Lineare Algebra IIa
Frühjahrssemester 2020
Mannheim

Claus Hertling

05.02.2020

Inhaltsverzeichnis

Die Vorlesung ist eine Fortsetzung der Vorlesung Lineare Algebra I im HWS 2019. Daher fängt sie mit Kapitel 10 und auf Seite 99 an.

hertling@math.uni-mannheim.de

10 Euklidische Ringe, Hauptidealringe und ZPE-Ringe

In diesem Kapitel ist R immer ein kommutativer Ring mit Eins, oft sogar ein Integritätsring (Definition 10.5).
Das Kapitel behandelt insbesondere die Struktur der Ringe $\mathbb{Z}$ und $K[x]$ und ihrer Elemente, unter anderem die Zerlegung von Zahlen oder Polynomen in Primzahlen bzw. irreduzible Polynome. Die Zerlegung von Polynomen in irreduzible Polynome wird bei der Konstruktion von Körpern als Quotientenringen (Kapitel 11) und bei der Jordannormalform (Kapitel 12) gebraucht werden.
Das Kapitel fängt mit einer Luxusversion (Luxus = die Teile (c) und (d) von Satz 10.2) des Euklidischen Algorithmus an.

Beispiel 10.1 Der Euklidische Algorithmus zur Bestimmung des ggT zweier ganzer Zahlen an einem Beispiel: $r_0 = 140, r_1 = 38$. Man bestimmt für $i \geq 2$ induktiv Zahlen q_{i-1} und r_i mit $0 \leq r_i < r_{i-1}$ und $r_{i-2} = q_{i-1}r_{i-1} + r_i$. Man stoppt, wenn $r_{n+1} = 0$ ist. Dann ist r_n der ggT von r_0 und r_1 (siehe Satz 10.2).

$$\begin{array}{l|l|l} 140 = 3 \cdot 38 + 26 & q_1 = 3 & r_2 = 26 \\ 38 = 1 \cdot 26 + 12 & q_2 = 1 & r_3 = 12 \\ 26 = 2 \cdot 12 + 2 & q_3 = 2 & r_4 = 2 \\ 12 = 6 \cdot 2 & q_4 = 6 & r_5 = 0. \end{array}$$

Hier ist $n = 4$ und $\mathrm{ggT}(140, 38) = r_4 = 2$.

Satz 10.2 *(Der Euklidische Algorithmus für $R = \mathbb{Z}$)*
(a) (Division mit Rest) Zu $a, b \in \mathbb{Z}$ mit $b \neq 0$ gibt es eindeutige Zahlen $q, r \in \mathbb{Z}$ mit $0 \leq r < |b|$ und $a = q \cdot b + r$.

(b) (Euklidischer Algorithmus) Zu r_0 und $r_1 \in \mathbb{Z}$ mit $r_1 \neq 0$ gibt es ein eindeutiges $n \in \mathbb{N}$ und eindeutige Zahlen $q_1 \in \mathbb{Z}$, $q_2 ..., q_n \in \mathbb{N}$, $r_2, ..., r_n \in \mathbb{N}$ und $r_{n+1} = 0$ mit:

$$\begin{aligned} & |r_1| > r_2 > ... > r_n > r_{n+1} = 0 \quad \text{und} \\ r_0 &= q_1 \cdot r_1 + r_2, \\ r_1 &= q_2 \cdot r_2 + r_3, \\ &\vdots \\ r_{n-2} &= q_{n-1} \cdot r_{n-1} + r_n, \\ r_{n-1} &= q_n \cdot r_n + r_{n+1} = q_n \cdot r_n \end{aligned}$$

Es ist

$$r_n = \mathrm{ggT}(r_0, r_1).$$

(c) (Erweiterter Euklidischer Algorithmus) In der Situation von (b) werden zwei weitere Zahlenfolgen $x_0, x_1, ..., x_n, x_{n+1} \in \mathbb{Z}$ *und* $y_0, y_1, ..., y_n, y_{n+1} \in \mathbb{Z}$ *definiert durch*

$$\begin{aligned} x_0 := 1,\ x_1 := 0,\ x_i := x_{i-2} - q_{i-1} \cdot x_{i-1} \quad &\text{für } i = 2, ..., n, n+1 \\ y_0 := 0,\ y_1 := 1,\ y_i := y_{i-2} - q_{i-1} \cdot y_{i-1} \quad &\text{für } i = 2, ..., n, n+1. \end{aligned}$$

Dann gilt für $i = 0, 1, ..., n, n+1$

$$r_i = x_i \cdot r_0 + y_i \cdot r_1.$$

Der Fall $i = n$ *sagt: Der* $\mathrm{ggT}(r_0, r_1)$ *ist Linearkombination von* r_0 *und* r_1*.*

(d) In der Situation von (b) und (c) und bei $r_0 \geq r_1 > 0$ *gilt außerdem*

$$\begin{aligned} (-1)^i x_i &> 0 \text{ für } i \neq 1, \\ x_1 = 0, \quad & x_0 = x_2 = 1 \leq |x_3| < |x_4| < ... < |x_n| < |x_{n+1}| = |\frac{r_1}{r_n}|, \\ (-1)^{i+1} y_i &> 0 \text{ für } i \neq 0, \\ y_0 = 0, \quad & y_1 = 1 \leq |y_2| < |y_3| < |y_4| < ... < |y_n| < |y_{n+1}| = |\frac{r_0}{r_n}|, \\ (-1)^i \cdot r_1 &= r_{i+1} \cdot x_i - r_i \cdot x_{i+1} \quad \text{für } i = 0, ..., n, \\ (-1)^i \cdot r_0 &= r_i \cdot y_{i+1} - r_{i+1} \cdot y_i \quad \text{für } i = 0, ..., n, \\ (-1)^i &= x_i \cdot y_{i+1} - y_i \cdot x_{i+1} \quad \text{für } i = 0, ..., n. \end{aligned}$$

Beweis: (a) Klar.

(b) Man wendet so lange Division mit Rest an, wie $r_i > 0$ ist. Wegen $|r_1| > r_2 > r_3 > ...$ bricht das Verfahren mit einem $r_{n+1} = 0$ ab.
Es ist $r_n = \mathrm{ggT}(r_0, r_1)$, denn:
1. r_n teilt alle r_i und insbesondere r_0 und r_1: man liest die Gleichungen von unten nach oben.
2. Jeder Teiler von r_0 und r_1 teilt alle r_i und insbesondere r_n: man liest die Gleichungen von oben nach unten.

(c) Beweis der Gleichung $r_i = x_i \cdot r_0 + y_i \cdot r_1$ mit vollständiger Induktion:
Induktionsanfang: Nach Definition von x_0, x_1, y_0, y_1 gilt sie für $i = 0$ und $i = 1$.
Induktionsschluß $i - 2 \& i - 1 \to i$:

$$\begin{aligned} r_i &= r_{i-2} - q_{i-1} \cdot r_{i-1} \\ &= (x_{i-2} \cdot r_0 + y_{i-2} \cdot r_1) - q_{i-1}(x_{i-1} \cdot r_0 + y_{i-1} \cdot r_1) \\ &= (x_{i-2} - q_{i-1} \cdot x_{i-1}) \cdot r_0 + (y_{i-2} - q_{i-1} \cdot y_{i-1}) \cdot r_1 \\ &= x_i \cdot r_0 + y_i \cdot r_1. \end{aligned}$$

(d) Die drei letzten Gleichungen, die Ungleichungen und die Vorzeichenaussagen beweist man mit vollständiger Induktion (Details: Übung). Aus zwei der drei letzten Gleichungen und aus $r_{n+1} = 0$ folgen dann für $i = n$ die Gleichungen $x_{n+1} = (-1)^{n+1} \frac{r_1}{r_n}$ und $y_{n+1} = (-1)^n \frac{r_0}{r_n}$. □

Beispiele 10.3 (a) Nochmal das Beispiel 10.1.

i	r_i	q_i	x_i	y_i	
0	140	–	1	0	$140 = 1 \cdot 140 + 0 \cdot 38$
1	38	3	0	1	$38 = 0 \cdot 140 + 1 \cdot 38$
2	26	1	1	-3	$26 = 1 \cdot 140 + (-3) \cdot 38$
3	12	2	-1	4	$12 = (-1) \cdot 140 + 4 \cdot 38$
4	2	6	3	-11	$2 = 3 \cdot 140 + (-11) \cdot 38$
5	0	–	$-\frac{r_1}{r_4} = -19$	$\frac{r_0}{r_4} = 70$	$0 = (-19) \cdot 140 + 70 \cdot 38$

also $n = 4$, $\quad ggT(r_0, r_1) = r_4 = 2$.

(b) Ein Beispiel mit größeren Zahlen: $r_0 = 272526$, $r_1 = 32574$.

$$
\begin{aligned}
r_{i-2} &= q_{i-1} \cdot r_{i-1} + r_i, \\
x_i &= x_{i-2} - q_{i-1} \cdot x_{i-1}, \\
y_i &= y_{i-2} - q_{i-1} \cdot y_{i-1},
\end{aligned}
$$

i	r_i	q_i	x_i	y_i
0	272526	–	1	0
1	32574	8	0	1
2	11934	2	1	-8
3	8706	1	-2	17
4	3228	2	3	-25
5	2250	1	-8	67
6	978	2	11	-92
7	294	3	-30	251
8	96	3	101	-845
9	6	16	-333	2786
10	0	–	$\frac{r_1}{r_9} = 5429$	$-\frac{r_0}{r_9} = -45421$

also $n = 9$, $\quad \mathrm{ggT}(r_0, r_1) = r_9 = 6 = -333 \cdot r_0 + 2786 \cdot r_1$.

Bemerkungen 10.4 (i) Die *Fibonacci-Zahlen* sind Zahlen $F_0, F_1, F_2, ... \in \mathbb{N} \cup \{0\}$, die durch die Rekursion

$$F_0 := 0, \ F_1 := 1 \text{ und } F_i := F_{i-2} + F_{i-1} \quad \text{für } i \geq 2$$

definiert sind,

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
F_i	0	1	1	2	3	5	8	13	21	34	55	89	144	233	377

(Beweis Übung:) Es gilt

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

(ii) **Satz** (Binet (1841), Lamé (1844), Beweis: Übung): Seien $r_0, r_1 \in \mathbb{Z}$ mit $0 < r_1 < F_{k+1}$. Dann ergibt der Euklidische Algorithmus in Satz 10.2 den $\mathrm{ggT}(r_0, r_1)$ in $n \leq k-1$ Schritten.

(iii) Man kann F_k mit (i) und $|\frac{1+\sqrt{5}}{2}| \approx 1,618$ und $|\frac{1-\sqrt{5}}{2}| \approx 0,618$ abschätzen und dann folgern (Details: Übung):

$$5 \cdot |\text{Dezimalstellen von } F_k| \geq k-1.$$

Deshalb und wegen (ii) ist bei $r_0, r_1 \in \mathbb{Z}$ mit $r_1 \geq 2$ die Anzahl n der Iterationsschritte im Euklidischen Algorithmus in Satz 10.2 kleiner oder gleich 5 mal die Zahl der Dezimalstellen von r_1: Es sei $k \in \mathbb{N}$ die Zahl mit $F_k \leq r_1 < F_{k+1}$.

$$n \leq k-1 \leq 5 \cdot |\text{Dezimalstellen von } F_k| \leq 5 \cdot |\text{Dezimalstellen von } r_1|.$$

(iv) Die drei Zahlenfolgen r_i, x_i, y_i werden mit der gleichen Rekursionsformel berechnet, nur mit unterschiedlichen Startwerten. Das ist gut für Implementierungen. Aber man kann sich die Berechnung der Folge y_i sparen, denn

$$y_i = \frac{r_i - x_i \cdot r_0}{r_1}.$$

(v) Der Euklidische Algorithmus funktioniert auch bei vielen anderen Ringen, zum Beispiel bei Polynomringen $K[t]$, und liefert wertvolle Strukturaussagen über diese Ringe. Das wird im folgenden entwickelt.

Definition 10.5 Sei $R \neq \{0\}$ ein kommutativer Ring mit Einselement 1 (Definition 2.1).
(a) Ein Element $a \in R - \{0\}$ heißt *Nullteiler*, falls ein Element $b \in R - \{0\}$ mit $a \cdot b = 0$ existiert.
(b) Der Ring R heißt *Integritätsring*, falls er keine Nullteiler hat.

Beispiele 10.6 (i) $\mathbb{Z}$ ist ein Integritätsring. Allgemeiner: Jeder Körper K ist ein Integritätsring (Lemma 2.4 (d)), und ebenso jeder Unterring $R \subset K$.

(ii) Ein Polynomring $K[x]$ über einem Körper K ist ein Integritätsring (Lemma 2.18 (f) in den Ergänzungen der großen Übung zu Kapitel 2).

(iii) Wenn m eine Primzahl ist, ist der Ring $\mathbb{Z}_m$ ein Körper (Satz 2.9 (c)). Wenn $m = a \cdot b$ zerlegbar ist mit $a, b \notin \{\pm 1\}$, dann sind $[a]_m$ und $[b]_m \in \mathbb{Z}_m$ Nullteiler; dann ist $\mathbb{Z}_m$ kein Integritätsring.

Lemma/Definition 10.7 *Sei R ein kommutativer Ring mit Einselement* 1.
(a) (Lemma) Die Menge

$$R^* := \{a \in R \mid \text{ es gibt ein } b \in R \text{ mit } a \cdot b = 1\}$$

ist eine Gruppe.
(Definition) Sie heißt Einheitengruppe *von R. Ihre Elemente heißen* Einheiten *von R.*

(b) (Definition) Ein Element $a \in R$ teilt *ein Element $b \in R$, falls es ein Element $c \in R$ mit $b = a \cdot c$ gibt. Notation: $a|b$, a ist ein* Teiler *von b.*

(c) (Definition) Zwei Elemente a und $b \in R$ heißen assoziiert *(Notation: $a \sim b$), wenn es eine Einheit $\varepsilon \in R$ mit $a = b \cdot \varepsilon$ gibt.*
*(Triviales Lemma) $\sim$ ist eine Äquivalenzrelation auf R. (*Äquivalenzrelationen *werden in Kapitel 11 behandelt/wiederholt.)*

(d) (Lemma) Wenn R ein Integritätsring ist, gilt

$$a \sim b \iff a|b \text{ und } b|a.$$

(e) Ein Element $b \in R$ ist ein größter gemeinsamer Teiler *von Elementen $a_1, ..., a_n \in R$, falls gilt:*

(i) b teilt $a_1, ..., a_n$.

(ii) Teilt $c \in R$ alle $a_1, ..., a_n$, so teilt c auch b.

(f) (Lemma) Sei R ein Integritätsring. Falls Elemente $a_1, ..., a_n$ einen größten gemeinsamen Teiler b besitzen, so ist die Menge $\mathrm{MggT}(a_1, ..., a_n)$ *aller größten gemeinsamen Teiler*

$$\mathrm{MggT}(a_1, ..., a_n) = \{c \in R \mid c \sim b\}.$$

Notation: Oft wird ein beliebiges oder ein ausgezeichnetes Element dieser Menge als $\mathrm{ggT}(a_1, ..., a_n)$ *bezeichnet (bei $R = \mathbb{Z}$ das positive Element).*

Beweis: (a) Ein Element $a \in R^*$ hat ein eindeutiges Inverses: Seien b und $\widetilde{b}$ Inverse von a, also $a \cdot b = 1$ und $a \cdot \widetilde{b} = 1$. Dann folgt $b = a \cdot \widetilde{b} \cdot b = a \cdot b \cdot \widetilde{b} = \widetilde{b}$. Dieses Inverse wird mit a^{-1} bezeichnet. R^* ist eine Gruppe wegen $(a^{-1})^{-1} = a$ und $(a \cdot b)^{-1} = b^{-1} \cdot a^{-1}$.
(b) Definition.
(c) Definition.
(d) $\Rightarrow$: Aus $a = \varepsilon \cdot b$ mit $\varepsilon \in R^*$ folgt $b = \varepsilon^{-1} \cdot a$; also gilt $b|a$ und $a|b$.
$\Leftarrow$: Der Fall $a = 0$ ist trivial ($\Rightarrow b = 0$). Sei $a \neq 0$. Aus $a = b \cdot c$ und $b = a \cdot d$ folgt $0 = a \cdot (1 - c \cdot d)$, also, weil R ein Integritätsring ist, $1 = c \cdot d$ und $c, d \in R^*$.
(e) Definition.
(f) mit (d) und (e) (ii). □

Beispiele 10.8 (i) Beispiel 10.23: es gibt Integritätsringe R und darin Elemente $a_1, ..., a_n \in R$, die keinen größten gemeinsamen Teiler besitzen.

(ii) $\mathbb{Z}^* = \{\pm 1\}$.

(iii) Ist K ein Körper, so ist $K^* = K - \{0\}$ und auch $K[x]^* = K - \{0\}$. Gleich wird gezeigt, daß dann größte gemeinsame Teiler existieren. Hier könnte man die Konvention setzen, daß $\mathrm{ggT}(a_1, ..., a_n) \in K[x]$ das eindeutige unitäre (d.h. Leitkoeffizient 1) Polynom in $\mathrm{MggT}(a_1, ..., a_n)$ sein soll.

(iv) $\mathbb{Z}_m^* = \{a \mid 0 < a < m, \mathrm{ggT}(a, m) = 1\}$ (Beweis Übung).

Definition 10.9 Ein Integritätsring R mit Einselement 1 ist ein *Euklidischer Ring* mit *Gradfunktion* $w : R - \{0\} \to \mathbb{N} \cup \{0\}$, wenn Division mit Rest möglich ist, d.h. wenn gilt:

> Zu $a, b \in R$ mit $b \neq 0$ gibt es $q, r \in R$ mit
> $a = q \cdot b + r$ und ($r = 0$ oder $w(r) < w(b)$).

Achtung: q und r müssen nicht eindeutig sein.

Beispiele 10.10 Die beiden wichtigsten Beispiele Euklidischer Ringe sind:
(i)

$$\mathbb{Z} \quad \text{mit} \quad w(n) := |n|.$$

In diesem Fall sind q und r nicht eindeutig; aber mit der Zusatzforderung $r \geq 0$ (also $0 \leq r < |b|$) erhält man eindeutige q und r.
(ii)

$$K[t] \quad \text{mit} \quad w(p(t)) := \deg p(t)$$

(Satz 2.21 (a) in den Ergänzungen der großen Übung zu Kapitel 2). Hier sind q und r eindeutig.

Satz 10.11 *Sei R ein Euklidischer Ring mit Gradfunktion $w : R - \{0\} \to \mathbb{N} \cup \{0\}$. Seien $r_0, r_1 \in R$ mit $r_1 \neq 0$.*

(a) (Euklidischer Algorithmus) Es gibt Elemente $r_0, r_1, ..., r_n \in R - \{0\}$, $r_{n+1} = 0$, und $q_1, q_2, ..., q_n \in R$ mit

$$w(r_1) > w(r_2) > ... > w(r_n), \quad r_{n+1} = 0,$$
$$r_{i-2} = q_{i-1} \cdot r_{i-1} + r_i \quad \text{für } i = 2, ..., n+1.$$

Dann ist r_n ein größter gemeinsamer Teiler von r_0 und r_1. (Achtung: es wird nicht Eindeutigkeit der r_i und q_i gefordert.)

