) = U$.

Dass $\operatorname{Im} f = \operatorname{Im} p$ gilt, ist ebenfalls eine direkte Konsequenz der Gleichheit $p = f \circ \pi$, weil $\pi$ surjektiv ist.

Weil $U \subseteq \operatorname{Ker}(p)$ ist, können wir den Quotientenvektorraum $\mathit{Ker}(p)/U$ bilden. Wir erhalten einen injektiven Vektorraum-Homomorphismus $\operatorname{Ker}(p)/U \to V/U$, $v + U \mapsto v + U$ (für $v \in \operatorname{Ker}(p)$), so dass wir $\operatorname{Ker}(p)/U$ als Untervektorraum von $V/U$ auffassen können. Es gilt dann

$$v + U \in \operatorname{Ker}(f) \Leftrightarrow p(v) = 0 \in W \Leftrightarrow v \in \operatorname{Ker}(p) \Leftrightarrow v + U \in \operatorname{Ker}(p)/U\ .$$

Damit haben wir gezeigt, dass $\operatorname{Ker} f = \operatorname{Ker}(p)/U$ gilt. Insbesondere erhalten wir

$$f \text{ injektiv} \Leftrightarrow \operatorname{Ker}(f) = 0 \Leftrightarrow \operatorname{Ker}(p)/U = 0 \Leftrightarrow U = \operatorname{Ker}(p).$$

□

Wir halten noch einen besonders wichtigen Spezialfall fest, der sich direkt aus dem Satz ergibt.

KOROLLAR 18.14. *Sei $f\colon V \to W$ ein Vektorraumhomomorphismus, $\pi\colon V \to V/\operatorname{Ker} f$ die kanonische Projektion, $\iota\colon \operatorname{Im} f \to W$ die Inklusion. Dann faktorisiert $f$ eindeutig als $f = \iota \circ g \circ \pi$ mit einem Isomorphismus $g\colon V/\operatorname{Ker} f \to \operatorname{Im} f$.*

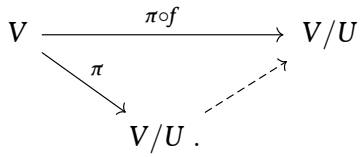
SATZ 18.15. *Sei $V$ endlich-dimensional, $U \subseteq V$ ein Untervektorraum. Dann ist* $\dim U + \dim V/U = \dim V$.

BEWEIS. Das ist eine unmittelbare Konsequenz der Dimensionsformel für die lineare Abbildung $\pi\colon V \to V/U$, denn $\pi$ ist surjektiv und hat Kern $U$. □

SATZ 18.16. *Seien $V$ ein $K$-Vektorraum, $f\colon V \to V$ ein Endomorphismus und $U \subseteq V$ ein $f$-invarianter Untervektorraum. Dann »induziert« $f$ einen Endomorphismus des Quotienten $V/U$, das heißt es gibt einen eindeutig bestimmten Endomorphismus $\bar{f}\colon V/U \to V/U$, so dass das folgende Diagramm kommutiert:*

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\ V/U & \xrightarrow{\bar{f}} & V/U\ . \end{array}$$

BEWEIS. Wir wenden den Homomorphiesatz auf das Diagramm




Weil $U \subseteq \operatorname{Ker}(\pi \circ f)$ gilt (hier benutzen wir die Voraussetzung $f(U) \subseteq U$), erhalten wir eine eindeutig bestimmte gestrichelte Abbildung $\bar{f}\colon V/U \to V/U$, so dass das Dreieck kommutativ ist. Mit dieser Abbildung ist dann auch das Quadrat in der Aussage des Satzes kommutativ. □

## 18.3. Der Quotient einer Gruppe nach einem Normalteiler

Die Quotientenkonstruktion kann man nicht nur für Vektorräume durchführen, sondern zum Beispiel auch im Kontext von Gruppen und von Ringen. In diesem Abschnitt behandeln wir Quotienten von Gruppen, danach kommen wir kurz zu Quotienten von Ringen. In beiden Fällen ist zunächst zu überlegen, nach welcher Art von Objekten man Quotienten konstruieren möchte. Jedenfalls soll es wieder einen surjektiven Homomorphismus (d.h. Gruppenhomomorphismus bzw. Ringhomomorphismus, je nachdem, in welchem Kontext wir arbeiten) geben soll, dessen Kern das »Objekt« ist, nach dem wir den Quotienten bilden. Wir haben gesehen, dass der Kern eines Ringhomomorphismus immer ein Ideal (aber im allgemeinen kein Unterring) ist. Deswegen werden wir im Fall von Ringen *Quotienten nach Idealen* betrachten.