(b) (Erweiterter Euklidischer Algorithmus) Es werden $x_0, x_1, ..., x_n, x_{n+1} \in R$ und $y_0, y_1, ..., y_n, y_{n+1} \in R$ definiert durch

$$x_0 := 1,\ x_1 := 0,\ x_i := x_{i-2} - q_{i-1} \cdot x_{i-1} \quad \text{für } i = 2, ..., n, n+1$$
$$y_0 := 0,\ y_1 := 1,\ y_i := y_{i-2} - q_{i-1} \cdot y_{i-1} \quad \text{für } i = 2, ..., n, n+1.$$

Sie erfüllen für $i \geq 0$

$$\begin{aligned} r_i &= x_i \cdot r_0 + y_i \cdot r_1 \\ (-1)^i \cdot r_1 &= r_{i+1} \cdot x_i - r_i \cdot x_{i+1}, \\ (-1)^i \cdot r_0 &= r_i \cdot y_{i+1} - r_{i+1} \cdot y_i, \\ (-1)^i &= x_i \cdot y_{i+1} - y_i \cdot x_{i+1}. \end{aligned}$$

Die erste Gleichung sagt im Fall $i = n$, daß r_n Linearkombination von r_0 und r_1 ist.

Beweis: Wie Satz 10.2. □

Beispiel 10.12 $R = \mathbb{Q}[x]$,

$r_{i-2} = q_{i-1} \cdot r_{i-1} + r_i,$
$x_i = x_{i-2} - q_{i-1} \cdot x_{i-1},$
$y_i = y_{i-2} - q_{i-1} \cdot y_{i-1},$

i	r_i	x_i	y_i	q_i
0	$x^5 + x^4 + x^3 + x^2 + x + 1$	1	0	−
1	$x^4 + x^3 + 2x^2 + x + 1$	0	1	x
2	$-x^3 + 1$	1	$-x$	$-x - 1$
3	$2(x^2 + x + 1)$	$x + 1$	$-x^2 - x + 1$	$-\frac{1}{2}x + \frac{1}{2}$
4	0	$\frac{1}{2}x^2 + \frac{1}{2} = \frac{r_1}{r_3}$	$-\frac{1}{2}x^3 - \frac{1}{2} = -\frac{r_0}{r_3}$	

also $n = 3$,
$r_3 =$ (ein) $\mathrm{ggT}(r_0, r_1) = (x + 1) \cdot r_0 + (-x^2 - x + 1) \cdot r_1$.

Definition 10.13 Sei R ein kommutativer Ring mit Einselement 1.

(a) Ein *Ideal* I in R ist eine additive Untergruppe $I \subset R$ mit

$$a \in I,\ b \in R \Rightarrow b \cdot a \in I.$$

(Bemerkung) Ein Ideal ist ein Unterring von R.

(b) Das von einem Tupel $(a_j)_{j \in J}$ erzeugte Ideal wird wie das Tupel notiert und ist

$$(a_j)_{j \in J} := \{\sum_{j \in K} b_j \cdot a_j \mid K \subset J \text{ endlich}, b_j \in R\}.$$

(Bemerkungen) Es ist das kleinste Ideal in R, das alle a_j enthält. Hier ist $1 \in R$ wichtig.

(c) Ein *Hauptideal* ist ein von einem Element erzeugtes Ideal.
(Bemerkung) Die Hauptideale in R sind die Ideale

$$(a) = \{b \cdot a \mid b \in R\}, \quad a \in R,$$

inklusive $(0) = \{0\}$ und $(1) = R$.

(d) Ein kommutativer Ring R mit Eins ist ein *Hauptidealring*, falls alle Ideale in R Hauptideale sind und falls er ein Integritätsring mit Einselement ist.

Satz 10.14 *Ein Euklidischer Ring ist ein Hauptidealring.*

Beweis: Sei R ein Euklidischer Ring mit Gradfunktion $w : R - \{0\} \to \mathbb{N} \cup \{0\}$. Sei $I \subset R$ ein Ideal mit $I \neq \{0\}$. Dann enthält I ein Element $b \neq 0$ mit minimalem $w(b)$.
Behauptung: $I = (b)$.
Beweis der Behauptung und damit des Satzes: Sei $a \in I$. Division mit Rest gibt $a = q \cdot b + r$ mit $r \in I$ und mit $w(r) < w(b)$ oder $r = 0$. Wegen $w(b)$ minimal ist $r = 0$, also $a = q \cdot b \in (b)$. □

Beispiele 10.15 (i) $\mathbb{Z}$ und $K[x]$ sind Euklidische Ringe, also auch Hauptidealringe.

(ii) Im Ring $\mathbb{Z}[x]$ ist das Ideal

$$\begin{aligned}(2, x) &\stackrel{!}{=} \{f(x) \in \mathbb{Z}[x] \mid f(0) \in 2\mathbb{Z}\} \\ &= \{\sum_{i=0}^{n} a_i x^i \mid n \in \mathbb{N}_0, a_i \in \mathbb{Z}, a_0 \in 2\mathbb{Z}\}\end{aligned}$$

kein Hauptideal (Beweis: leichte Übung). Daher ist $\mathbb{Z}[x]$ kein Hauptidealring und kein Euklidischer Ring.

(iii) Hier ohne Beweis: es gibt Hauptidealringe, die keine Euklidischen Ringe sind.

(iv) Sei $R = K[x]$. Die Menge

$$\{\sum_{i=0}^{n} a_{2i} x^{2i} \mid n \in \mathbb{N}_0, a_{2i} \in K\}$$

ist ein Unterring von $K[x]$, aber kein Ideal in $K[x]$.

Lemma 10.16 *Sei R ein Hauptidealring. Dann haben beliebige Elemente $a_1, ..., a_n \in R$ einen größten gemeinsamen Teiler. Er ist Linearkombination der Elemente $a_1, ..., a_n$.*

Beweis: Sei a ein Erzeugendes des Hauptideals $(a_1, ..., a_n)$. Dann teilt a alle a_i wegen $a_i \in (a)$. Weil a Linearkombination der a_i ist, teilt jeder Teiler der a_i auch a. Daher ist a ein ggT von $a_1, ..., a_n$. □

Definition 10.17 Sei R ein Integritätsring mit Einselement 1.

(a) Ein Element $a \in R - (R^* \cup \{0\})$ heißt *irreduzibel* (oder *unzerlegbar*), falls gilt

$$a = b \cdot c \Rightarrow b \in R^* \text{ oder } c \in R^* (\text{d.h. } a \sim c \text{ oder } a \sim b).$$

(b) Ein Element $a \in R - (R^* \cup \{0\})$ heißt *Primelement*, falls gilt

$$a | (b \cdot c) \Rightarrow a | b \text{ oder } a | c.$$

Lemma 10.18 *Sei R ein Integritätsring mit Einselement* 1*.*

(a) Jedes Primelement ist irreduzibel.

(b) Falls R ein Hauptidealring ist, ist jedes irreduzible Element ein Primelement.

Beweis: (a) Sei $a \in R - (R^* \cup \{0\})$ ein Primelement und $a = b \cdot c$. Dann gilt $a|(b \cdot c)$, also $a|b$ oder $a|c$. Weil natürlich auch $b|a$ und $c|a$ gilt, folgt mit Lemma 10.7 (d) $a \sim b$ oder $a \sim c$.

(b) Sei $a \in R - (R^* \cup \{0\})$ irreduzibel, $a|(b \cdot c)$ und nicht $a|c$. Zu zeigen ist $a|b$. Weil a irreduzibel ist, sind seine einzigen Teiler die Einheiten und die zu a assoziierten Elemente. Wegen nicht $a|c$ ist

$$1 = \text{ (ein) } \operatorname{ggT}(a, c).$$

Nach Lemma 10.16 gibt es $\lambda_1, \lambda_2 \in R$ mit

$$1 = \lambda_1 \cdot a + \lambda_2 \cdot c.$$

Daher ist $b = \lambda_1 \cdot a \cdot b + \lambda_2 \cdot b \cdot c$. Daraus und aus $a|(b \cdot c)$ folgt $a|b$. □

Definition 10.19 Ein Integritätsring R ist ein *ZPE-Ring* (oder *Gaußscher Ring*), falls folgende Bedingungen erfüllt sind:

(i) Die irreduziblen Elemente und die Primelemente stimmen überein.

(ii) Jedes Element $a \in R - (R^* \cup \{0\})$ läßt sich als Produkt

$$a = p_1 \cdot \ldots \cdot p_n$$

von irreduziblen Elementen p_i schreiben.

(iii) Hat man zwei solche Produktdarstellungen

$$p_1 \cdot \ldots \cdot p_n = q_1 \cdot \ldots \cdot q_m,$$

so ist $n = m$, und es gibt eine Permutation $\sigma \in S_n$ mit $p_i \sim q_{\sigma(i)}$.

Satz 10.20 *Ein Hauptidealring ist ein ZPE-Ring.*

Beweis: (i) folgt aus Lemma 10.18.

(ii) **Annahme:** *Ein gegebenes $a \in R - (R^* \cup \{0\})$ läßt sich nicht als Produkt irreduzibler Elemente schreiben.*

Weil a dann selber nicht irreduzibel ist, ist $a = a_1 \cdot a_2$ mit $a_i \in R - (R^* \cup \{0\})$, und mindestens eines von a_1 und a_2, etwa a_2, läßt sich auch nicht als Produkt von

irreduziblen Elementen schreiben. Man setzt das fort und erhält (bei geeigneter Indizierung) eine unendliche Kette von immer größeren Hauptidealen

$$(a) \overset{\subset}{\neq} (a_2) \overset{\subset}{\neq} (a_4) \overset{\subset}{\neq} \ldots .$$

Hier wird benutzt (Lemma 10.7 (d)): bei $b_1, b_2 \in R - \{0\}$ ist

$$(b_1) = (b_2) \iff b_1 | b_2 \text{ und } b_2 | b_1 \iff b_1 \sim b_2.$$

Die Vereinigungsmenge $\bigcup_{i=1}^{\infty}(a_{2i})$ ist auch ein Ideal: Zu je zwei Elementen b und c in ihr gibt es ein i mit $b, c \in (a_{2i})$ und auch $b + c \in (a_{2i})$.
Weil R ein Hauptidealring ist, ist die Vereinigungsmenge ein Hauptideal mit einem Erzeugenden $\widetilde{a}$. Das Erzeugende $\widetilde{a}$ muß in einem der (a_{2i}) liegen. Aber dann wird die Kette von Hauptidealen bei (a_{2i}) konstant. Widerspruch. Also ist die Annahme falsch.
(iii) Sei $p_1 \cdot \ldots \cdot p_n = q_1 \cdot \ldots \cdot q_m$ mit p_i und q_j irreduzible Elemente. Nach (i) sind sie auch Primelemente.
p_1 teilt q_1 oder $q_2 \cdot \ldots \cdot q_m$. Induktiv erhält man, daß p_1 (mindestens) eines der q_j teilt, etwa $q_{\sigma(1)}$. Weil dieses irreduzibel ist, ist $p_1 \sim q_{\sigma(1)}$, also $q_{\sigma(1)} = \varepsilon_1 \cdot p_1$. Weil R ein Integritätsring ist, folgt aus

$$0 = p_1 \cdot \left(p_2 \ldots \cdot p_n - \varepsilon_1 \cdot \prod_{j \in \{1,\ldots,m\} - \{\sigma(1)\}} q_j \right)$$

auch

$$0 = p_2 \ldots \cdot p_n - \varepsilon_1 \cdot \prod_{j \in \{1,\ldots,m\} - \{\sigma(1)\}} q_j.$$

Nun macht man genauso mit p_2 weiter. Induktiv erhält man die Behauptungen. $\square$.

Bemerkungen 10.21 (i) **Satz** von Gauß (hier ohne Beweis): Ist R ein ZPE-Ring, so ist auch $R[x]$ ein ZPE-Ring.

(ii) Die Sätze 10.14 und 10.20 kann man in dem Diagramm

$$\{\text{Euklidische Ringe}\} \subset \{\text{Hauptidealringe}\} \subset \{\text{ZPE-Ringe}\}$$

zusammenfassen.
Ohne Beweis: Es gibt Hauptidealringe, die keine Euklidischen Ringe sind (Beispiel 10.15 (iii)).
Ohne Beweise: Der Ring $\mathbb{Z}[x]$ ist ein ZPE-Ring wegen (i); aber das Ideal $(2, x)$ ist kein Hauptideal in $\mathbb{Z}[x]$ (Beispiel 10.15 (ii)). Der Ring $K[x_1, x_2]$ ist ein ZPE-Ring wegen (i); aber das Ideal (x_1, x_2) ist kein Hauptideal in $K[x_1, x_2]$.

(iii) Als wichtige Folgerung soll man sich merken, daß die Polynomringe $K[x]$ und der Ring $\mathbb{Z}$ ZPE-Ringe sind.

(iv) **Frage:** *Welche Polynome in $K[x]$ sind irreduzibel (= irreduzible Elemente = Primelemente) ?.*
Bei $K = \mathbb{C}$ und $K = \mathbb{R}$ ist die Frage einfach zu beantworten: In $\mathbb{C}[x]$ sind es genau die linearen (d.h. Grad 1) Polynome (Satz 2.23 (a) in den Ergänzungen der großen Übung zu Kapitel 2); in $\mathbb{R}[x]$ sind es die linearen Polynome und die quadratischen Polynome ohne reelle Nullstellen (Satz 2.23 (b)).
Aber bei $K = \mathbb{Q}$ und $K = \mathbb{F}_p$ gibt es neben den linearen Polynomen irreduzible Polynome beliebig hohen Grades. Die Beantwortung der Frage ist in beiden Fällen schwer und führt auf reiche Strukturen.

(v) Bei Polynomen kleinen Grades kann man die Irreduzibilität leicht feststellen. Sei $f(x) \in K[x]$ mit $\deg f(x) \geq 1$. Dann gilt:

(a) $\deg f(x) = 1 \Rightarrow f(x)$ ist irreduzibel.

(b) $\deg f(x) \in \{2, 3\}$: $f(x)$ ist irreduzibel $\iff$ $f(x)$ hat keine Nullstelle.

(c) $\deg f(x) \geq 4$: $f(x)$ ist irreduzibel $\Rightarrow$ $f(x)$ hat keine Nullstelle.

Die Umkehrung $\Leftarrow$ gilt nicht, denn $f(x)$ könnte eine Zerlegung in zwei Faktoren ohne Nullstellen mit Graden ≥ 2 besitzen.

Beispiel 10.22 Sei $K = \mathbb{F}_2$. Die Anzahl der Polynome in $\mathbb{F}_2[x]$ vom Grad $0, 1, 2, 3, ..., n$ ist $1, 2, 4, 8, ..., 2^n$, denn bei den Polynomen $x^n + a_{n-1}x^{n-1} + ... + a_1x + a_0$ vom Grad n sind die Koeffizienten $a_{n-1}, ..., a_1, a_0 \in \mathbb{F}_2$ frei wählbar.
Wegen der Bemerkungen 10.21 (v) sind alle Polynome vom Grad 1 und genau die Polynome der Grade 2 und 3, die keine Nullstellen in $\mathbb{F}_2$ haben, irreduzibel.

x	irr
$x + 1$	irr
$x^2 = x \cdot x$	red
$x^2 + 1 = (x + 1)^2$	red
$x^2 + x = x(x + 1)$	red
$x^2 + x + 1$	irr
$x^3 = x \cdot x \cdot x$	red
$x^3 + 1 = (x + 1)(x^2 + x + 1)$	red
$x^3 + x = x(x + 1)^2$	red
$x^3 + x + 1$	irr
$x^3 + x^2 = x^2(x + 1)$	red
$x^3 + x^2 + 1$	irr
$x^3 + x^2 + x = x(x^2 + x + 1)$	red
$x^3 + x^2 + x + 1 = (x + 1)^3$	red

Auch die irreduziblen Polynome in $\mathbb{F}_2[x]$ vom Grad 4 kann man nun relativ leicht bestimmen (Übung): Es sind alle Polynome vom Grad 4, die keine Nullstellen haben und die auch nicht Produkte von zwei irreduziblen Polynomen vom Grad 2 sind.

Beispiel 10.23 Zum Abschluß ein Beispiel eines Ringes, wo viele der oben diskutierten Eigenschaften nicht erfüllt sind. Die Menge

$$\mathbb{Z}[\sqrt{-5}] := \{a + b \cdot \sqrt{-5} \mid a, b \in \mathbb{Z}\}$$

ist offensichtlich ein Ring mit Eins. Wegen $\mathbb{Z}[\sqrt{-5}] \subset \mathbb{C}$ ist der Ring ein Integritätsring. Er erfüllt:

(i) (Beweis: Algebra-Lehrbücher) Die vier Elementen 2, 3, $1 + \sqrt{-5}$, $1 - \sqrt{-5}$ sind alle irreduzibel, und keine zwei von ihnen sind assoziiert.

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

(iii) Wegen (i) und (ii) ist $\mathbb{Z}[\sqrt{-5}]$ kein ZPE-Ring.

(iv) Wegen (i) und (ii) sind die Elemente 2, 3, $1+\sqrt{-5}$, $1-\sqrt{-5}$ zwar irreduzibel, aber keine Primelemente.

(v) Die Elemente $2 \cdot (1+\sqrt{-5})$ und 6 besitzen keinen ggT. Denn 2 und $1+\sqrt{-5}$ sind zwei gemeinsame Teiler, aber es gibt keinen gemeinsamen Teiler, den beide teilen.

11 Quotienten bei Gruppen, Ringen und Vektorräumen

11.1 Äquivalenzrelationen

Definition 11.1 (a) Eine *(zweistellige) Relation* auf einer Menge $M \neq \emptyset$ ist eine Teilmenge R von $M \times M$.

(b) Eine Relation $R \subset M \times M$ ist eine *Äquivalenzrelation*, falls sie folgende drei Eigenschaften erfüllt:

(i) *Reflexivität:* Für alle $x \in M$ gilt: $(x, x) \in R$.

(ii) *Symmetrie:* Für alle $x, y \in M$ gilt: $(x, y) \in R \Rightarrow (y, x) \in R$.

(iii) *Transitivität:* Für alle $x, y, z \in M$ gilt: $(x, y) \in R$ und $(y, z) \in R \Rightarrow (x, z) \in R$.

(b) Notation: Bei einer Äquivalenzrelation R schreibt man oft $x \sim y$ (oder $x \sim_R y$) statt $(x, y) \in R$. Und man denkt sich x und y als *äquivalent* (in irgendeinem Sinne). Dann sehen die drei Eigenschaften so aus:

$$\begin{aligned} \text{Reflexivität:} &\quad \forall\, x \in M \quad x \sim x, \\ \text{Symmetrie:} &\quad \forall\, x, y \in M \quad x \sim y \Rightarrow y \sim x, \\ \text{Transitivität:} &\quad \forall\, x, y, z \in M \quad x \sim y \wedge y \sim z \Rightarrow x \sim z. \end{aligned}$$

(c) Sei R eine Äquivalenzrelation. Die *Äquivalenzklasse* $[x]$ eines Elementes $x \in M$ ist

$$[x] := \{y \in M \mid x \sim y\}.$$

(d) (Triviales Lemma) Sei R eine Äquivalenzrelation. Wegen der Transitivität und der Symmetrie gilt offenbar

$$x \sim y \iff [x] = [y] \iff [x] \cap [y] \neq \emptyset.$$

Also zerfällt M in lauter disjunkte Äquivalenzklassen.