Im Fall von Gruppen wissen wir, dass der Kern eines Gruppenhomomorphismus eine Untergruppe ist. Wir werden unten sehen, dass nicht jede Untergruppe wirklich als Kern auftreten kann, aber wir beginnen unsere Betrachtungen, indem wir die Überlegungen aus dem Vektorraumfall auf Gruppen und Untergruppen übertragen. Wenn nichts anderes gesagt wird, schreiben wir alle auftretenden Gruppen multiplikativ.

Die Definition von Nebenklassen $v + U$ eines Untervektorraums $U \subseteq V$ können wir leicht übertragen, indem wir die Vektorraumaddition durch die Gruppenverknüpfung ersetzen. Weil wir nicht voraussetzen, dass diese kommutativ ist, erhalten wir aber zwei (in der Regel unterschiedliche) Begriffe von Nebenklassen.

Definition 18.17. (1) Für $g \in G$ heißt

$$gH = \{gh;\ h \in H\}$$

die *Linksnebenklasse* von $g$ bezüglich $H$, und $Hg := \{hg;\ h \in H\}$ die *Rechtsnebenklasse* von $g$ bezüglich $H$.

(2) Die Menge der Linksnebenklassen von $H$ in $G$ wird mit $G/H$ bezeichnet. Die Menge der Rechtsnebenklassen bezeichnen wir mit $H\backslash G$.

⊣

Die Linksnebenklassen von $H$ in $G$ sind genau die Äquivalenzklassen bezüglich der Äquivalenzrelation

$$g \sim g' \quad \Longleftrightarrow \quad g^{-1}g' \in H.$$

Insbesondere gilt für $g, g' \in G$ entweder $gH = g'H$ oder $gH \cap g'H = \emptyset$. Sind $gH, g'H$ Linksnebenklassen, so ist die Abbildung $x \mapsto g'g^{-1}x$ eine Bijektion $gH \to g'H$ (mit Umkehrabbildung $y \mapsto (g')^{-1}gy$). Entsprechende Aussagen gelten für Rechtsnebenklassen. Als Folgerung erhalten wir:

SATZ 18.18 (Lagrange). *Sei $G$ eine endliche Gruppe und $H \subseteq G$ eine Untergruppe. Dann gilt*

$$\#G = \#H \cdot \#(G/H).$$

*Insbesondere ist $\#H$ ein Teiler von $\#G$.*

BEWEIS. Wir zählen die Elemente von $G$, indem wir die Anzahl $\#(G/H)$ der Nebenklassen multiplizieren mit der Anzahl der Elemente jeder Nebenklasse (diese Anzahl ist, wie wir soeben bemerkt haben, von der Nebenklasse unabhängig und ist gleich $\#H$, denn $H = 1\,H$ ist ja eine der Nebenklassen). □

Als nützliches Korollar des Satzes von Lagrange halten wir noch die folgende Aussage fest. Siehe Abschnitt I.8.5.1 für einige Anwendungen in der elementaren Zahlentheorie.

KOROLLAR 18.19. *Sei $G$ eine (multiplikativ geschriebene) endliche Gruppe mit $n$ Elementen und neutralem Element $e$, und sei $g \in G$. Dann gilt $g^n = e$.*

BEWEIS. Sei

$$H := \langle g \rangle = \{g^i;\ i \in \mathbb{Z}\}$$

die von $g$ erzeugte Untergruppe. Nach dem Satz von Lagrange ist $m := \#H$ ein Teiler von $G$. Es ist leicht zu sehen, dass dann $H = \{1, g, g^2, \ldots, g^{m-1}\}$ und $g^m = 1$ gilt. Damit folgt die Behauptung. □

BEMERKUNG 18.20. Unser Ziel ist nun, analog zum Vektorraumfall, die Menge $G/H$ mit einer Gruppenstruktur zu versehen, so dass die kanonische Projektion $\pi\colon G \to G/H, g \mapsto gH$ ein Gruppenhomomorphismus mit Kern $H$ ist. Damit das gelingen kann, müssen wir aber eine weitere Bedingung an $H$ stellen! Denn dass $\pi$ ein Gruppenhomomorphismus sein soll, bedeutet, dass die Multiplikation auf $G/H$ durch

$$(g_1H)(g_2H) := (g_1g_2)H$$

definiert werden müsste. Damit das wohldefiniert ist, muss aus $g_1H = g_1'H$ und $g_2H = g_2'H$ folgen, dass $(g_1g_2)H = g_1'g_2'H$ ist, mit anderen Worten muss gelten

$$g_i^{-1}g_i' \in H,\ i = 1, 2 \quad \Longrightarrow (g_1g_2)^{-1}g_1'g_2' \in H.$$

Es ist leicht zu sehen, dass das dazu äquivalent ist, dass für alle $h \in H$ und $g \in G$ auch $ghg^{-1} \in H$ gilt. In der Tat ist klar, dass jeder Kern eines Gruppenhomomorphismus diese Eigenschaft hat. Es ist nicht schwierig Beispiele von Gruppen $G$ und Untergruppen $H$ zu finden, für die diese Bedingung nicht gilt (schon in der symmetrischen Gruppe $G = S_3$ gibt es Beispiele). In kommutativen Gruppen tritt dieses Problem natürlich nicht auf; daher haben wir es auch beim Vektorraumquotienten nicht gesehen. ◊

Aufgrund der Überlegungen in der vorherigen Bemerkung treffen wir die folgende Definition.

DEFINITION 18.21. Sei $G$ eine Gruppe. Eine Untergruppe $H \subseteq G$ heißt *Normalteiler*, wenn die folgenden äquivalenten Bedingungen erfüllt sind:

(i) für alle $h \in H$ und $g \in G$ gilt $ghg^{-1} \in H$,

(ii) für alle $g \in G$ gilt

$$H = gHg^{-1} := \{ghg^{-1}; h \in H\}.$$

(iii) für alle $g \in H$ gilt $gH = Hg$.

⊣

Die Äquivalenz dieser Bedingungen ist nicht schwer zu zeigen. Dass die Normalteilereigenschaft eine notwendige Bedingung an $H$ ist, um einen Gruppenhomomorphismus mit Kern $H$ zu konstruieren, halten wir noch einmal explizit fest.

LEMMA 18.22. *Ist $f\colon G \to G'$ ein Gruppenhomomorphismus, dann ist* $\operatorname{Ker}(f)$ *ein Normalteiler von $G$.*

BEWEIS. Sind $h \in \operatorname{Ker}(f)$ und $g \in G$, so gilt

$$f(ghg^{-1}) = f(g)f(h)f(g)^{-1} = 1,$$

also haben wir $ghg^{-1} \in \operatorname{Ker}(f)$. □

Umgekehrt ist auch jeder Normalteiler $H \subseteq$ der Kern eines geeigneten Gruppenhomomorphismus, wie die Konstruktion des Quotienten $G/H$ zeigt.

DEFINITION 18.23 (Quotient einer Gruppe nach einem Normalteiler). Seien $G$ eine Gruppe und $H \subseteq G$ ein Normalteiler. Dann ist die Abbildung

$$G/H \times G/H \to G/H\,, \quad (g_1H, g_2H) \mapsto g_1g_2H)$$

wohldefiniert und definiert auf $G/H$ die Struktur einer Gruppe, die man als den *Quotienten von $G$ nach $H$* bezeichnet.

Die Abbildung $\pi\colon G \to G/H$ ist ein surjektiver Gruppenhomomorphismus mit Kern $H$, der als die *kanonische Projektion* bezeichnet wird. ⊣

BEWEIS. Zum Beweis der Wohldefiniertheit seien $g_1, g_1', g_2, g_2' \in G$ mit $g_iH = g_i'H$, $i = 1, 2$ gegeben. Wir wollen zeigen, dass $g_1g_2H = g_1'g_2'H$ ist, also dass $g_2^{-1}g_1^{-1}g_1'g_2' \in H$ gilt. Aber es ist

$$g_2^{-1}g_1^{-1}g_1'g_2' = g_2^{-1}(g_1^{-1}g_1'g_2'g_2^{-1})g_2 \in H,$$

weil $g_1^{-1}g_1'g_2'g_2^{-1} = g_1^{-1}g_1'(g_2(g_2')^{-1})^{-1}$ in $H$ liegt und $H$ ein Normalteiler ist.