(e) Notationen: Sei R eine Äquivalenzrelation. Dann ist jedes Element x ein *Repräsentant* seiner Äquivalenzklasse. Die Menge aller Äquivalenzklassen in M wird mit M/R (oder manchmal mit $M/\sim$) bezeichnet. Man hat die offensichtliche Abbildung

$$\pi_R : M \to M/R, \quad x \mapsto [x],$$

die jedes Element auf seine Äquivalenzklasse abbildet. Das Urbild einer Äquivalenzklasse (als Element von M/R) ist die Äquivalenzklasse (als Teilmenge von M).

(f) (Triviales Lemma, eine Umkehrung zu (d)) Ist $M \neq \emptyset$, und hat man eine Zerlegung von M in lauter disjunkte nichtleere Teilmengen,

$$M = \dot{\bigcup_{j \in J}} M_j,$$

(J ist irgendeine Indexmenge, nicht notwendig endlich; die M_j sind paarweise disjunkt), so definiert das eine eindeutige Äquivalenzrelation R mit den M_j als Äquivalenzklassen:

$$x \sim_R y \iff_{\text{Def}} \exists\, j \in J \text{ mit } x \in M_j \wedge y \in M_j.$$

Bemerkungen/Beispiele 11.2 (i) Bei konkreten Äquivalenzrelationen hat man konkrete Elemente und sieht die Elemente, die gewisse Eigenschaften teilen, als *äquivalent* an. Ob man das mit dem Begriff der Äquivalenz*relation* fasst, oder ob man einfach die Äquivalenz*klassen* vor Augen hat, kommt auf das gleiche raus (wegen (d) und (f)).

(ii) In den nächsten drei Unterkapiteln werden viele Äquivalenzrelationen gegeben, die von zusätzlicher mathematischer Struktur auf einer Menge M kommen. Aber es ist auch nicht schlecht, sich naive Äquivalenzklassen auszudenken. Eine steht in (iii).

(iii) Auf der Menge aller lebenden Menschen kann man die Äquivalenzrelation *(gleich viele Jahre alt)* betrachten. Eine Äquivalenzklasse besteht dann aus allen Menschen mit dem gleichen Lebensalter (in Jahren). Die Anzahl dieser Äquivalenzklassen liegt ein gutes Stück über 100. Jeder Mensch eines Alters ist ein Repräsentant der Äquivalenzklasse aller Menschen mit diesem Alter. Eine Äquivalenzklasse wird auch durch die gemeinsame Eigenschaft ihrer Repräsentanten, das Lebensalter, charakterisiert.

(iv) (Erinnerung) Die rationalen Zahlen kann man mit Hilfe von Äquivalenzklassen aus den ganzen Zahlen konstruieren. Auf der Menge $M := \{(a, b) \in \mathbb{Z}^2 \mid b \neq 0\}$ wird eine Äquivalenzrelation R durch

$$(a, b) \sim_R (c, d) \iff_{\text{Def}} ad = bc \quad \left(\iff \frac{a}{b} = \frac{c}{d} \right)$$

definiert. Dann wird $\mathbb{Q} := M/R$ als die Menge der Äquivalenzklassen definiert. Statt $[(a, b)]$ wird $\frac{a}{b}$ geschrieben. Als nächstes muß man diskutieren, wie auf $\mathbb{Q}$ die Standard-Rechenoperationen (wohl)definiert sind. Übrigens hat jede Äquivalenzklasse hier einen ausgezeichneten Repräsentanten. Das ist das eindeutige Paar (a, b) in der Äquivalenzklasse mit $b > 0$ und $\text{ggT}(a, b) = 1$.

(v) Das Programm in den nächsten drei Unterkapiteln ist verwandt zu (iv). Gegeben sind eine Menge M mit Struktur (e.g. eine Verknüpfung) und eine Zerlegung in disjunkte Mengen, die Äquivalenzklassen einer Äquivalenzrelation R. Dann wird untersucht, ob sich diese Struktur auf die Menge M/R der Äquivalenzklassen

übertragen läßt. Das Problem hat im Fall einer Verknüpfung $*$ immer die Gestalt: Ist für $a, b, c, d \in M$ mit $[a] = [c]$ und $[b] = [d]$ auch $[a * b] = [c * d]$? Falls ja, kann man eine Verknüpfung $\overline{*}$ auf M/R durch

$$[a]\overline{*}[b] := [a * b]$$

definieren. Denn diese Definition funktioniert genau dann, wenn die rechte Seite von der Wahl der Repräsentanten unabhängig ist.

Das folgende triviale Lemma trifft eine nah verwandte Aussage.

Lemma 11.3 *(Homomorphiesatz für Äquivalenzrelationen) Sei $M \neq \emptyset$ eine Menge und R eine Äquivalenzrelation auf M. Sei $N \neq \emptyset$ eine zweite Menge und $f : M \to N$ eine Abbildung. Dann gibt es genau dann eine Abbildung $\overline{f} : M/R \to N$ mit $f = \overline{f} \circ \pi_R$, falls äquivalente Elemente das gleiche Bild unter f haben (in Formeln: falls $a \sim b \Rightarrow f(a) = f(b)$ gilt). In dem Fall ist $\overline{f}([x]) = f(y)$ für einen beliebigen Repräsentanten y von $[x]$. Insbesondere ist dann $\overline{f}([x]) = f(x)$.*

In dem Fall sagt man, dass das Diagramm

$$\begin{array}{ccc} M & \xrightarrow{f} & N \\ {\scriptstyle \pi_R}\downarrow & \nearrow_{\overline{f}} & \\ M/R & & \end{array}$$

kommutiert.

11.2 Quotienten bei Gruppen

Notation 11.4 Sei $(G, \cdot)$ eine Gruppe.

(i) Für $A_1, A_2 \subset G$ nichtleer sei

$$A_1A_2 := \{a_1a_2 \mid a_1 \in A_1, a_2 \in A_2\}.$$

Dies definiert ein Produkt (d.h. eine Verknüpfung) auf der Menge der nichtleeren Teilmengen von G. Es ist assoziativ, weil das Produkt auf G assoziativ ist. Daher werden beim Produkt von ≥ 3 Mengen die Klammern weggelassen.

(ii) Es seien $c_1, c_2, c_3, ... \in G$ und $A_1, A_2, A_3, ... \subset G$.

$$c_1A_1 := \{c_1a_1 \mid a_1 \in A_1\} = \{c_1\}A_1, \quad A_1c_1 := \{a_1c_1 \mid a_1 \in A_1\} = A_1\{c_1\},$$

und analog z.B.

$$c_1A_1A_2c_2c_3A_3 := \{c_1a_1a_2c_2c_3a_3 \mid a_1 \in A_1, a_2 \in A_2, a_3 \in A_3\}.$$

(iii) Ist $U \subset G$ eine Untergruppe, so ist offenbar $eU = U = Ue = UU$.

Definition 11.5 Sei G eine Gruppe.

(a) Sei $a \in G$. Die Abbildung $l_a : G \to G,\ b \mapsto ab$, ist die Linksmultiplikation mit a. Die Abbildung $r_a : G \to G,\ b \mapsto ba$, ist die Rechtsmultiplikation mit a.

(b) Sei $U \subset G$ eine Untergruppe. Die Mengen $l_a(U) = aU$ für $a \in G$ heißen *Linksnebenklassen von U*. Die Mengen $r_a(U) = Ua$ sind die *Rechtsnebenklassen von U*. G/U bezeichnet die Menge $\{aU \mid a \in G\}$ der Linksnebenklassen von U. Analog bezeichnet $U\backslash G$ die Menge der Rechtsnebenklassen (aber das wird weniger gebraucht).

(c) Ist $(G, +)$ eine abelsche Gruppe mit additiv geschriebener Verknüpfung, so stimmen die Links- und Rechtsnebenklassen paarweise überein; sie werden mit $a + U := \{a + u \mid u \in U\}$ bezeichnet, manchmal (wenn klar ist, welche Untergruppe U gemeint ist) auch mit $[a]$.

Satz 11.6 *Sei G eine Gruppe.*

(a) Für jedes $a \in G$ sind die Abbildungen l_a und r_a bijektiv.

(b) Sei U eine Untergruppe, $a, b \in G$. Es ist entweder $aU = bU$ oder $aU \cap bU = \emptyset$; analog für die Rechtsnebenklassen. Daher ist G die disjunkte Vereinigung der Linksnebenklassen von U. Daher erhält man so eine Äquivalenzrelation $\sim_l$ mit den Linksnebenklassen als Äquivalenzklassen. Es ist $G/U = G/\sim_l$. Konkret sieht sie so aus:

$$a \sim_l b \iff \exists\ u \in U \text{ mit } a = bu.$$

Das gleiche gilt für die Rechtsnebenklassen: G ist disjunkte Vereinigung der Rechtsnebenklassen. Das gibt eine Äquivalenzrelation $\sim_r$ mit den Rechtsnebenklassen als Äquivalenzklassen. Es ist $U\backslash G = G/\sim_r$. Konkret sieht sie so aus:

$$a \sim_r b \iff \exists\ u \in U \text{ mit } a = ub.$$

Beweis: (a) l_a ist injektiv: $ab = ac \Rightarrow b = c$ (Kürzungsregel, Lemma 1.3 (c)).
l_a ist surjektiv: die Gleichung $a \cdot x = b$, wo x gesucht ist, hat die Lösung $x = a^{-1}b$.
Analog für r_a.
(b) Sei $aU \cap bU \neq \emptyset$; dann existieren $u_1, u_2 \in U$ mit $au_1 = bu_2 \in aU \cap bU$.
Für jedes $u \in U$ ist $au = au_1 \cdot u_1^{-1}u = bu_2 \cdot u_1^{-1}u \in bU$, also $aU \subset bU$; analog ist $bU \subset aU$; also ist $aU = bU$. Die Aussagen zu $\sim_l$ sind klar.
Der Beweis für die Rechtsnebenklassen ist analog. □

Der Satz liefert zwei Äquivalenzrelationen $\sim_l$ und $\sim_r$. Die große Frage im Sinne des Programms in Bemerkung 11.2 (v) ist, wann/ob die Gruppenstruktur auf G eine Gruppenstruktur auf G/U (oder auf $U\backslash G$) induziert. Sie wird nach Lemma 11.7 und den Beispielen 11.8 in Satz 11.9 beantwortet.

Lemma/Definition 11.7 *Sei U eine Untergruppe einer Gruppe G.*

(a) (Lemma) Für jedes $a \in G$ ist aUa^{-1} eine Untergruppe von G. Sie ist isomorph zu U.

(Definition) Diese Untergruppen heißen die zu U konjugierten Untergruppen.

(b) (Definition) U ist ein Normalteiler von G *genau dann, wenn $U = aUa^{-1}$ für alle $a \in G$ ist.*

(c) (Lemma) U ist ein Normalteiler von G genau dann, wenn $aU = Ua$ für alle $a \in G$ ist, d.h. wenn für jedes Element seine Linksnebenklasse und seine Rechtsnebenklasse übereinstimmen.

Beweis: (a) Man muß prüfen, daß aUa^{-1} abgeschlossen unter Produkt und Inversen-Bildung ist:
$au_1a^{-1}, au_2a^{-1} \in aUa^{-1} \Rightarrow (au_1a^{-1})(au_2a^{-1}) = au_1u_2a^{-1} \in aUa^{-1}$;
$aua^{-1} \in aUa^{-1} \Rightarrow (aua^{-1})^{-1} = au^{-1}a^{-1} \in aUa^{-1}$.
Die Abbildung $U \to aUa^{-1}$, $u \mapsto aua^{-1}$, ist offenbar ein Gruppenisomorphismus.
(b) Definition.
(c) Man multipliziere $U = aUa^{-1}$ von rechts mit a bzw $aU = Ua$ von rechts mit a^{-1}.
□

Beispiele 11.8 (i) Ist G abelsch, so ist jede Untergruppe von G ein Normalteiler von G.

(ii) Ist $f : G \to H$ ein Gruppenhomomorphismus, so ist $U := \ker(f)$ ein Normalteiler von G. (Daß $\ker(f)$ eine Untergruppe ist, war schon in Lemma 1.19 bewiesen worden.)

Beweis:
$$f(aUa^{-1}) = f(a)f(U)f(a^{-1}) = f(a)\{e_H\}f(a)^{-1} = \{e_H\},$$
also $aUa^{-1} \subset U$. Für a^{-1} statt a gibt das $a^{-1}Ua \subset U$, also $U \subset aUa^{-1}$. Also ist $aUa^{-1} = U$. □

(iii) Die Untergruppen $Z_1, Z_2, Z_3 \subset S_3$ von Beispiel 1.16 (iii) sind konjugiert zueinander; sie sind daher keine Normalteiler. Die Untergruppe $A_3 \subset S_3$ ist ein Normalteiler von S_3.

Satz 11.9 *(a) Sei G eine Gruppe und U eine Untergruppe. Auf G/U gibt es eine Verknüpfung $*$ mit $aU * bU = (ab)U$ für beliebige $a, b \in G$ genau dann, wenn U ein Normalteiler ist.*

(b) In dem Fall ist $(G/U, *)$ ein Gruppe. Sie ist die Quotientengruppe von G nach U. *Dann ist $*$ einfach die naive Multiplikation in der Notation 11.4 (i), $aU * bU = aUbU(= abU)$. Die Projektion $\pi_U : G \to G/U, a \mapsto aU$, ist ein surjektiver Gruppenhomomorphismus. Sein Kern ist* $\ker(\pi_U) = U$.

Beweis: (a) $\Rightarrow$: Sei $*$ eine Verknüpfung auf G/U mit $aU * bU = (ab)U$. Es soll gezeigt werden, dass $\pi_U : G \to G/U$ ein Gruppenhomomorphismus ist. Dann folgt mit Bemerkung 11.8 (ii), dass U ein Normalteiler ist.
Die Eigenschaft $\pi_U(a) * \pi_U(b) = aU * bU = (ab)U = \pi_U(ab)$ eines Gruppenhomomorphismus ist schon bekannt. Es bleibt zu zeigen, dass G/U eine Gruppe ist.

Die Assoziativität von $*$ auf G/U folgt aus der Assoziativität von $\cdot$ auf G:

$$(aU * bU) * cU = (ab)U * cU = (ab)cU = a(bc)U = aU * (bc)U = aU * (bU * cU).$$

Auch die anderen Gruppeneigenschaften übertragen sich direkt von G auf G/U.
$U = eU$ ist das neutrale Element:

$$aU * eU = aeU = aU \quad \text{und} \quad eU * aU = eaU = aU.$$

$a^{-1}U$ ist das inverse Element zu aU:

$$aU * a^{-1}U = aa^{-1}U = eU \quad \text{und} \quad a^{-1}U * aU = a^{-1}aU = eU.$$

Daher ist $(G/U, \cdot)$ eine Gruppe.
$\Leftarrow$: Sei $U \subset G$ ein Normalteiler. Man betrachtet die naive Verknpfung auf den nichtleeren Teilmengen von G, die in der Notation 11.4 (i) definiert war. Es seien $a, b \in G$. Man rechnet:

$$aUbU = abUU = abU.$$

Weil das Produkt wieder eine Linksnebenklasse ist, schränkt sich die naive Verknüpfung auf G/U ein. Man wählt $*$ als diese naive Verknüpfung.

(b) Das ist schon im Beweis von (a) gezeigt worden. □

Der nächste Satz ergänzt Satz 11.9. Er ist zugleich ein bißchen abstrakt und ein bißchen trivial.

Satz 11.10 *Sei $f : G \to H$ ein Gruppenhomomorphismus. Weil $\ker(f)$ ein Normalteiler ist (Beispiel 11.8 (ii)), ist $G/\ker(f)$ eine Gruppe (Satz 11.9). Auch $f(G)$ ist eine Gruppe (Lemma 1.19).*
Die Vorschrift $a\ker(f) \mapsto f(a)$ definiert eine Abbildung

$$\widetilde{f} : G/\ker(f) \to f(G).$$

$\widetilde{f}$ ist ein Gruppenisomorphismus.

Beweis: (1) $\widetilde{f}$ ist *wohldefiniert*: zu zeigen ist

$$a\ker(f) = b\ker(f) \Rightarrow f(a) = f(b).$$

Das gilt, denn falls $a\ker(f) = b\ker(f)$ gilt, existiert ein $u \in \ker(f)$ mit $a = bu$. Dann ist $f(a) = f(bu) = f(b)f(u) = f(b)e_H = f(b)$.
(2) $\widetilde{f}$ ist surjektiv: klar.
(3) $\widetilde{f}$ ist injektiv: sei $\widetilde{f}(a_1\ker(f)) = \widetilde{f}(a_2\ker(f))$. Dann ist $f(a_1) = f(a_2)$; also $f(a_2^{-1}a_1) = f(a_2)^{-1}f(a_1) = e_H$; also $a_2^{-1}a_1 \in \ker(f)$; also $a_1 = a_2(a_2^{-1}a_1) \in a_2\ker(f)$; also ist $a_1\ker(f) = a_2\ker(f)$.

(4) $\widetilde{f}$ ist ein Gruppenhomomorphismus:

$$\widetilde{f}\left((a\ker(f))\cdot(b\ker(f))\right) \overset{\text{Satz 11.9}}{=} \widetilde{f}(ab\ker(f)) \overset{\text{Def. von } \widetilde{f}}{=} f(ab)$$
$$\overset{f \text{ Gruppenhom.}}{=} f(a)f(b) \overset{\text{Def. von } \widetilde{f}}{=} \widetilde{f}(a\ker(f))\cdot\widetilde{f}(b\ker(f)). \qquad \square$$

Auch bei einer Untergruppe, die kein Normalteiler ist, ist es interessant, die Linksnebenklassen zu betrachten, insbesondere, wenn die Gruppe endlich ist.

Satz 11.11 *Sei G eine endliche Gruppe*

(a) (Satz von Lagrange) Sei U eine Untergruppe von G.
Dann sind auch U und G/U endlich, und alle Linksnebenklassen (und Rechtsnebenklassen) haben genauso viele Elemente wie U. Daher ist

$$|G| = |U|\cdot|G/U|.$$

Also teilt die Ordnung $|U|$ von U die Ordnung $|G|$ von G.

(b) Sei G eine endliche Gruppe. Für jedes $a \in G$ existiert eine kleinste Zahl $o(a) \in \mathbb{N}$ mit $a^{o(a)} = e$. Sie teilt $|G|$.
(Definition:) Diese Zahl heißt Ordnung von a.

Beweis: (a) G ist die Vereinigungsmenge der nach Satz 11.6 (b) paarweise disjunkten Linksnebenklassen. Daher ist die Menge G/U endlich, und es gibt geeignete $a_2, ..., a_{|G/U|} \in G$ mit

$$G/U = \{U, a_2U, ..., a_{|G/U|}U\}.$$

$l_{a_j} : U \to a_jU$ ist bijektiv. Daher ist $|G| = |U|\cdot|G/U|$.