Alternativ kann man sich davon überzeugen, dass die folgende Gleichheit von Teilmengen von $G$ gilt (wobei die Schreibweise $gH$ in naheliegender Weise verallgemeinert wird):

$$(g_1H)(g_2H) = g_1(Hg_2)H = g_1(g_2H)H = (g_1g_2)H,$$

und dass auch daraus die Wohldefiniertheit folgt.

Dass die Gruppenaxiome gelten, ist dann eine einfache Folgerung. Für das Assoziativgesetz haben wir

$$((g_1H)(g_2H))(g_3H) = (g_1g_2H)(g_3H) = (g_1g_2g_3H) = (g_1H)((g_2H)(g_3H)).$$

Das neutrale Element ist $H = 1H$, das inverse Element von $gH$ ist $g^{-1}H$.

Es ist eine direkte Folge der Definitionen, dass $\pi$ ein surjektiver Gruppenhomomorphismus ist. Es gilt $\pi(g) = 1_{G/H} = H$ genau dann, wenn $gH = H$ ist, also wenn $g$ in $H$ liegt. □

SATZ 18.24. (1) *(Universelle Eigenschaft des Quotienten) Sei $T$ eine abelsche Gruppe und $p\colon G \to T$ ein Gruppenhomomorphismus mit $H \subseteq \operatorname{Ker} p$. Dann existiert ein Gruppenhomomorphismus $f\colon G/H \to T$ mit $f \circ \pi = p$.*

(2) *(Homomorphiesatz) Sei $T$ eine abelsche Gruppe und $p\colon G \to T$ ein Gruppenhomomorphismus. Es existiert ein Gruppenhomomorphismus $f\colon G/H \to T$ mit $f \circ \pi = p$ genau dann, wenn $H \subseteq \operatorname{Ker} p$. In diesem Fall ist $f$ eindeutig bestimmt und es gilt $\operatorname{Im} f = \operatorname{Im} p$, und die Abbildung $f$ ist genau dann injektiv wenn $H = \operatorname{Ker} p$.*

BEWEIS. Den Beweis führt man genau wie im Vektorraumfall (siehe Satz 18.13). □

BEMERKUNG 18.25. In dem Fall, dass $G$ abelsch ist, ist jede Untergruppe $H$ von $G$ ein Normalteiler. Der Quotient $G/H$ ist dann auch eine abelsche Gruppe. ◊

ANHANG F

# Bemerkungen zur Literatur *

Die Bemerkungen zur Literatur im Skript zur Linearen Algebra 1, Abschnitt I.D, haben natürlich weiterhin Gültigkeit und die dort angegebenen Bücher und Skripte (beziehungsweise gegebenenfalls die zweiten Bände/Teile, die teilweise dort auch schon verlinkt sind) versorgen Sie mit allem Stoff (und noch deutlich mehr), den wir in der Linearen Algebra 2 behandeln werden.

Was hier noch ergänzt werden soll, sind einige Bemerkungen dazu, welche Bücher/Texte einen ähnlichen Ansatz verfolgen wie wir in der Vorlesung (und wo man vielleicht einen anderen Blickwinkel findet). Denn im Vergleich zur Linearen Algebra 1 ist der Stoff von Teil 2 schon etwas weniger standardisiert. Um den Satz über die Jordansche Normalform zu beweisen, gibt es unterschiedliche Möglichkeiten, der Quotientenvektorraum wird oft schon früher behandelt, oft schon im ersten Semester zur Linearen Algebra, und die anderen Universalkonstruktionen, die wir erwähnen (Tensorprodukt und äußere Potenz) gehören nicht unbedingt zum Standardstoff. Bei den Bi- und Sesquilinearformen gibt es vor allem insofern Unterschiede, ob ausschließlich über den Körpern $\mathbb{R}$ und $\mathbb{C}$ gearbeitet wird, oder ob der Fall eines allgemeinen Grundkörpers zu Beginn ebenfalls betrachtet wird.

## F.1. Literaturverweise zu einigen Vorlesungsthemen

**F.1.1. Die Jordansche Normalform.** Die Vorlesung richtet sich nicht genau nach einer Vorlage, aber die Darstellung in den Büchern von Brieskorn (Lineare Algebra und Analytische Geometrie II), Fischer (Lernbuch Lineare Algebra und Analytische Geometrie) sind nicht so weit davon entfernt, wie wir es machen. Ebenso kann ich das Buch [**Vi**] von Vinberg, Kap. 6.4, empfehlen.

Ein anderer Zugang wird beispielsweise von Bosch [**Bo**] gewählt. Dort wird der Satz über die Jordansche Normalform aus dem »Struktursatz für endlich erzeugte Moduln über Hauptidealringen« gefolgert. Dieser Zugang ist konzeptioneller, erfordert aber einen beträchtlichen Aufwand zur Entwicklung dieser allgemeinen Theorie.

**F.1.2. Universalkonstruktionen.** Tensorprodukte und die äußere Algebra werden zum Beispiel auch in den Büchern von Bosch [**Bo**] und Waldmann (Lineare Algebra 2, https://doi.org/10.1007/978-3-662-53348-2) und im Skript von Löh (Lineare Algebra II[I]) besprochen.

**F.1.3. Bilinearformen und Sesquilinearformen.** Wie gesagt variiert hier der Grad der Allgemeinheit, in der das Thema durchgenommen wird. Ich habe mich für einen Mittelweg entschieden, bei dem die Theorie solange, wie es mathematisch keinen Unterschied macht, über allgemeinen Körpern aufgebaut wird, denn das hat – zum Beispiel für die Zahlentheorie – durchaus einen Nutzen. Ähnlich ist es auch im Buch von Lorenz (Lineare Algebra 2), jedenfalls soweit es die Bilinearformen betrifft. Brieskorn (Lineare Algebra und Analytische Geometrie II) geht noch einen Schritt weiter und lässt nicht nur Körper, sondern beliebige Schiefkörper als »Grundkörper« zu, und erhält so die Theorie in der letztendlich

[I] http://www.mathematik.uni-regensburg.de/loeh/teaching/linalg2_ss17/lecture_notes.pdf

richtigen Allgemeinheit. Für den ersten Kontakt erschien mir das aber sozusagen zuviel des Guten.

# Literaturverzeichnis


[Bo-A] S. Bosch, *Algebra* https://doi.org/10.1007/978-3-662-61649-9

[Bo] S. Bosch, *Lineare Algebra*, Springer Spektrum 2014, https://doi.org/10.1007/978-3-642-55260-1

[Fi] G. Fischer, *Lineare Algebra*, Springer Spektrum 2014, https://doi.org/10.1007/978-3-658-03945-5

[Fi-AG] G. Fischer, *Analytische Geometrie*, Vieweg+Teubner, 7. Aufl., 2001

[Hi] S. Hildebrandt, *Analysis 1*, Springer, 2. Aufl., 2006, https://doi.org/10.1007/3-540-29285-3

[Hu] T. Hungerford, *Algebra*, Springer Graduate Texts in Math. **73**, 1974.

[Jä] K. Jähnich, *Lineare Algebra*, Springer Hochschultext, 2. Aufl., 1981.

[Kl] W. Klingenberg, *Lineare Algebra und Geometrie*, Springer, 3. Aufl., 1992.

[LM] J. Liesen, V. Mehrmann, *Lineare Algebra*, Springer 2015. https://doi.org/10.1007/978-3-658-06610-9

[Lo] F. Lorenz, *Lineare Algebra 1*, Spektrum Akad. Verlag 2004.

[Ma] J. Matoušek, *Thirty-three Miniatures*, Mathematical and Algorithmic Applications of Linear Algebra, Student Math. Library **35**, AMS 2010. Siehe auch preliminary version[2].

[Pi] R. Pink, *Lineare Algebra I und II*, Zusammenfassung, 2016. https://people.math.ethz.ch/~pink/ftp/Lineare-Algebra-Zusammenfassung-20161006.pdf

[Sch] A. Schmidt, *Einführung in die algebraische Zahlentheorie*, Springer 2007, https://doi.org/10.1007/978-3-540-45974-3

[Vi] E. Vinberg, *A Course in Algebra*, Graduate Studies in Math. **56**, AMS 2003.

[Wa] S. Waldmann, *Lineare Algebra 2*, Springer 2017, https://doi.org/10.1007/978-3-662-53348-2)

[Zi] H. Zieschang, *Lineare Algebra und Geometrie*, Vieweg+Teubner 1997, https://doi.org/10.1007/978-3-322-80093-0


[2] https://kam.mff.cuni.cz/%7Ematousek/stml-53-matousek-1.pdf

# Index