(b) Die Menge $\{e, a, a^2, a^3, ...\} \subset G$ ist endlich. Daher gibt es ein $k_0 \in \mathbb{N}\cup\{0\}$ und ein $k_1 > k_0$ mit $a^{k_0} = a^{k_1}$. Also ist (Kürzungsregel) $a^{k_1-k_0} = e$. Also existiert $o(a)$. Die Menge $\{e, a, a^2, ..., a^{o(a)-1}\}$ ist eine Untergruppe von G. Ihre Ordnung ist $o(a)$. Also teilt $o(a)$ die Gruppenordnung $|G|$. $\square$

Beispiele 11.12 (i) Die Ordnung einer zyklischen Permutation $(a_1...a_k) \in S_n$ ist k. Die Ordnung einer beliebigen Permutation ist das kgV der Ordnungen der zyklischen Permutationen mit disjunkten Trägern, deren Produkt die Permutation ist (vgl. Satz 1.25 in den Ergänzungen der großen Übung zu Kapitel 1).

(ii) $S_n = A_n \cup (12)A_n, \quad |A_n| = \frac{n!}{2}$.

(iii) Die Ordnungen der Untergruppen der S_3 sind 1,2,3,6 (Beispiel 1.16 (iii)).

(iv) Sei $m \in \mathbb{N}$; es ist $(m\mathbb{Z}, +) \subset (\mathbb{Z}, +)$ eine Untergruppe. Die Menge $\mathbb{Z}$ ist die disjunkte Vereinigung der m Nebenklassen von $m\mathbb{Z}$:

$$\begin{aligned}\mathbb{Z} &= m\mathbb{Z}\cup(1+m\mathbb{Z})\cup...\cup((m-1)+m\mathbb{Z}),\\ \mathbb{Z}/m\mathbb{Z} &= \{m\mathbb{Z}, 1+m\mathbb{Z}, ..., (m-1)+m\mathbb{Z}\}.\end{aligned}$$

Die Quotientengruppe $(\mathbb{Z}/m\mathbb{Z}, +)$ ist natürlich isomorph zur Gruppe $(\mathbb{Z}_m, +_m)$ von Satz 2.9. Die Konstruktion hier als Quotientengruppe (mit den Äquivalenzklassen als neuen Elementen und der induzierten Addition) ist befriedigender als die alte naive Konstruktion (mit den Zahlen 0 bis $m-1$ in $\mathbb{Z}_m$ und der künstlichen Addition $+_m$). Im nächsten Unterkapitel wird auch die Multiplikation auf $\mathbb{Z}/m\mathbb{Z}$ definiert.

Notationen 11.13 (i) Vgl. Def. 11.5 (c): Ist $(G, +)$ eine abelsche Gruppe, so sind die Nebenklassen von U die Mengen $\{a+u \mid u \in U\} =: a+U =: [a]$. Die Verknüpfung $+$ von Satz 11.9 auf G/U ist gegeben durch

$$[a] + [b] = [a + b].$$

(ii) Im Fall $(G, +) = (\mathbb{Z}, +)$ und $U = m\mathbb{Z}$ für ein $m \in \mathbb{N}$ sagt man für $a+m\mathbb{Z} = b+m\mathbb{Z}$ auch

$$a \equiv b \mod m, \quad \text{in Worten: } a \textit{ ist kongruent zu } b \textit{ modulo } m.$$

Das ist äquivalent zu (m *teilt* $a - b$). Kongruent zu sein ist eine Äquivalenzrelation, die *Kongruenzrelation*. Die Äquivalenzklassen sind die Mengen $a + m\mathbb{Z}$; sie heißen *Kongruenzklassen*.

Definition/Lemma 11.14 *(a) (Definition) Eine Gruppe $(G, \cdot)$ heißt zyklische Gruppe der Ordnung $m \in \mathbb{N}$, falls ein $a \in G$ existiert mit $G = \{e, a, a^2, ..., a^{m-1}\}$ und $m = o(a)$. Dann ist $a^m = e$ und $G = \{e, a, a^2, ...\}$. Jedes solche Element ist ein Erzeugendes der Gruppe.*

(b) (Lemma) Dann ist

$$(G, \cdot) \to (\mathbb{Z}/m\mathbb{Z}, +),\ a^k \mapsto k + m\mathbb{Z},$$

ein Gruppenisomorphismus. Auch

$$(\mathbb{Z}/m\mathbb{Z}, +) \to (\{e^{2\pi ik/m} \mid k \in \mathbb{Z}_m\}, \cdot),\ k \mapsto e^{2\pi ik/m},$$

und

$$(\mathbb{Z}_m, +_m) \to (\mathbb{Z}/m\mathbb{Z}, +),\ k \mapsto k + m\mathbb{Z} = [k],$$

sind Gruppenisomorphismen. Sowohl die additive Version $(\mathbb{Z}/m\mathbb{Z}, +)$ als auch die multiplikative Version $(\{e^{2\pi ik/m} \mid k \in \mathbb{Z}_m\}, \cdot)$ sind wichtig.

(c) (Lemma) Die Erzeugenden der Gruppe $(\mathbb{Z}_m, +_m)$ sind genau die Elemente von $\mathbb{Z}_m^*$, also die Elemente $k \in \mathbb{Z}_m$ mit $\mathrm{ggT}(k, m) = 1$.

(d) (Definition) Sei $(G, \cdot)$ eine nicht notwendig endliche Gruppe, und sei $a \in G$. Falls eine Potenz von a gleich e ist, so gibt es ein minimales $o(a) \in \mathbb{N}$ mit $a^{o(a)} = e$. Das heißt die Ordnung *von a.*
(Lemma) In dem Fall ist $\{e, a, a^2, ..., a^{o(a)-1}\} \subset G$ die von a erzeugte zyklische Untergruppe von G.

Beweis: (a) Definition. (b) Klar. (d) Klar.
(c) Ein Element $k \in \mathbb{Z}_m$ ist genau dann ein Erzeugendes der Gruppe $(\mathbb{Z}_m, +_m)$, wenn die von k erzeugte zyklische Untergruppe die ganze Gruppe $\mathbb{Z}_m$ ist. Das ist äquivalent dazu, dass 1 in dieser Untergruppe liegt. Das ist äquivalent dazu, dass es ein $b \in \mathbb{Z}_m$ mit $b \cdot_m a = 1$ gibt. Und das ist äquivalent zu $\mathrm{ggT}(a, m) = 1$ (vgl. Beispiel 10.8 (iv) und Aufgabe 2 von Blatt 2). □

11.3 Quotienten bei Ringen

Im folgenden werden nur kommutative Ringe betrachtet, da wir nur die brauchen. Einiges des folgenden Materials läßt sich auch leicht für nichtkommutative Ringe entwickeln.
Wir werden im folgenden eine Analogie zwischen Gruppen und (kommutativen) Ringen feststellen, die ziemlich weit geht:

Gruppe	$\sim$	kommutativer Ring
Untergruppe	$\sim$	Unterring
Normalteiler	$\sim$	Ideal
Beispiel 11.8 (ii)	$\sim$	Lemma 11.15 (a)
Quotientengruppe nach Satz 11.9	$\sim$	Quotientenring nach Satz 11.17
Satz 11.10	$\sim$	Satz 11.18

Lemma 11.15 *(a) Seien R und S kommutative Ringe, und sei $f : R \to S$ ein Ringhomomorphismus (siehe Definition 2.5 (c): f respektiert Addition und Multiplikation). Dann ist* $\ker(f)$ $(= \{a \in R \mid f(a) = 0\})$ *ein Ideal.*

(b) Seien R, S und f wie in (a). Dann ist f genau dann injektiv, wenn $\ker(f) = \{0\}$ *ist.*

(c) Die einzigen Ideale in einem Körper sind $\{0\}$ und K.

(d) Jeder Körperhomomorphismus ist injektiv.

Beweis: (a) Wegen Lemma 1.19 ist $\ker(f) \subset R$ eine additive Untergruppe. Es bleibt die Idealeigenschaft zu zeigen. Sei $a \in \ker(f)$ und $b \in R$. Dann ist

$$f(b \cdot a) = f(b) \cdot f(a) = f(b) \cdot 0 = 0, \quad \text{also } b \cdot a \in \ker(f).$$

(b) Das ist ganz analog zu Satz 5.3 (d) (eine lineare Abbildung ist genau dann injektiv, wenn ihr Kern gleich $\{0\}$ ist). Es ist $f(0) = 0$ wegen Lemma 1.19.
$\Rightarrow$: f injektiv und $f(a) = 0 \Rightarrow a = 0$.
$\Leftarrow$: $f(a) = f(b) \Rightarrow f(a - b) = 0 \Rightarrow$ (wegen $\ker(f) = \{0\}$) $a - b = 0 \Rightarrow a = b$.

(c) Sei $I \subset K$ ein Ideal in einem Körper K mit $I \neq \{0\}$. Sei $a \in I - \{0\}$. Dann existiert a^{-1}. Also ist $1 = a^{-1} \cdot a \in I$. Also ist $I = K$.

(d) Nach Definition ist ein Körperhomomorphismus $f : K \to L$ ein Ringhomomorphismus mit $f(1_K) = 1_L$. Daher ist $\ker(f) \neq K$. Wegen (a) und (c) ist $\ker(f) = \{0\}$. Wegen (b) ist f injektiv. □

Bemerkung 11.16 Sei R ein kommutativer Ring und $I \subset R$ eine (additive) Untergruppe. Weil die Addition kommutativ ist, stimmen die Linksnebenklassen und die Rechtsnebenklassen überein und werden einfach Nebenklassen genannt und mit $[a] = a + I$ (für $a \in R$) bezeichnet. Nach Satz 11.9 ist $(R/I, +)$ eine abelsche Gruppe mit $[a] + [b] = [a + b]$, und die Projektion $\pi_I : R \to R/I,\ a \mapsto [a]$, ist ein surjektiver Gruppenhomomorphismus mit $\ker(\pi_I) = I$.

Satz 11.17 *Sei R ein kommutativer Ring und $I \subset R$ eine additive Untergruppe.*

(a) Es gibt genau dann eine Verknüpfung $*$ auf R/I mit $[a] * [b] = [a \cdot b]$, wenn I ein Ideal ist.

(b) In dem Fall ist $(R/I, +, *)$ ein kommutativer Ring, der Quotientenring von R nach I. *Dann ist $\pi_I : R \to R/I,\ a \mapsto [a]$, ein surjektiver Ringhomomorphismus mit Kern I. Später wird $\cdot$ statt $*$ geschrieben.*

Beweis: (a) $\Rightarrow$: Sei $*$ eine Verknüpfung auf R/I mit $[a] * [b] = [a \cdot b]$. Es soll gezeigt werden, dass $\pi_I : R \to R/I$ ein Ringhomomorphismus ist. Dann folgt mit Lemma 11.15 (a), dass I ein Ideal ist.
Die Eigenschaften $[a] + [b] = [a + b]$ und $[a] * [b] = [a \cdot b]$ eines Ringhomomorphismus sind schon bekannt. Es bleibt zu zeigen, dass R/I ein Ring ist.
Die Assoziativität von $*$ auf R/I folgt aus der Assoziativität von $\cdot$ auf R:

$$([a] * [b]) * [c] = [ab] * [c] = [(ab)c] = [a(bc)] = [a] * [bc] = [a] * ([b] * [c]).$$

Weil $*$ kommutativ ist, bleibt nur ein Distributivgesetz zu zeigen. Das folgt aus dem Distributivgesetz für R:

$$([a] + [b]) * [c] = [a + b] * [c] = [(a + b)c] = [ac + bc] = [ac] + [bc] = [a] * [c] + [b] * [c].$$

$\Leftarrow$: Sei $I \subset R$ ein Ideal. Sei $[a] = [c]$ und $[b] = [d]$, also $a = c + u_1$ und $b = d + u_2$ mit $u_1, u_2 \in I$. Dann ist $ab = cd + (u_1 d + cu_2 + u_1 u_2) \in cd + I$, also $[ab] = [cd]$. Daraus folgt, dass die Klasse $[ab] \in R/I$ unabhängig von der Wahl der Repräsentanten a von $[a]$ und b und $[b]$ ist. Daher gibt der Ansatz $[a] * [b] := [ab]$ eine wohldefinierte Verknüpfung.

(b) Das ist schon im Beweis von (a) gezeigt worden. □

Satz 11.18 *Seien R und S kommutative Ringe, und sei $f : R \to S$ ein Ringhomomorphismus. Weil* $\ker(f)$ *ein Ideal ist (Lemma 11.15 (a)), ist* $R/\ker(f)$ *ein kommutativer Ring (Satz 11.17). Auch $f(R)$ ist ein kommutativer Ring.*
Die Vorschrift $[a] \mapsto f(a)$ definiert eine Abbildung

$$\widetilde{f} : R/\ker(f) \to f(R).$$

$\widetilde{f}$ ist ein Ringisomorphismus.

Beweis: Von Satz 11.10 wissen wir schon, daß $\widetilde{f}$ wohldefiniert und ein Isomorphismus additiver Gruppen ist. Es bleibt zu zeigen, daß (i) $f(R)$ ein Unterring von S ist und daß (ii) $\widetilde{f}$ die Multiplikation in $R/\ker(f)$ auf die Multiplikation in S abbildet.

Zu (i): Es reicht zu zeigen, daß $f(R)$ bezüglich der Multiplikation in S abgeschlossen ist. Das ist aber offensichtlich: $f(a) \cdot f(b) = f(ab) \in f(R)$.

Zu (ii):

$$\widetilde{f}([a][b]) \overset{\text{Satz 11.17}}{=} \widetilde{f}([ab]) \overset{\text{Def. von } \widetilde{f}}{=} f(ab) \overset{f \text{ Ringhom.}}{=} f(a)f(b) \overset{\text{Def. von } \widetilde{f}}{=} \widetilde{f}([a]) \cdot \widetilde{f}([b]). \quad \square$$

Mit der Quotientenkonstruktion in Satz 11.17 kann man aus bekannten Ringen und Idealen in ihnen neue Ringe als Quotientenringe erstellen. Es ist eine kraftvolle Konstruktion.
Der Teil (a) des nächsten Satzes 11.20 wird zeigen, wann so ein Quotientenring ein Körper ist. Teil (b) wird den Fall betrachten, wo der Ausgangsring ein Polynomring $K[t]$ ist.
Den Fall werden wir benutzen, um den größeren Teil des bisher ziemlich geheimnisvolllen Satzes 2.11 zu beweisen, der behauptet hat, dass es zu jeder Primzahlpotenz $q = p^l$ (p Primzahl, $l \in \mathbb{N}$) bis auf Isomorphie genau einen Körper K mit $|K| = q$ gibt. Der war dann $\mathbb{F}_q$ genannt worden.

Lemma/Definition 11.19 *Sei R ein kommutativer Ring.*

(a) (Lemma) Sind $I_1 \subset R$ und $I_2 \subset R$ zwei Ideale, so ist nach der Notation 11.4 (i) ihre Summe $I_1 + I_2 = \{a + b \,|\, a \in I_1, b \in I_2\}$. Die Summe ist wieder ein Ideal.

(b) (Definition) Ein Ideal $I \subset R$ ist ein maximales Ideal*, falls $I \subsetneqq R$ ist und falls kein Ideal $J \subset R$ mit $I \subsetneqq J \subsetneqq R$ existiert.*

(c) (Triviales Lemma) Ein Ideal $I \subsetneqq R$ ist genau dann maximal, wenn für jedes $a \in R - I$ gilt: $(a) + I = R$.

Beweis: (a) Die Idealeigenschaft von $I_1 + I_2$ ist ziemlich klar: bei $a \in I_1$, $b \in I_2$ und $c \in R$ ist $ca \in I_1$, $cb \in I_2$, also $c(a + b) = ca + cb \in I_1 + I_2$. Der Rest ist noch klarer. (b) Definition. (c) Klar. $\square$

Satz 11.20 *(a) Sei R ein kommutativer Ring mit Eins, und sei $I \subsetneqq R$ ein Ideal. Der Quotientenring R/I ist genau dann ein Körper, wenn I ein maximales Ideal ist.*

(b) Ein Ideal $(f(t)) \subsetneqq K[t]$ im Polynomring $K[t]$ über einem Körper K ist genau dann ein maximales Ideal, wenn $f(t)$ ein irreduzibles Polynom ist.

Beweis: (a) $\Rightarrow$: Sei R/I ein Körper. Wegen Lemma 11.19 (c) reicht es zu zeigen, dass für jedes $a \in R - I$ gilt: $(a) + I = R$. Sei $a \in R - I$. Dann ist $[a] \in R/I - \{0\}$. Daher ist $[a]$ invertierbar. Also gibt es ein $b \in R$ mit $[a][b] = 1_{R/I} = [1_R]$. Daher ist $1_R \in ab + I \subset (a) + I$. Daher ist $R = (a) + I$.

$\Leftarrow$: Sei I ein maximales Ideal. R/I ist ein kommutativer Ring mit Einselement $1_{R/I} = [1_R]$. Zu zeigen ist, dass jedes Element $[a] \in R/I - \{0\}$ invertierbar ist. Sei $[a] \in R/I - \{0\}$. Dann ist $a \in R - I$. Weil I ein maximales Ideal ist, ist $(a) + I = R$. Also gibt es es ein $b \in R$ und ein $u \in I$ mit $ab + u = 1_R$. Es folgt $[a][b] = [1_R] = 1_{R/I}$. Also ist $[a]$ invertierbar, und das Inverse ist $[b]$.

(b) Jedes Ideal in $K[t]$ ist ein Hauptideal. Sei $(f) \subsetneqq K[t]$ ein Ideal ungleich $\{0\}$. Dann ist $f \in K[t] - K$, also $\deg f \geq 1$.

$\Rightarrow$: Sei (f) ein maximales Ideal. Wäre $f(t)$ reduzibel, so gäbe es Polynome f_1 und f_2 mit $f = f_1 f_2$ und $1 \leq \deg f_1 < \deg f$ und $1 \leq \deg f_2 < \deg f$. Dann wäre $(f) \subsetneqq (f_1) \subsetneqq K[t]$, also (f) kein maximales Ideal, ein Widerspruch.

$\Leftarrow$: Sei f irreduzibel. Zu zeigen ist wegen Lemma 11.19 (c), dass für jedes $g \in K[t] - (f)$ gilt: $(g) + (f) = K[t]$. Sei $g \in K[t] - (f)$. Weil f irreduzibel ist und weil g wegen $g \notin (f)$ kein Vielfaches von f ist, ist $\mathrm{ggT}(f, g) = 1$. Wegen Lemma 10.16 ist dann 1 eine Linearkombination von f und g, also $1 \in (f, g) = (g) + (f)$. Daraus folgt $(g) + (f) = K[t]$. □

Bemerkungen/Beispiele 11.21 (i) Satz 11.20 (a) und (b) zusammen erlauben es, viele interessante Körper zu konstruieren. Voraussetzung ist eine gute Kontrolle darüber, welche Polynome in $K[t]$ irreduzibel sind. Man kann sich da auf die unitären Polynome beschränken. In (ii) bis (vii) kommen nun Bemerkungen zu den wichtigsten Fällen.

(ii) $K = \mathbb{C}$. Die einzigen irreduziblen unitären Polynome in $\mathbb{C}[t]$ sind die linearen Polynome, also die Polynome $t - a$ mit $a \in \mathbb{C}$ (Bemerkung 10.21 (iv)). Mit ihnen erhält man keinen neuen Körper, denn $\mathbb{C}[t]/(t - a) \cong \mathbb{C}$.

(iii) $K = \mathbb{R}$. Die einzigen irreduziblen unitären Polynome sind die linearen Polynome, also die Polynome $t - a$ mit $a \in \mathbb{R}$, und die quadratischen Polynome ohne reelle Nullstellen, also die Polynme $t^2 + at + b$ mit $a, b \in \mathbb{R}$ und $a^2 - 4b < 0$ (Bemerkung 10.21 (iv)). Auch hier erhält man keine neuen Körper, denn

$$\frac{\mathbb{R}[t]}{(t-a)} \cong \mathbb{R}, \quad \frac{\mathbb{R}[t]}{(t^2+at+b)} \cong \mathbb{C}.$$

Sind $\lambda_1, \lambda_2 \in \mathbb{C} - \mathbb{R}$ (mit $\overline{\lambda_1} = \lambda_2$) die Nullstellen von $t^2 + at + b$, so hat man in der zweiten Gleichung die Wahl zwischen den beiden Körperisomorphismen für $i = 1$ oder $i = 2$ mit

$$\frac{\mathbb{R}[t]}{(t^2+at+b)} \to \mathbb{C}, \quad 1 \mapsto 1, \ [t] \mapsto \lambda_i.$$

(iv) (Ausblick in die Algebra und algebraische Zahlentheorie) $K = \mathbb{Q}$. Hier gibt viele irreduzible Polynome beliebigen Grades ≥ 1. Sei $f(t) \in \mathbb{Q}[t]$ ein irreduzibles unitäres Polynom vom Grad n. In $\mathbb{C}[t]$ zerfällt es in Linearfaktoren $f(t) = (t - \lambda_1)...(t - \lambda_n)$ mit $\lambda_1, ..., \lambda_n \in \mathbb{C}$. Sei $f' := df/dt$. Wegen $\deg f'(t) = \deg f(t) - 1$ ist $\mathrm{ggT}(f, f') = 1$.

Daher sind $\lambda_1, ..., \lambda_n$ keine Nullstellen von f'. Daher sind sie alle verschieden. Man hat für jedes $i = 1, ..., n$ einen Körperisomorphismus

$$\begin{array}{rcl} \dfrac{\mathbb{Q}[t]}{(f(t))} & \to & \mathbb{Q}[\lambda_i] := \mathbb{Q} \oplus \mathbb{Q}\lambda_i \oplus \mathbb{Q}\lambda_i^2 \oplus ... \oplus \mathbb{Q}\lambda_i^{n-1} \subset \mathbb{C}, \\ [t] & \mapsto & \lambda_i. \end{array}$$

Die Körper $\mathbb{Q}[t]/(f(t))$ und $\mathbb{Q}[\lambda_i]$ sind $\mathbb{Q}$-Vektorräume der Dimension n. Die Körper $\mathbb{Q}[\lambda_i] \subset \mathbb{C}$ heißen *Zahlkörper*.

(v) (Ausblick in die Algebra) $K = \mathbb{F}_p$. Hier ist p eine Primzahl, und der endliche Körper $\mathbb{F}_p \cong \mathbb{Z}_p \cong \mathbb{Z}/p\mathbb{Z}$ ist seit Satz 2.9 bekannt. Sei

$$J_{n,p} := \{f(t) \in \mathbb{F}_p[t] \mid \deg f(t) = n,\ f(t) \text{ ist irreduzibel und unitär}\}.$$

Man weiß, dass $J_{n,p} \neq \emptyset$ für alle n und p ist. Man hat sogar gute Kontrolle über die Anzahlen $|J_{n,p}|$. Beispiel 10.22 gibt die Zahlen $|J_{1,2}| = 2, [J_{2,2}| = 1, |J_{3,2}| = 2$ in der folgenden Tabelle.

$n =$	1	2	3	4	5	6
$\lvert J_{n,2}\rvert =$	2	1	2	3	6	9
$\lvert J_{n,3}\rvert =$	3	3	8	18	48	116
$\lvert J_{n,5}\rvert =$	5	10	40	125	624	2580

Bei $f \in J_{n,p}$ ist $\mathbb{F}_p[t]/(f)$ ein $\mathbb{F}_p$-Vektorraum der Dimension n und ein Körper mit p^n Elementen. Das gibt die Existenz eines solchen Körpers in Satz 2.11. Beispiel 10.22 gibt uns die Körper

$$\mathbb{F}_4 = \frac{\mathbb{F}_2[t]}{(t^2+t+1)} \quad \text{und} \quad \mathbb{F}_8 = \frac{\mathbb{F}_2[t]}{(t^3+t+1)}$$

mit 4 bzw. 8 Elementen.

(vi) (Fortsetzung von (v)) Es ist auch im Prinzip klar (aber gewöhnungsbedürftig), wie man in den so als Quotientenringen gewonnenen Körpern rechnet: Man muß immer modulo $f(t)$ rechnen. Im Beispiel von $\mathbb{F}_4$ sieht das so aus. Man setzt $\alpha := [t]$. Dann ist

$$\begin{array}{rcl} \mathbb{F}_4 & = & \mathbb{F}_2 \cdot 1 \oplus \mathbb{F}_2 \cdot \alpha \quad \text{als } \mathbb{F}_2\text{-Vektorraum} \\ & = & \{0, 1, \alpha, 1+\alpha\}, \\ \text{und } \alpha^2+\alpha+1 & = & 0, \text{ also } \alpha^2 = -\alpha - 1 = \alpha + 1, \\ \text{also } \alpha(\alpha+1) & = & 1, \quad (\alpha+1)^2 = \alpha. \end{array}$$

Das gibt die Additionstafel und die Multiplikationstafel

$+$	0	1	α	$1+\alpha$
0	0	1	α	$1+\alpha$
1	1	0	$1+\alpha$	α
α	α	$1+\alpha$	0	1
$1+\alpha$	$1+\alpha$	α	1	0

$\cdot$	0	1	α	$1+\alpha$
0	0	0	0	0
1	0	1	α	$1+\alpha$
α	0	α	$1+\alpha$	1
$1+\alpha$	0	$1+\alpha$	1	α

(vii) (Fortsetzung von (v) und (vi)) Warum ist jeder Körper K mit p^n Elementen zu jedem Körper $\mathbb{F}_p[t]/(f)$ mit $f \in J_{n,p}$ isomorph? Seit Aufgabe 4 von Blatt 5 im HWS 19 wissen wir, dass die Charakteristik von K gleich p ist und er ein $\mathbb{F}_p$-Vektorraum der Dimension n ist.
Es gelten die beiden folgenden bemerkenswerten Formeln,

$$\begin{aligned} t^{p^n} - t &= \prod_{a \in K} (t-a) \qquad \text{in } K[t], \\ t^{p^n} - t &= \prod_{r|n} \prod_{g \in J_{r,p}} g(t) \qquad \text{in } \mathbb{F}_p[t]. \end{aligned}$$

Der Beweis der zweiten Formel ist nicht sehr lang, aber er braucht mit Vorbereitungen schon mehrere Seiten und wird hier nicht gegeben. Die erste Formel folgt so: $(K^*, \cdot)$ ist eine multiplikative Gruppe der Ordnung $p^n - 1$. Daher und wegen des Satzes von Lagrange, Satz 11.11, erfüllen alle Elemente $a \in K^*$ die Gleichung $a^{p^n-1} = 1$. Die 0 erfüllt $0^{p^n} - 0 = 0$. Daher sind alle p^n Elemente von K Nullstellen des Polynoms $t^{p^n} - t$. Daher zerfällt es in die Linearfaktoren $t - a$ mit $a \in K$.
Wegen der zweiten Formel ist jedes $f \in J_{n,p}$ ein Teiler von $t^{p^n} - t$. Daher und wegen der ersten Formel enthält K n Nullstellen von f. Sei $\alpha \in K$ eine solche Nullstelle. Daher hat man den Körperisomorphismus

$$\frac{\mathbb{F}_p[t]}{(f(t))} \xrightarrow{\cong} K = \mathbb{F}_p[\alpha] = \mathbb{F}_p 1 \oplus \mathbb{F}_p \alpha \oplus ... \oplus \mathbb{F}_p \alpha^{n-1}, \quad [t] \mapsto \alpha.$$

Daher gibt es bis auf Isomorphie nur einen Körper mit p^n Elementen. Der wird $\mathbb{F}_{p^n}$ genannt.

Am Ende dieses Unterkapitels werden noch zwei Versionen eines Satzes gegeben, der nicht Körper betrifft, sondern Quotientenringe von Hauptidealringen.

Bemerkung 11.22 Seien $R_1, ..., R_n$ kommutative Ringe mit Eins. Dann ist das Produkt $R := R_1 \times ... \times R_n$ mit der komponentenweisen Addition und Multiplikation ein kommutativer Ring mit Einselement $(1_{R_1}, ..., 1_{R_n})$ ist. Man sieht auch leicht, dass die Einheitengruppe R^* von R gerade

$$R^* = R_1^* \times ... \times R_n^*$$

ist.

Satz 11.23 *(Chinesischer Restsatz, moderne Version für beliebige Hauptidealringe) Es sei R ein Hauptidealring, und $a_1, ..., a_k \in R - \{0\}$ seien paarweise teilerfremde Elemente, also $\mathrm{ggT}(a_i, a_j) = 1$ für $i \neq j$. Es sei $a_0 := a_1 \cdot ... \cdot a_k$. Die Klasse von $b \in R$ im Quotientenring $R/(a_j)$ für $j \in \{0, 1, ..., k\}$ wird mit $[b]_j$ bezeichnet.*
Es gibt wohldefinierte Ringhomomorphismen $\pi_i : R/(a_0) \to R/(a_i), [b]_0 \mapsto [b]_i$, für $i \in \{1, ..., k\}$. Zusammen geben sie einen Ringisomorphismus

$$\pi : \frac{R}{(a_0)} \to \frac{R}{(a_1)} \times ... \times \frac{R}{(a_k)}, \quad [b]_0 \mapsto ([b]_1, ..., [b]_k).$$

Beweis: Ist $[b]_0 = [c]_0$, so gibt es ein $d \in R$ mit $b = c + d \cdot a_0$. Wegen $a_i | a_0$ ist dann auch $[b]_i = [c]_i$ für $i \in \{1, ..., k\}$. Daher ist π_i eine wohldefinierte Abbildung. Dass es eine Ringhomomorphismus ist, ist ziemlich offensichtlich. Daher ist auch $\pi = \pi_1 \times ... \times \pi_k$ ein Ringhomomorphismus.

π ist injektiv: Zu zeigen ist $\ker \pi = \{0\}$. Sei $[b]_0 \in \ker \pi$. Dann ist $[b]_i = \pi_i([b]_0) = 0$, also $b_i \in (a_i)$, also $a_i | b$. Wegen $\mathrm{ggT}(a_i, a_j) = 1$, und weil man im Hauptidealring eine eindeutige Primfaktorzerlegung hat, gilt $a_0 | b$, also $b \in (a_0)$, also $[b]_0 = 0$. Daher ist $\ker \pi = \{0\}$.

π ist surjektiv: Man sucht ein Urbild unter π eines beliebigen Elementes

$$([c_1]_1, ..., [c_k]_k) \in \frac{R}{(a_1)} \times ... \times \frac{R}{(a_k)}.$$

Sei $A_i := \prod_{j \neq i} a_j$, also $A_i \cdot a_i = a_0$. Es gilt $\mathrm{ggT}(A_i, a_i) = 1$. Daher gibt es $d_i, e_i \in R$ mit $1 = d_i A_i + e_i a_i$, also $[1]_i = [d_i A_i]_i$. Man wählt

$$b := \sum_{j=1}^{k} d_j A_j c_j \in R.$$

Dann ist wegen $[A_j]_i = 0$ für $j \neq i$

$$\pi_i([b]_0) = [b]_i = \sum_j [d_j A_j c_j]_i = [d_i A_i]_i [c_i] = [1]_i [c_i] = [c_i].$$

Also ist $\pi([b]_0) = ([c_1]_1, ..., [c_k]_k)$. □

Satz 11.24 *(Chinesischer Restsatz, klassische Version für $\mathbb{Z}$)*
Es seien $n_1, ..., n_k \in \mathbb{N}$ mit $\mathrm{ggT}(n_i, n_j) = 1$ für $i \neq j$, und es seien $c_1, ..., c_k \in \mathbb{Z}$. Gesucht sind Lösungen $b \in \mathbb{Z}$ der Gleichungen

$$b \equiv c_i \mod n_i \quad \text{für alle } i = 1, ..., k.$$

(i) Die Lösungsmenge ist nicht leer.

(ii) Ist $b \in \mathbb{Z}$ eine Lösung, so ist $\widetilde{b} \in \mathbb{Z}$ genau dann eine Lösung, wenn $\widetilde{b} \equiv b \mod (n_1 \cdot ... \cdot n_k)$ ist.

Beweis: Man wendet Satz 11.23 im Fall $R = \mathbb{Z}$ und $n_i = a_i$ an. Ein Element $b \in \mathbb{Z}$ ist genau dann eine Lösung der Gleichungen, wenn es $\pi([b]_0) = ([c_1]_1, ..., [c_k]_k)$ erfüllt. Weil π surjektiv ist, gibt es Lösungen. Ist b eine Lösung, so ist $\widetilde{b}$ genau dann eine Lösung, wenn $[b]_0 = [\widetilde{b}]_0$ ist, denn π ist ein Isomorphismus. □

Beispiele 11.25 (i) Genau die Zahlen in $\{..., 2, 5, 8, 11, 14, 17, ...\} = 2 + 3\mathbb{Z}$ erfüllen die Gleichung

$$b \equiv 2 \mod 3.$$

(ii) Gesucht ist eine Zahl $b \in \mathbb{Z}$ mit

$$\begin{aligned} b &\equiv 2 \mod 3, \\ b &\equiv 4 \mod 5. \end{aligned}$$

Die Zahlen oben, $2, 5, 8, 11, 14, 17, ...$, die ja $b \equiv 2 \mod 3$ erfüllen, haben modulo 5 die Reste $2, 0, 3, 1, 4, 2, ...$ Daher tun's $b = 14$ und jede Zahl in $14 + 15\mathbb{Z}$.

(iii) Gesucht ist eine Zahl $b \in \mathbb{Z}$ mit

$$\begin{aligned} b &\equiv 2 \mod 3, \\ b &\equiv 4 \mod 5, \\ b &\equiv 3 \mod 7. \end{aligned}$$

Die Zahlen $14, 14+15, 14+2\cdot 15, 14+3\cdot 15, ...$ erfüllen die ersten beiden Gleichungen und haben modulo 7 die Reste $0, 1, 2, 3, ...$ Daher tun's $b = 14 + 3 \cdot 15 = 59$ und jede Zahl in $59 + 105\mathbb{Z}$.
Wie findet man b? Ein Weg ist durch induktives Probieren wie oben. Man startet mit $b_0 = 14$, das die ersten beiden Gleichungen löst, und addiert Vielfache von 15, bis man auch modulo 7 den richtigen Rest hat. Wegen $14 \equiv 0 \mod 7$ und $15 \equiv 1 \mod 7$ ist es einfach, $b = 14 + 3 \cdot 15 = 59$ zu finden.
Ein anderer Weg ist, systematisch dem Beweis des chinesischen Restsatzes zu folgen:

i	a_i	c_i	A_i	d_i
1	3	2	35	2
2	5	4	21	1
3	7	3	15	1

Hier ist $A_i = \prod_{j \neq i} a_j$, und d_i ist so gewählt, dass $d_i A_i \equiv 1 \mod a_i$ ist. Dann ist eine Lösung

$$b_1 = \sum_i d_i A_i c_i = 2 \cdot 35 \cdot 2 + 1 \cdot 21 \cdot 4 + 1 \cdot 15 \cdot 3 = 140 + 84 + 45 = 2 \cdot 105 + 59,$$

und die Menge aller Lösungen ist die Menge

$$b_1 + (\prod_i a_i)\mathbb{Z} = 2 \cdot 105 + 59 + 105\mathbb{Z} = 59 + 105\mathbb{Z}.$$

Korollar 11.26 *Die Eulersche φ-Funktion (siehe auch Aufgabe 3 von Blatt 2)*

$$\begin{aligned} \varphi : \mathbb{N} &\to \mathbb{N}, \\ m &\mapsto |\mathbb{Z}_m^*| = |(\mathbb{Z}/m\mathbb{Z})^*| \end{aligned}$$

erfüllt

$$\begin{aligned} \varphi(p^k) &= (p-1) \cdot p^{k-1}, && \text{falls } p \text{ eine Primzahl ist,} \\ \varphi(n_1 \cdot n_2) &= \varphi(n_1) \cdot \varphi(n_2), && \text{falls } \mathrm{ggT}(n_1, n_2) = 1 \text{ ist,} \\ \varphi(p_1^{k_1} \cdot ... \cdot p_l^{k_l}) &= (p_1 - 1)p_1^{k_1 - 1} \cdot ... \cdot (p_l - 1)p_l^{k_l - 1}, \\ & && \text{falls } p_1, ..., p_l \text{ verschiedene Primzahlen sind.} \end{aligned}$$

Beweis: 1. Gleichung: Aufgabe 3 (a) von Blatt 2.
2. Gleichung: Aus der Ring-Version des chinesischen Restsatzes folgt

$$\begin{aligned} \mathbb{Z}/(n_1 \cdot n_2)\mathbb{Z} &\cong (\mathbb{Z}/n_1\mathbb{Z}) \times (\mathbb{Z}/n_2\mathbb{Z}), \\ (\mathbb{Z}/(n_1 \cdot n_2)\mathbb{Z})^* &\cong (\mathbb{Z}/n_1\mathbb{Z})^* \times (\mathbb{Z}/n_2\mathbb{Z})^*, \\ \varphi(n_1 \cdot n_2) &= \varphi(n_1) \cdot \varphi(n_2). \end{aligned}$$

3. Gleichung: Sie folgt aus der 1. und der 2. Gleichung. □

11.4 Quotienten bei Vektorräumen

Die Quotientenkonstruktion bei Vektorräumen birgt nach den Quotientenkonstruktionen bei Gruppen und bei kommutativen Ringen keine Überraschungen mehr. Im Gegenteil. Vektorräume sind viel einfachere Objekte als Gruppen oder Ringe (die Isomorphieklasse eines Vektorraums ist durch den zugrundeliegenden Körper und seine Dimension bestimmt). Auch die Quotientenkonstruktion liefert nur wieder Vektorräume.
Die generelle Schwierigkeit, Äquivalenzklassen als eigenständige Objekte zu behandeln, mit denen man wie mit Zahlen oder wie mit Vektoren im $\mathbb{R}^n$ rechnen kann, ist eine Schwierigkeit des 1. Semesters (und einiger Mathematiker im 19. Jahrhundert). Nach den vorherigen drei Unterkapiteln sollte sie nicht mehr auftreten.

Satz 11.27 *Sei V ein K-Vektorraum und U ein Untervektorraum. Nach Satz 11.9 ist V/U eine additive abelsche Gruppe. Die Äquivalenzklasse von $v \in V$ wird als $[v] \in V/U$ geschrieben.*
Die Abbildung $K \times V/U \to V/U$, $(\lambda, [v]) \mapsto [\lambda \cdot v]$, definiert eine skalare Multiplikation auf V/U und macht V/U zu einem K-Vektorraum.
Die Projektion $\pi_U : V \to V/U, v \mapsto [v]$, ist eine surjektive lineare Abbildung. Ihr Kern ist $\ker(\pi_U) = U$. Im Fall von $\dim U < \infty$ ist $\dim V/U = \dim V - \dim U$.

Beweis: Der größere Teil folgt aus Satz 11.9. Der Rest ist eine Übung. □

Der nächste Satz ergänzt Satz 11.27.

Satz 11.28 *(Homomorphiesatz für Vektorräume)*

(a) Sei $f : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen. Weil $\ker(f)$ ein Untervektorraum ist (Satz 5.3 (c)), ist $V/\ker(f)$ ein Vektorraum (Satz 11.26). $f(V)$ ist ein Untervektorraum von W (Satz 5.3 (c)).
Der Gruppenisomorphismus

$$\widetilde{f} : V/\ker(f) \to f(V), \quad [v] \mapsto f(v),$$

von Satz 11.10 ist ein Vektorraumisomorphismus.

(b) Sei $f : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen, und sei $U \subset V$ ein Untervektorraum. Es gibt genau dann eine lineare Abbildung $\overline{f} : V/U \to W$ mit $f = \overline{f} \circ \pi_U$ wenn $U \subset \ker(f)$ ist.

Beweis: (a) Der größere Teil folgt aus Satz 11.10. Der Rest ist eine Übung.
(b) Das folgt leicht mit Lemma 11.3 und der Linearität der beteiligten Abbildungen. Die Details sind eine Übung. □

Bemerkung 11.29 Sei V ein Vektorraum und U ein Untervektorraum. Dann gibt es immer einen *Komplementärraum*, das ist ein Untervektorraum $\widetilde{U} \subset V$ mit $U \oplus \widetilde{U} = V$. Er ist nur in den uninteressanten Fällen $U = V$ oder $U = \{0\}$ eindeutig.
Aber für jeden solchen Raum $\widetilde{U}$ ist die Einschränkung der Projektion $\pi_U : V \to V/U$ ein Vektorraumisomorphismus $\pi_U : \widetilde{U} \mapsto V/U, \quad v \mapsto [v]$.

12 Jordannormalform

Dieses Kapitel setzt Kapitel 8 fort. Wieder bezeichnet K irgendeinen Körper. Es geht wieder um Normalformen für Endomorphismen. Im Zentrum des Kapitels steht die Jordannormalform. Sie läßt sich herstellen, wenn das charakteristische Polynom in Linearfaktoren zerfällt.
Am Ende des Kapitels kommen Bemerkungen zu einer allgemeineren Normalform.
Im ersten Teil des Kapitels (bis Satz 12.10) werden die Begriffe und Sätze des Kapitels 8 wiederholt.

Definition/Lemma 12.1 *(= 8.1) (a) (Definition) Sei $A \in M(n \times n, K)$. Das* charakteristische Polynom *von A ist das Polynom*

$$\begin{aligned} P_A(t) := \det(t \cdot E_n - A) &= (-1)^n \det(A - t \cdot E_n) \\ &= (-1)^n \cdot \begin{vmatrix} a_{11} - t & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - t & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & a_{nn} - t \end{vmatrix}. \end{aligned}$$

(Lemma) Es ist

$$\begin{aligned} P_A(t) &= t^n - (a_{11} + a_{22} + ... + a_{nn})t^{n-1} \\ &\quad +(...)t^{n-2} + ... + (...)t + (-1)^n \det A \in K[t]. \end{aligned}$$

Es ist also ein unitäres (d.h. Leitkoeffizient 1) Polynom vom Grad n. Der Koeffizient

$$\mathrm{Spur}(A) := a_{11} + a_{22} + ... + a_{nn}$$

von $-t^{n-1}$ heißt Spur *von A.*

(b) (Lemma) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$. Sei $\mathcal{B}$ eine Basis von V. Das Polynom

$$P_f(t) := (-1)^n \det(M(\mathcal{B}, f, \mathcal{B}) - t \cdot E_n) = P_{M(\mathcal{B},f,\mathcal{B})}(t)$$

ist unabhängig von der Wahl der Basis $\mathcal{B}$.
(Definition) Es heißt charakteristisches Polynom *von f.*
Auch die Zahl

$$\mathrm{Spur}(f) := \mathrm{Spur}(M(\mathcal{B}, f, \mathcal{B}))$$

ist unabhängig von der Wahl von $\mathcal{B}$. Sie heißt Spur *von f.*

Beispiele 12.2 ($\subset$ 8.2) (i) (Obere Dreiecksmatrix) Ist $A = (a_{ij}) \in M(n \times n, K)$ eine obere Dreiecksmatrix (also $a_{ij} = 0$ für $i > j$), so ist auch $t \cdot E_n - A$ eine obere Dreiecksmatrix, mit Diagonaleinträgen $t - a_{11}, ..., t - a_{nn}$. Daher ist

$$P_A(t) = \det(t \cdot E_n - A) = \prod_{i=1}^{n} (t - a_{ii}).$$

Ein wichtiger Spezialfall: A eine *Diagonalmatrix*, d.h. $a_{ij} = 0$ für $i \neq j$,

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}.$$

(ii) (Jordanblock) Eine Matrix $A \in M(n \times n, K)$ ist ein *Jordanblock*, falls sie die Gestalt hat

$$A = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}.$$

Das ist auch ein Spezialfall einer oberen Dreiecksmatrix, mit

$$P_A(t) = (t - \lambda)^n.$$

Definition 12.3 (= 8.3) (a) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V. Ein Element $v \in V - \{0\}$ heißt *Eigenvektor* von f, falls es ein $\lambda \in K$ gibt mit

$$f(v) = \lambda \cdot v.$$

Ein solches λ heißt *Eigenwert* von f.

(b) (Lemma) Für jedes $\lambda \in K$ ist die Menge

$$\mathrm{Eig}(f, \lambda) := \{v \in V \mid f(v) = \lambda \cdot v\} = \ker(f - \lambda \cdot \mathrm{id})$$

offenbar ein Untervektorraum von V. Und offenbar ist λ genau dann ein Eigenwert von f, wenn $\mathrm{Eig}(f, \lambda) \neq \{0\}$ ist.
(Definition) Der Vektorraum $\mathrm{Eig}(f, \lambda)$ heißt *Eigenraum* von f bezüglich λ.

(c) Eigenwerte, Eigenvektoren und Eigenräume einer Matrix $A \in M(n \times n, K)$ sind die Eigenwerte, Eigenvektoren und Eigenräume des Endomorphismus

$$l_A : M(n \times 1, K) \to M(n \times 1, K), \quad b \mapsto A \cdot b.$$
$$(l_A = \text{ Linksmultiplikation mit } A)$$

Äquivalent und konkreter: $v \in M(n \times 1, K) - \{0\}$ heißt *Eigenvektor* von A, falls es ein $\lambda \in K$ gibt mit

$$A \cdot v = \lambda \cdot v.$$

Ein solches λ heißt *Eigenwert* von A.

Für *jedes* $\lambda \in K$ ist

$$\begin{aligned} \mathrm{Eig}(A,\lambda) &:= \{v \in M(n\times 1,K) \mid A\cdot v = \lambda\cdot v\} \\ &= \ker(l_A - \lambda\cdot \mathrm{id}) = \mathrm{Lös}(A-\lambda\cdot E_n, 0) \end{aligned}$$

der Eigenraum von A zum Wert λ. Es ist offenbar ein Untervektorraum von $M(n\times 1,K)$. Und offenbar ist $\lambda\in K$ genau dann ein Eigenwert von A, wenn $\mathrm{Eig}(A,\lambda)\neq\{0\}$ ist.

Satz 12.4 *(= 8.4)* **(Bestimmung von Eigenwerten und Eigenvektoren)**
(a) Sei $f: V\to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n\in\mathbb{N}$. Dann gilt folgende Äquivalenz:

$$\lambda\in K \text{ ist ein Eigenwert von } f \iff P_f(\lambda)=0.$$

(b) Sei $A\in M(n\times n,K)$. Dann gilt folgende Äquivalenz:

$$\lambda\in K \text{ ist ein Eigenwert von } A \iff P_A(\lambda)=0.$$

Beispiele 12.5 ($\subset$ 8.5) Es werden dieselben Beispiele wie in 12.2 betrachtet.
(i) (Obere Dreiecksmatrix) Wegen $P_A(t)=\prod_{i=1}^n(t-a_{ii})$ sind die Eigenwerte $a_{11},...,a_{nn}$.
Aber die Eigenvektoren sind schwerer zu bestimmen. Sofort sieht man nur den Eigenvektor $e_1=(1,0...,0)^{tr}$ zum Eigenwert a_{11}. Wenn mehrere a_{ii} übereinstimmen, gibt es zu ihnen eventuell nur einen Eigenvektor, siehe (ii).
Nur im Fall einer Diagonalmatrix sieht man sofort viele Eigenvektoren: e_i ist Eigenvektor zum Eigenwert a_{ii}. Mehr dazu kommt in 12.6.

(ii) (Jordanblock) Hier ist $P_A(t)=(t-\lambda)^n$, also hat man nur einen Eigenwert. Tatsächlich ist hier

$$A-\lambda\cdot E_n = \begin{pmatrix} 0 & 1 & & 0\\ & \ddots & \ddots & \\ & & \ddots & 1\\ 0 & & & 0 \end{pmatrix},$$

also $\mathrm{rang}(A-\lambda\cdot E_n)=n-1$, also $\dim\mathrm{Eig}(A,\lambda)=1$.
Genauer:

$$\mathrm{Eig}(A,\lambda)=\mathrm{Lös}(A-\lambda\cdot E_n,0)=K\cdot e_1.$$

Definition/Lemma 12.6 *(= 8.6) (a) (Definition) Ein Endomorphismus $f: V\to V$ eines K-Vektorraums V mit $\dim V = n\in\mathbb{N}$ heißt* diagonalisierbar*, falls es eine Basis von V aus Eigenvektoren von f gibt.*
(b) (Lemma) Ist $\mathcal{B}=(b_1,...,b_n)$ eine solche Basis mit $f(b_i)=\lambda_i$, so ist

$$M(\mathcal{B},f,\mathcal{B}) = \begin{pmatrix} \lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n, \end{pmatrix}.$$

(c) (Definition) Eine Matrix $A \in M(n \times n, K)$ heißt diagonalisierbar*, falls es es eine Basis von $M(n \times 1, K)$ aus Eigenvektoren von A gibt.*
(d) (Lemma) Eine Matrix A ist diagonalisierbar genau dann, wenn es eine Matrix $B \in GL(n, K)$ gibt mit

$$B^{-1} \cdot A \cdot B = \text{ Diagonalmatrix.}$$

Bemerkung 12.7 ($\sim$ 8.7) Wenn ein Endomorphismus diagonalisierbar ist, ist das etwas ganz besonderes. Viele Endomorphismen lassen sich nicht diagonalisieren.
Aber auch dann gibt es *Normalformen*. Das sind Matrizen von besonders guter Gestalt, die nach Wahl von besonders guten Basen die Endomorphismen repräsentieren. Wenn das charakteristische Polynom in Linearfaktoren zerfällt, hat man die *Jordannormalform*. Sie ist das Hauptthema dieses Kapitels. Aber auch, wenn es nicht zerfällt, hat man eine interessante Normalform, die die Jordannormalform verallgemeinert. Dazu kommen Bemerkungen am Ende des Kapitels.

Satz/Definition 12.8 *($\supset$ 12.8) (a) (Notation) Seien $V_i \subset V$ $(i = 1, ..., k)$ Untervektorräume eines Vektorraums V. Der von ihnen erzeugte Untervektorraum von V ist*

$$\sum_{i=1}^{k} V_i := \{v_1 + ... + v_k \mid v_i \in V_i\}.$$

(b) (Satz) Folgende drei Bedingungen sind äquivalent:

- *(α) Jedes Element von $\sum_{i=1}^k V_i$ lässt sich auf eindeutige Weise als Linearkombination von Elementen der V_i schreiben, d.h.*

$$v_1 + ... + v_k = \widetilde{v}_1 + ... + \widetilde{v}_k \text{ mit } v_i, \widetilde{v}_i \in V_i \Rightarrow \text{ für alle } i \quad v_i = \widetilde{v}_i.$$

- *(β) Beliebige Basen der V_i bilden zusammen eine Basis von $\sum_{i=1}^k V_i$.*
- *(γ) Für alle $i = 1, ..., k$ ist $V_i \cap (\sum_{j \neq i} V_j) = \{0\}$.*

(c) (Definition) Wenn die drei äquivalenten Bedingungen in (b) erfüllt sind, ist der von den V_i erzeugte Untervektorraum die direkte Summe *der V_i; Notation: $\bigoplus_{i=1}^k V_i$ oder $V_1 \oplus ... \oplus V_k$.*
(d) (Neuer Satz) Sei V ein Vektorraum und $U \subset V$ ein Untervektorraum. Dann gibt es einen Untervektorraum $\widetilde{U}$ mit $U \oplus \widetilde{U} = V$. Er ist nur in den Fällen $U = V$ und $U = \{0\}$ eindeutig (dann ist er $\widetilde{U} = \{0\}$ bzw. $\widetilde{U} = V$).
(Definition) Er heißt Komplementärraum*.*

Beweis des neuen Teils (d): Existenz im Spezialfall $\dim V < \infty$: mit dem Basisergänzungssatz Satz 3.19. Man ergänzt eine Basis $\mathcal{B}_1$ von U mit einem geeigneten Tupel $\mathcal{B}_2$ von Vektoren zu einer Basis von V. Dann kann man $\widetilde{U} := \text{span } \mathcal{B}_2$ wählen. Existenz im allgemeinen Fall: hier ohne Beweis. Es ist verwandt zur Existenz von Basen im allgemeinen Fall.

Nicht-Eindeutigkeit: Sei $\{0\} \subsetneqq U \subsetneqq V$, und sei $\widetilde{U}$ ein Komplementärraum zu U. Dann ist auch $\{0\} \subsetneqq \widetilde{U} \subsetneqq V$. Sei $u_0 \in U - \{0\}$ und sei $(\widetilde{u}_j)_{j\in J}$ mit $1 \in J$ eine Basis von $\widetilde{U}$. Dann ist $\operatorname{span}(\widetilde{u}_1 + u_0; \widetilde{u}_j, j \in J - \{1\})$ auch ein Komplementärraum und $\neq \widetilde{U}$. □

Definition 12.9 (= 8.10) (a) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V. Ein Element $v \in V - \{0\}$ heißt *verallgemeinerter Eigenvektor* oder *Hauptvektor*, falls es ein $\lambda \in K$ und ein $m \in \mathbb{N}$ gibt mit

$$(f - \lambda \cdot \mathrm{id})^m(v) = 0.$$

(b) Für jedes $\lambda \in K$ ist die Menge

$$\mathrm{Hau}(f, \lambda) := \{v \in V \mid \text{ es gibt ein } m \in \mathbb{N} \text{ mit } (f - \lambda \cdot \mathrm{id})^m(v) = 0\}$$

offenbar ein Untervektorraum von V. Er heißt *Hauptraum* von f bezüglich λ. Offenbar ist $\mathrm{Hau}(f, \lambda) \supset \mathrm{Eig}(f, \lambda)$. Es gilt

$$\lambda \text{ Eigenwert} \iff \mathrm{Eig}(f, \lambda) \neq \{0\} \iff \mathrm{Hau}(f, \lambda) \neq \{0\}.$$

Die erste Äquivalenz und die Richtung $\Rightarrow$ in der zweiten Äquivalenz sind klar. $\mathrm{Eig}(f, \lambda) \neq \{0\} \Leftarrow \mathrm{Hau}(f, \lambda) \neq \{0\}$ folgt so: Sei $v \in \mathrm{Hau}(f, \lambda) - \{0\}$, und sei $m \in \mathbb{N}$ minimal mit $(f - \lambda \cdot \mathrm{id})^m(v) = 0$. Dann ist $(f - \lambda \cdot \mathrm{id})^{m-1}(v) \in \mathrm{Eig}(f, \lambda) - \{0\}$.

Der folgende Satz wurde in Kapitel 8 nicht bewiesen. Er wird nach Lemma 12.11 bewiesen. Er ist ein wichtiger Schritt zur Jordannormalform.

Satz 12.10 *(=8.11) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums mit $\dim_K V = n \in \mathbb{N}$. Sein charakteristisches Polynom zerfalle in Linearfaktoren,*

$$P_f(t) = \prod_{i=1}^{k} (t - \lambda_i)^{d_i}$$

mit λ_i paarweise verschieden und $d_i \in \mathbb{N}$.
Dann gilt:

$$V = \bigoplus_{i=1}^{k} \mathrm{Hau}(f, \lambda_i) \quad \text{und} \quad \dim \mathrm{Hau}(f, \lambda_i) = d_i.$$

Hier endet die Wiederholung zu Kapitel 8.

Lemma 12.11 *Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim V = n \in \mathbb{N}$. Sei $\lambda \in K$ und $d(\lambda) := \dim \mathrm{Hau}(f, \lambda)$. Dann ist*

$$\mathrm{Hau}(f, \lambda) = \ker(f - \lambda \cdot \mathrm{id})^m \quad \textit{für jedes} \quad m \geq d(\lambda).$$

Beweis: Das Problem ist zu sehen, daß die $d(\lambda)$-te Potenz reicht. Sei $H_j := \ker(f - \lambda \cdot \mathrm{id})^j \subset V$. Es ist $\{0\} = H_0 \subset H_1 \subset H_2 \subset ....$ Wegen $\dim \mathrm{Hau}(f, \lambda) = d(\lambda)$ gibt es ein $l \leq d(\lambda)$ mit $H_l = H_{l+1}$.
Behauptung: $H_l = H_{l+1} = H_{l+2} = ...$
Annahme: $v \in H_{l+2} - H_{l+1}$.
Dann ist $(f - \lambda \cdot \mathrm{id})(v) \in H_{l+1} - H_l$. Widerspruch.
Also ist $H_{l+1} = H_{l+2}$. Die Behauptung folgt induktiv. Also ist

$$\mathrm{Hau}(f, \lambda) \stackrel{\mathrm{Def}}{=} \bigcup_{j \in \mathbb{N}} H_j = H_l = H_{d(\lambda)} = \ker(f - \lambda \cdot \mathrm{id})^{d(\lambda)}. \qquad \square$$

Beweis von Satz 12.10:
Der Beweis wird induktiv nach $\dim V$ geführt. Nach Lemma 12.11 ist

$$U := \mathrm{Hau}(f, \lambda_1) \stackrel{!}{=} \ker(f - \lambda_1 \cdot \mathrm{id})^n.$$

Sei

$$W := \mathrm{Bild}(f - \lambda_1 \cdot \mathrm{id})^n = (f - \lambda_1 \cdot \mathrm{id})^n(V).$$

Behauptungen:

(i) $\dim U + \dim W = n = \dim V$.

(ii) $U \cap W = \{0\}$ und $U \oplus W = V$.

(iii) W ist f-invariant, d.h. $f(W) \subset W$.

(iv) $P_{f|_U}(t) = (t - \lambda_1)^{d_1}$ und $P_{f|_W}(t) = \prod_{i=2}^{k} (t - \lambda_i)^{d_i}$.

Anwendung: Die Induktionsannahme läßt sich wegen (iii) und (iv) auf W anwenden. Sie gibt

$$W = \bigoplus_{i=2}^{k} \mathrm{Hau}(f, \lambda_i) \quad \text{und} \quad \dim \mathrm{Hau}(f, \lambda_i) = d_i \quad \text{für } i \geq 2.$$

Daraus und aus $V = U \oplus W$ folgt der Satz.

Beweis der Behauptungen: (i) Satz 5.3 (g) für die lineare Abbildung $(f - \lambda_1 \cdot \mathrm{id})^n$.

(ii) Sei $a \in U \cap W$, $a = (f - \lambda_1 \cdot \mathrm{id})^n(b)$ mit $b \in V$. Wegen $a \in U$ ist

$$0 = (f - \lambda_1 \cdot \mathrm{id})^n(a) = (f - \lambda_1 \cdot \mathrm{id})^{2n}(b),$$

also $b \in \mathrm{Hau}(f, \lambda_1)$. Wegen Lemma 12.11 ist schon die n-te Potenz $(f - \lambda_1 \cdot \mathrm{id})^n(b) = 0$, also $a = 0$. Daher ist $U \cap W = \{0\}$ und $U + W = U \oplus W$. Mit (i) folgt $U \oplus W = V$.

(iii) Klar, denn f und $(f - \lambda_1 \cdot \mathrm{id})^n$ kommutieren.

(iv) Wegen $U \oplus W = V$, und weil U und W f-invariante Untervektorräume sind, ist

$$P_f(t) = P_{f|_U}(t) \cdot P_{f|_W}(t).$$

Das sieht man, indem man Basen von U und W wählt. Zusammen geben sie eine Basis $\mathcal{B}$ von V. Die Matrix $M(\mathcal{B}, f, \mathcal{B})$ ist eine Blockdiagonalmatrix mit 2 Blöcken. Man wendet Beispiel 8.2 (iii) an.
Wäre λ_1 eine Nullstelle von $P_{f|_W}$, so würde W einen Eigenvektor zum Eigenwert λ_1 enthalten, im Widerspruch zu $U \cap W = \{0\}$. Wäre ein λ_i mit $i \geq 2$ eine Nullstelle von $P_{f|_U}$, so würde U einen Eigenvektor u zum Eigenwert λ_i enthalten. Dann wäre

$$0 = (f - \lambda_1 \cdot \mathrm{id})^n(u) = (\lambda_i - \lambda_1)^n \cdot u \neq 0,$$

eine Unmöglichkeit. Daher spaltet sich $P_f(t)$ in $P_{f|_U}(t)$ und $P_{f|_W}(t)$ auf wie in (iv) behauptet. Hier ist die Eindeutigkeit der Zerlegung in Linearfaktoren wichtig, die ein Spezialfall der ZPE-Eigenschaft von $K[t]$ ist. □

Bemerkungen 12.12 (i) Eine *Normalform* eines Endomorphismus $f : V \to V$ eines endlich-dimensionalen Vektorraums V ist eine Matrix $M(\mathcal{B}, f, \mathcal{B})$ für eine geschickt gewählte Basis $\mathcal{B}$ von V, so dass die Matrix besonders hübsch ist.

(ii) In der Situation von Satz 12.10 kann und sollte man auf jeden Fall eine Basis $\mathcal{B}$ wählen, die sich aus Basen der Untervektorräume $\mathrm{Hau}(f, \lambda_i)$ zusammensetzt. Dann ist $M(\mathcal{B}, f, \mathcal{B})$ schon mal in Blockdiagonalgestalt (mit je einem Block zu einem Hauptraum), und man muß sich nur noch Gedanken über Normalformen für die Einschränkungen von f auf diese Haupträume Gedanken machen.

(iii) Auf so einem Hauptraum $\mathrm{Hau}(f, \lambda_i)$ kann man statt f genausogut $f - \lambda_i \cdot \mathrm{id}$ betrachten, denn $M(\mathcal{B}, f, \mathcal{B}) = M(\mathcal{B}, f - \lambda_i \cdot \mathrm{id}, \mathcal{B}) + \lambda_i \cdot E_{d_i}$.

Definition 12.13 (a) Eine Matrix $A \in M(n \times n, K)$ ist in *Jordannormalform*, falls sie die Gestalt hat

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_l \end{pmatrix}.$$

Hier sind

$$A_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix} \in M(r_i \times r_i, K)$$

Jordanblöcke (Beispiel 12.2 (ii)); alle anderen Einträge von A sind Null. Die Zahl r_i ist die *Größe* des Jordanblocks A_i. Natürlich ist $n = r_1 + ... + r_l$. Die λ_i müssen hier nicht alle verschieden sein.

(b) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$. Er *läßt sich in Jordannormalform bringen*, falls eine Basis $\mathcal{B}$ von V existiert, so daß $M(\mathcal{B}, f, \mathcal{B})$ in Jordannormalform ist.
Dann heißt die Matrix $M(\mathcal{B}, f, \mathcal{B})$ eine *Jordannormalform* von f.

(c) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$. Er heißt *nilpotent*, falls $f^n := f \circ ... \circ f = 0$ ist.

Bemerkungen 12.14 (i) In der Situation von Satz 12.10 ist die Einschränkung von $f - \lambda_i \cdot \mathrm{id}$ auf $\mathrm{Hau}(f, \lambda_i)$ wegen Lemma 12.11 nilpotent.

(ii) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$. Er ist genau dann diagonalisierbar, wenn er sich in eine Jordannormalform bringen läßt, bei der alle Jordanblöcke Größe 1 haben.

(iii) Sei $f : V \to V$ ein Endomorphismus und $\mathcal{B} = (b_1, ..., b_n)$ eine Basis von V, so daß $M(\mathcal{B}, f, \mathcal{B})$ in Jordannormalform ist, mit Jordanblöcken $A_i \in M(r_i \times r_i, K)$ $(i = 1, ..., l)$ wie oben.
Um die Eigenschaften leichter zu beschreiben, werden die Basiselemente umbenannt und in Basen von Teilräumen zusammengefaßt:

$$\begin{aligned}
(b_1, .., b_{r_1}) &=: (\beta_1^{(1)}, ..., \beta_{r_1}^{(1)}) =: \mathcal{B}_1, \\
(b_{r_1+1}, .., b_{r_1+r_2}) &=: (\beta_1^{(2)}, ..., \beta_{r_2}^{(2)}) =: \mathcal{B}_2, \\
&\vdots \\
(b_{n-r_l+1}, .., b_n) &=: (\beta_1^{(l)}, ..., \beta_{r_l}^{(l)}) =: \mathcal{B}_l, \\
V_1 := \mathrm{span}_K \mathcal{B}_1, &\quad ... \quad , V_l := \mathrm{span}_K \mathcal{B}_l.
\end{aligned}$$

Dann ist $f(V_i) \subset V_i$, und

$$\begin{aligned}
f(\beta_{r_i}^{(i)}) &= \lambda_i \cdot \beta_{r_i}^{(i)} + \beta_{r_i-1}^{(i)}, \\
f(\beta_{r_i-1}^{(i)}) &= \lambda_i \cdot \beta_{r_i-1}^{(i)} + \beta_{r_i-2}^{(i)}, \\
&\vdots \\
f(\beta_2^{(i)}) &= \lambda_i \cdot \beta_2^{(i)} + \beta_1^{(i)}, \\
f(\beta_1^{(i)}) &= \lambda_i \cdot \beta_1^{(i)},
\end{aligned}$$

denn

$$(f(\beta_1^{(i)}), ..., f(\beta_{r_i}^{(i)})) = (\beta_1^{(i)}, ..., \beta_{r_i}^{(i)}) \cdot \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix}.$$

Also wirkt $f - \lambda_i \cdot \mathrm{id}$ auf V_i so:

$$\beta_{r_i}^{(i)} \mapsto \beta_{r_i-1}^{(i)} \mapsto ... \mapsto \beta_2^{(i)} \mapsto \beta_1^{(i)} \mapsto 0.$$

Man sieht, daß die Potenz $(f - \lambda_i \cdot \mathrm{id})^{r_i}$ ganz V_i auf Null abbildet, d.h.

$$((f - \lambda_i \cdot \mathrm{id})|_{V_i})^{r_i} = 0.$$

Also ist $(f - \lambda_i \cdot \mathrm{id})|_{V_i}$ nilpotent. Man sieht auch

$$\begin{aligned}
\mathrm{Eig}(f,\lambda) &= \bigoplus_{i:\lambda_i=\lambda} K \cdot \beta_1^{(i)}, \\
\dim \mathrm{Eig}(f,\lambda) &= (\text{Anzahl der Jordanblöcke mit Eigenwert } \lambda), \\
\mathrm{Hau}(f,\lambda) &= \bigoplus_{i:\lambda_i=\lambda} V_i, \\
\dim \mathrm{Hau}(f,\lambda) &= \sum_{i:\lambda_i=\lambda} r_i, \\
P_f(t) &= \prod_{\lambda \text{ Eigenwert}} (t-\lambda)^{d(\lambda)} \quad \text{mit} \quad d(\lambda) := \dim \mathrm{Hau}(f,\lambda).
\end{aligned}$$

Nun kommt der Hauptsatz dieses Kapitels 12. Er wird nach dem Korollar 12.16 bewiesen.

Satz 12.15 *(a) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$. Er läßt sich genau dann in Jordannormalform bringen, wenn sein charakteristisches Polynom in Linearfaktoren zerfällt, d.h. wenn es $\lambda_1, ..., \lambda_k \in K$ und $d_1, ..., d_k \in \mathbb{N}$ gibt mit*

$$P_f(t) = \prod_{i=1}^{k} (t - \lambda_i)^{d_i}$$

(also genau in der Situation von Satz 12.10).

(b) Je zwei Jordannormalformen eines Endomorphismus haben die gleichen Jordanblöcke; nur in der Reihenfolge der Blöcke können sie sich unterscheiden.

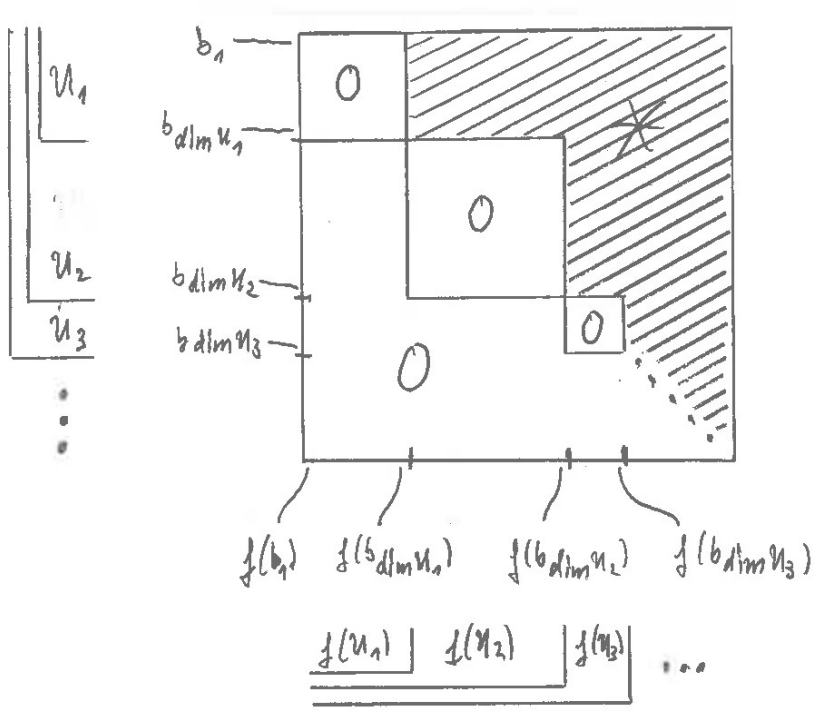
Korollar 12.16 *(a) Im Fall $K = \mathbb{C}$ läßt sich jeder Endomorphismus eines endlichdimensionalen Vektorraums in Jordannormalform bringen (im Fall $K = \mathbb{R}$ gilt das nicht).*

(b) Jeder nilpotente Endomorphismus f eines n-dimensionalen Vektorraums erfüllt $P_f(t) = t^n$ und läßt sich daher in Jordannormalform bringen. Hier sind alle $\lambda_i = 0$.

Beweis: (a) Aufgrund des Fundamentalsatzes der Algebra zerfällt jedes nichtkonstante Polynom in $\mathbb{C}[t]$ in Linearfaktoren. Man kann Satz 12.15 (a) anwenden.

(b) Sei $f : V \to V$ nilpotent und $\dim V = n \in \mathbb{N}$. Sei $U_j := \ker f^j \subset V$. Wegen Lemma 12.11 ist $\{0\} = U_0 \subset U_1 \subset U_2 \subset ... \subset U_n = V$. Eine Familie von Unterräumen, die einander so enthalten, heißt eine *Filtrierung* von V. Wegen des Basisergänzungssatzes Satz 3.19 gibt es eine Basis $\mathcal{B} = (b_1, ..., b_n)$ von V, die die Filtrierung durch die Unterräume U_j *respektiert*, d.h so dass $(b_1, ..., b_{\dim U_j})$ eine Basis

von U_j ist. Wegen $f(U_j) \subset U_{j-1}$ (denn $f^{j-1}(f(U_j)) = f^j(U_j) = \{0\}$) ist die Matrix $M(\mathcal{B}, f, \mathcal{B})$ eine obere Blockdreiecksmatrix, deren Diagonalblöcke nur die Einträge Null haben.



Daher und wegen Bemerkung 8.2 (iii) ist $P_f(t) = t^n$, also zerfällt $P_f(t)$ in die Linearfaktoren $t, ..., t$. Man kann Satz 12.15 (a) anwenden. □

Beweis von Satz 12.15: (a) $\Rightarrow$: Sei $M(\mathcal{B}, f, \mathcal{B})$ in Jordannormalform. Mit Beispiel 12.14 (iii) folgt $P_f(t) = \prod_{\lambda \text{ Eigenwert}} (t-\lambda)^{d(\lambda)}$ mit $d(\lambda) = \dim \mathrm{Hau}(f, \lambda)$.

$\Leftarrow$: Seien $\lambda_1, ..., \lambda_k \in K$ paarweise verschieden, seien $d_i \in \mathbb{N}$, und sei $P_f(t) = \prod_{i=1}^k (t-\lambda_i)^{d_i}$. Das ist die Situation von Satz 12.10. Es reicht, für jeden Hauptraum $\mathrm{Hau}(f, \lambda_i)$ eine Basis $\mathcal{B}_i$ zu finden, so dass die Matrix $M(\mathcal{B}_i, f|_{\mathrm{Hau}(f,\lambda_i)}, \mathcal{B}_i)$ in Jordannormalform ist. Das ist äquivalent dazu, dass die Matrix $M(\mathcal{B}_i, (f-\lambda_i \cdot \mathrm{id})|_{\mathrm{Hau}(f,\lambda_i)}, \mathcal{B}_i)$ in Jordannormalform ist. $(f - \lambda_i \cdot \mathrm{id})|_{\mathrm{Hau}(f,\lambda_i)}$ ist wegen Lemma 12.11 nilpotent.
Daher muß man nun im Prinzip nur noch den 2. Teil von Korollar 12.16 (b) nochmal beweisen, aber nun ohne Benutzung von Satz 12.15 (a). Das wird im folgenden gemacht, es ist ziemlich viel Arbeit.
Ab jetzt wird $k = 1, d_1 = n, \lambda_1 = 0$ angenommen. Also ist nun $f : V \to V$ nilpotent und $P_f(t) = t^n$.
Wie im Beweis von Korollar 12.16 (b) sei $U_j := \ker(f^j) \subset V$. Wieder ist $\{0\} = U_0 \subset U_1 \subset ... \subset U_n = V$. Aus dem Beweis von Lemma 12.11 folgt, dass es ein eindeutiges $m \in \mathbb{N}$ mit $0 < m \leq n$ und

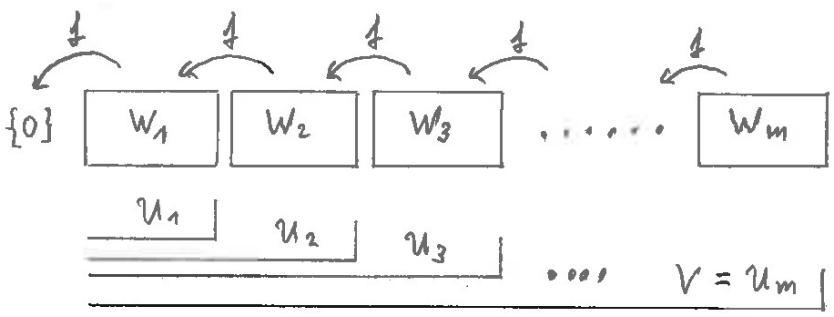
$$\{0\} = U_0 \subsetneqq U_1 \subsetneqq ... \subsetneqq U_m = U_{m+1} = ... = U_n = V$$

gibt.

Behauptung: Es gibt eine Zerlegung von V in eine direkte Summe $V = \bigoplus_{i=1}^{m} W_i$ von Unterräumen W_i mit den Eigenschaften:

$\alpha)$ $U_j = \bigoplus_{i=1}^{j} W_i$ für $j = 1, ..., m$, also $U_1 = W_1$ und für $j \geq 2$ $U_j = U_{j-1} \oplus W_j$.

$\beta)$ $f(W_i) \subset W_{i-1}$ für $i = 2, ..., m$.
$W_1 = U_1$, also $f(W_1) = \{0\}$.



Beweis der Behauptung: Die Räume W_j werden in der Reihenfolge $W_m, W_{m-1}, ..., W_1$ konstruiert.
1. Schritt: Wegen Satz 12.8 (d) kann man $W_m \subset V$ so wählen, daß

$$V = U_{m-1} \oplus W_m.$$

2. Schritt: Dann ist $f(W_m) \cap U_{m-2} = \{0\}$; denn bei $a \in W_m$ mit $f(a) \in U_{m-2}$ ist $0 = f^{m-2}(f(a)) = f^{m-1}(a)$, also $a \in U_{m-1} \cap W_m = \{0\}$. Daher ist

$$U_{m-2} + f(W_m) = U_{m-2} \oplus f(W_m) \subset U_{m-1}.$$

Nun kann man $W_{m-1} \subset U_{m-1}$ so wählen, daß

$$U_{m-1} = U_{m-2} \oplus W_{m-1} \quad \underline{\text{und}} \quad f(W_m) \subset W_{m-1} \quad \text{gilt.}$$

Man wählt nämlich $\widetilde{W}_{m-1} \subset U_m$ so, daß $U_{m-1} = (U_{m-2} \oplus f(W_m)) \oplus \widetilde{W}_{m-1}$ ist und setzt dann $W_{m-1} := f(W_m) \oplus \widetilde{W}_{m-1}$.
Weitere Schritte: Iteration des 2. Schritts liefert die Behauptung. $(\square)$

Beobachtung: Für $i = 2, ..., m$ ist $f : W_i \to W_{i-1}$ injektiv, denn

$$W_i \cap \ker f = W_i \cap U_1 = \{0\},$$

wegen

$$U_1 \subset U_{i-1} \quad \text{und} \quad U_{i-1} + W_i = U_{i-1} \oplus W_i.$$

Nun kann man eine Basis $\mathcal{B}$ von V wählen, so daß $M(\mathcal{B}, f, \mathcal{B})$ in Jordannormalform ist. Das wird mit dem folgenden Diagramm und der Erläuterung danach gemacht.

Basis von	W_m	W_{m-1}	...	W_1
	$\gamma_1^{(m)}$	$f(\gamma_1^{(m)})$	...	$f^{m-1}(\gamma_1^{(m)})$
	$\vdots$	$\vdots$		$\vdots$
	$\gamma_{s_m}^{(m)}$	$f(\gamma_{s_m}^{(m)})$	...	$f^{m-1}(\gamma_{s_m}^{(m)})$
		$\gamma_1^{(m-1)}$	...	$f^{m-2}(\gamma_1^{(m-1)})$
		$\vdots$	$\vdots$	
		$\gamma_{s_{m-1}}^{(m-1)}$	...	$f^{m-2}(\gamma_{s_{m-1}}^{(m-1)})$
			$\ddots$	$\vdots$
				$\gamma_1^{(1)}$
				$\vdots$
				$\gamma_{s_1}^{(1)}$

Zuerst wählt man eine Basis $\gamma_1^{(m)}, ..., \gamma_{s_m}^{(m)}$ von W_m. Dann ergänzt man $f(\gamma_1^{(m)}), ..., f(\gamma_{s_m}^{(m)})$ mit $\gamma_1^{(m-1)}, ..., \gamma_{s_{m-1}}^{(m-1)}$ zu einer Basis von W_{m-1}. Dann ergänzt man die Bilder unter f dieser Basis zu einer Basis von W_{m-2}. Und so weiter.
Im Diagramm geben die Zeilen von rechts nach links gelesen Basen von Jordanblöcken einer Jordannormalform.

(b) Auch hier reicht es, den Fall eines nilpotenten Endomorphismus zu betrachten. Sei $s_l \geq 0$ die Anzahl der Jordanblöcke der Größe l $(1 \leq l \leq n)$. Es ist natürlich $\sum_{l=1}^{n} l \cdot s_l = n$. Es ist

$$\begin{aligned} s_l &= \dim W_l - \dim f(W_{l+1}) = \dim W_l - \dim W_{l+1} \\ &= (\dim U_l - \dim U_{l-1}) - (\dim U_{l+1} - \dim U_l) \\ &= 2\dim U_l - \dim U_{l+1} - \dim U_{l-1} \\ &= 2\dim \ker f^l - \dim \ker f^{l+1} - \dim \ker f^{l-1}. \end{aligned}$$

Also ist s_l unabhängig von allen Wahlen. Daher haben je zwei Jordannormalformen bis auf die Reihenfolge die gleichen Jordanblöcke. □

Bemerkungen 12.17 (i) Wenn das charakteristische Polynom eines Endomorphismus $f : V \to V$ in Linearfaktoren zerfällt, $P_f(t) = \prod_{\lambda \text{ Eigenwert}} (t - \lambda)^{d(\lambda)}$, weiß man nach Satz 12.15, dass f eine Jordannormalform besitzt.
Aber allein aus dem charakteristischen Polynom kann man selten die Anzahl und Größe der Jordanblöcke zu einem Eigenwert λ bestimmen (natürlich hat man im Fall $d(\lambda) = 1$ nur einen Jordanblock zum Eigenwert λ, und der hat Größe 1). Die Anzahl ergibt sich aus

$$\dim \mathrm{Eig}(f, \lambda) = (\text{Anzahl der Jordanblöcke mit Eigenwert } \lambda)$$

(Bemerkung 12.14 (iii)).
Aber oft reichen $P_f(t)$ und $\dim \mathrm{Eig}(f,\lambda)$ auch zusammen nicht, um die Größen der Jordanblöcke zu bestimmen. Glück hat man in den beiden Extremfällen $\dim \mathrm{Eig}(f,\lambda) = d(\lambda)$ und $\dim \mathrm{Eig}(f,\lambda) = 1$.
Im 1. Fall hat man $d(\lambda)$ viele Jordanblöcke der Größe 1 zum Eigenwert λ.
Im 2. Fall hat man nur einen Jordanblock der Größe $d(\lambda)$ zum Eigenwert λ. Aber in anderen Fällen ist die Bestimmung der Größen der Jordanblöcke mühsam. Das unten definierte *Minimalpolynom* $M_f(t)$ hilft ein wenig: Es gibt für jeden Eigenwert die Größe des größen Jordanblocks (Definition/Lemma 12.21 (c)).

(ii) Der 2. Fall tritt zum Beispiel bei einer *Begleitmatrix* auf: Sei $f(t) = t^n + a_{n-1}t^{n-1} + \ldots + a_1 t + a_0 \in K[t]$ ein unitäres Polynom mit $\deg f(t) = n \geq 1$. Die Matrix

$$A^{(f)} := \begin{pmatrix} & & & -a_0 \\ 1 & & & -a_1 \\ & \ddots & & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix} \in M(n \times n, K)$$

heißt *Begleitmatrix zum Polynom* $f(t)$. Ihr charakteristisches Polynom ist $P_A(t) \stackrel{!}{=} f(t)$ (Aufgabe 3 von Blatt 10 im HWS 2019). Wenn $f(t)$ in $K[t]$ in Linearfaktoren zerfällt, sind alle Eigenräume von $A^{(f)}$ eindimensional, denn $\mathrm{rang}(A - \lambda \cdot E_n) = n-1$ oder n, wegen der Einsen in der Nebendiagonalen. Daher ist A zu einer Matrix in Jordannormalform konjugiert, die zu jedem Eigenwert nur einen Jordanblock hat.

(iii) (Beweis in einer Algebra-Vorlesung) Zu jedem Körper K gibt es einen (bis auf Isomorphie eindeutigen) kleinsten Körper $\overline{K} \supset K$, in dem jedes Polynom in $K[t]$ in Linearfaktoren zerfällt. Er heißt *algebraischer Abschluß* von K. Es ist $\overline{\mathbb{R}} = \mathbb{C}$. Aber es ist $\overline{\mathbb{Q}} \subsetneq \mathbb{C}$.

(iv) Eine Matrix $A \in M(n \times n, K)$ liegt auch in $M(n \times n, \overline{K})$. Daher gibt es nach Satz 12.15 (a) eine invertierbare Matrix $B \in M(n \times n, \overline{K})$, so daß $B^{-1} \cdot A \cdot B$ in Jordannormalform ist. Das ist auch für das Verständnis der Linksmultiplikation l_A mit A auf $M(n \times 1, K)$ sehr nützlich.

(v) Andererseits kann man den Begriff der Jordannormalform auch in einer geeigneten Weise erweitern und dann innerhalb von $M(n \times n, K)$ bleiben. Das wird am Ende des Kapitels in Satz 12.23 diskutiert. Vorher kommen der Satz von Cayley-Hamilton und das Minimalpolynom. Beide erfordern nicht, dass das charakteristische Polynom in Linearfaktoren zerfällt.

Bemerkung 12.18 Ein Endomorphismus $f : V \to V$ eines K-Vektorraums V induziert eine Abbildung

$$\begin{array}{rcl} \Phi_f : K[t] & \to & \mathrm{End}(V), \\ \sum_{i=0}^{m} a_i t^i & \mapsto & \sum_{i=0}^{m} a_i f^i. \end{array}$$

Hier ist $f^0 := \mathrm{id}_V$. Die Abbildung Φ_f ist ein Ringhomomorphismus und ein Vektorraumhomomorphismus. Das Bild $\Phi_f(K[t])$ ist ein *kommutativer* Unterring mit Eins von $\mathrm{End}(V)$.
Sei $\dim V = n \in \mathbb{N}$. Dann ist $\dim \mathrm{End}(V) = n^2$. Es ist $\dim K[t]_{\leq n^2} = n^2 + 1$. Daher ist schon die Einschränkung von Φ_f auf $K[t]_{\leq n^2}$ nicht injektiv. Satz 12.19 sagt, daß $P_f(t)$ im Kern liegt.

Satz 12.19 *(Satz von Cayley und Hamilton) Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$. Es ist*

$$P_f(f) = 0 \in \mathrm{End}(V).$$

Beweis: Sei $\mathcal{B} = (b_1, ..., b_n)$ irgendeine Basis von V und $M(\mathcal{B}, f, \mathcal{B}) =: A = (a_{ij})$, also

$$f(b_j) = \sum_{i=1}^{n} a_{ij} b_i.$$

Wir wenden auf jeden Eintrag der Matrix

$$C(t) := (A - t \cdot E_n)^{tr} \in M(n \times n, K[t])$$

die Abbildung Φ_f an und erhalten die Matrix

$$C(f) = \begin{pmatrix} a_{11} \cdot \mathrm{id}_V - f & a_{21} \cdot \mathrm{id}_V & \cdots & a_{n1} \cdot \mathrm{id}_V \\ a_{12} \cdot \mathrm{id}_V & a_{22} \cdot \mathrm{id}_V - f & & a_{n2} \cdot \mathrm{id}_V \\ \vdots & & \ddots & \\ a_{1n} \cdot \mathrm{id}_V & & & a_{nn} \cdot \mathrm{id}_V - f \end{pmatrix} \in M(n \times n, \Phi_f(K[t])).$$

Es ist $\det C(f) = (-1)^n P_f(f)$. Multipliziert man von rechts den Spaltenvektor $(b_1, ..., b_n)^{tr} \in M(n \times 1, V)$ dran, so erhält man

$$C(f) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} -f(b_1) + \sum_{i=1}^{n} a_{i1} b_i \\ \vdots \\ -f(b_n) + \sum_{i=1}^{n} a_{in} b_i \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

$C(f)^\sharp$ ist die Komplementärmatrix zu $C(f)$ von Definition 7.13. Mit Satz 7.15 (c) erhält man

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = C(f)^\sharp \cdot C(f) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} \det C(f) & & 0 \\ & \ddots & \\ 0 & & \det C(f) \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Also ist $P_f(f)(b_i) = 0$ für alle $i = 1, ..., n$; also $P_f(f)(v) = 0$ für alle $v \in V$. Daher ist $P_f(f) = 0 \in \mathrm{End}(V)$. □

Bemerkung 12.20 Bei einem Endomorphismus $f : V \to V$ mit einer Jordannormalform $M(\mathcal{B}, f, \mathcal{B})$ sieht man Satz 12.19 einfacher. In dem Fall ist (mit den Notationen von 12.14 (iii))

$$V = \bigoplus_{\lambda} \mathrm{Hau}(f, \lambda) = \bigoplus_{\lambda} \left(\bigoplus_{i:\lambda_i=\lambda} V_i \right),$$

und es ist

$$(f - \lambda_i \cdot \mathrm{id})^{r_i}(V_i) = \{0\},$$

also auch

$$P_f(f)(V_i) = \left(\prod_{j \neq i}^{k} (f - \lambda_j \cdot \mathrm{id})^{r_j} \right) \circ (f - \lambda_i \cdot \mathrm{id})^{r_i}(V_i) = \{0\}.$$

Definition/Lemma 12.21 *Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit $\dim_K V = n \in \mathbb{N}$.*

(a) (Definition) Der Kern der Abbildung Φ_f von Bemerkung 12.18 enthält ein unitäres Polynom $M_f(t)$ mit kleinstem positiven Grad. Es heißt Minimalpolynom *von f.*

(b) (Lemma) $M_f(t)$ teilt jedes andere Element von $\ker \Phi_f$, insbesondere teilt es das charakteristische Polynom $P_f(t)$.

(c) (Lemma) Läßt sich f in Jordannormalform bringen, so ist (mit den Notationen von 12.14 (iii))

$$M_f(t) = \prod_{\lambda \text{ Eigenwert}} (t - \lambda)^{r(\lambda)}$$
$$\text{mit} \qquad r(\lambda) = \max(r_i \mid \lambda_i = \lambda).$$

Beweis: (a) Definition.
(b) Würde $M_f(t)$ ein Element $g(t)$ von $\ker \Phi_f$ nicht teilen, so würde man mit Polynomdivision $g(t) = q(t)M_f(t) + r(t)$ ein Polynom $r(t)$ von kleinerem Grad erhalten, das auch in $\ker \Phi_f$ liegen würde. Widerspruch.
(c) Wegen Bemerkung 12.14 (iii) ist $(f - \lambda_i \cdot \mathrm{id})^{r_i}$ die kleinste Potenz von $f - \lambda_i \cdot \mathrm{id}$, die ganz V_i auf Null abbildet. Daher ist $M_f(t)$ das kleinste unitäre Polynom, das für alle i durch $(t - \lambda_i)^{r_i}$ teilbar ist. Das ist gerade

$$M_f(t) = \prod_{\lambda \text{ Eigenwert}} (t - \lambda)^{r(\lambda)}.$$

□

Beispiel 12.22 Läßt sich $f : V \to V$ in Jordannormalform bringen mit $\dim V = 15$ und λ_i und r_i wie in der Tabelle (mit α, β und $\gamma \in K$ paarweise verschieden)

λ_i	α	α	β	β	β	γ
r_i	1	5	2	3	3	1

so ist

$$\begin{aligned} P_f(t) &= (t-\alpha)^6(t-\beta)^8(t-\gamma), \\ M_f(t) &= (t-\alpha)^5(t-\beta)^3(t-\gamma). \end{aligned}$$

Satz 12.23 *(Ausblick, Verallgemeinerung der Jordannormalform, ohne Beweis) Sei $f : V \to V$ ein Endomorphismus eines n-dimensionalen K-Vektorraums V, und sei*

$$P_f(t) = \prod_{i=1}^{k} p_i(t)^{d_i} \qquad \text{mit } d_i \in \mathbb{N} \quad \text{und}$$

$$p_i(t) \in K[t] \quad \text{irreduzibel und unitär und paarweise verschieden}$$

die eindeutige Zerlegung des charakteristischen Polynoms von f in ein Produkt irreduzibler und unitärer Polynome ($K[t]$ ist ein ZPE-Ring). Sei $n_i := \deg p_i \geq 1$.

(a) Das Minimalpolynom $M_f(t)$ hat die Gestalt

$$M_f(t) = \prod_{i=1}^{k} p_i(t)^{e_i} \quad \text{mit } 1 \leq e_i \leq d_i.$$

(b) ($\supset$ Satz 12.10) Sei

$$\mathrm{Hau}(f, p_i) := \ker(p_i(f)^{e_i}) \subset V.$$

Dann ist

$$\begin{aligned} \mathrm{Hau}(f, p_i) &= \ker(p_i(f)^m) \quad \text{für alle } m \geq e_i, \\ \dim \mathrm{Hau}(f, p_i) &= d_i \cdot n_i, \\ V &= \bigoplus_{i=1}^{k} \mathrm{Hau}(f, p_i). \end{aligned}$$

(c) ($\supset$ Satz 12.15) Es gibt eindeutige Zahlen $s_i, r_{i1}, ..., r_{is_i} \in \mathbb{N}$ mit $\sum_{j=1}^{s_i} r_{ij} = d_i$ und eine Basis $\mathcal{B}_i$ von $\mathrm{Hau}(f, p_i)$, so dass gilt:

$$M(\mathcal{B}_i, f|_{\mathrm{Hau}(f,p_i)}, \mathcal{B}_i) = \begin{pmatrix} A^{(p_i)}_{(r_{i1})} & & \\ & \ddots & \\ & & A^{(p_i)}_{(r_{is_i})} \end{pmatrix} \in M(d_i n_i \times d_i n_i, K)$$

mit

$$A^{(p_i)}_{(r_{ij})} = \begin{pmatrix} A^{(p_i)} & E_{n_i} & & \\ & \ddots & \ddots & \\ & & \ddots & E_{n_i} \\ & & & A^{(p_i)} \end{pmatrix} \in M(r_{ij}n_i \times r_{ij}n_i, K),$$

$$A^{(p_i)} = (\textit{Begleitmatrix zu } p_i) \in M(n_i \times n_i, K).$$

Die Matrix $M(\mathcal{B}_i, f|_{\mathrm{Hau}(f,p_i)}, \mathcal{B}_i)$ verallgemeinert eine Jordannormalform, die Matrix $A^{(p_i)}_{(r_{ij})}$ verallgemeinert einen Jordanblock (Bemerkung 12.24 (i)).

(d) Sei $\overline{K} \supset K$ der algebraische Abschluß von K (Bemerkung 12.17 (iii)), und sei $p_i(t) = \prod_{l=1}^{n_i}(t - \lambda_{il})$ mit $\lambda_{i1}, ..., \lambda_{in_i} \in \overline{K}$. Dies sind die Eigenwerte von $A^{(p_i)}_{(r_{ij})}$ als Matrix in $M(r_{ij}n_i \times r_{ij}n_i, \overline{K})$, sie sind alle verschieden. Es gibt eine Matrix $B_{i,j} \in GL(r_{ij}n_i, \overline{K})$, so dass $B_{i,j}A^{(p_i)}_{(r_{ij})}B^{-1}_{i,j}$ in Jordannormalform ist und zu jedem λ_{il} genau einen Jordanblock der Größe r_{ij} hat.

Bemerkungen 12.24 (i) In Prosa: Wenn man mit K arbeitet, sind die Nullstellen von $p_i(t)$ und die Jordanblöcke nicht sichtbar. Aber eine Blockdiagonalmatrix über $\overline{K}$ aus gleich großen Jordanblöcken zu allen diesen Nullstellen ist konjugiert zu einer gemeinsamen Matrix $A^{(p_i)}_{(r_{ij})}$, die über K definiert ist.
Diese Matrix verallgemeinert einen Jordanblock: Die Eigenwerte λ in der Diagonalen werden durch die Begleitmatrizen $A^{(p_i)} \in M(n_i \times n_i, K)$ ersetzt, die 1'en in der Nebendiagonalen werden durch die Matrizen $E_{n_i} \in M(n_i \times n_i, K)$ ersetzt.

(ii) Es gibt auch eine andere Normalform in $M(n \times n, K)$ für $f : V \to V$, die *rationale Normalform*. Bei ihr faßt man möglichst viele Jordanblöcke zusammen. Die Matrix besteht dann aus möglichst wenigen und möglichst großen Begleitmatrizen von gewissen Polynomen $f_1, ..., f_s \in K[t]$ mit $f_1|f_2|...|f_s$ und $f_1 \cdot ... \cdot f_s = p_f(t)$. Diese Polynome sind dann eindeutig und heißen *Elementarteiler*.

(iii) Der beste Beweis von Satz 12.23 benutzt einen Satz über die Struktur von *Moduln über Hauptidealringen*. Hier ist V ein $K[t]$-Modul: $g(t) \in K[t]$ gibt den Endomorphismus $g(f) \in \mathrm{End}(V)$ (Bemerkung 12.18).

13 Bilinearformen und Sesquilinearformen

Am Ende von Kapitel 9 waren drei Sätze zitiert worden. Der erste wird nun bewiesen. Auf die anderen beiden wird hier nicht mehr eingegangen.

Satz 13.1 *(∼ Satz 9.24) (Spektralsatz für reelle symmetrische Matrizen)*
Sei $A \in M(n \times n, \mathbb{R})$ symmetrisch. Es gilt:

(i) $P_A(t) = \prod_{i=1}^{n}(t - \lambda_i)$ mit $\lambda_i \in \mathbb{R}$. Also sind alle Eigenwerte von A (als komplexer Matrix) reell.

(ii) Die Matrix A ist diagonalisierbar. Mit Satz 12.10 folgt: $M(n \times 1, \mathbb{R})$ ist direkte Summe aller Eigenräume.

(iii) Die Eigenräume zu verschiedenen Eigenwerten sind orthogonal bezüglich des Standardskalarproduktes auf $M(n \times 1, \mathbb{R})$.

(iv) Daher gibt es eine ON-Basis von $M(n \times 1, \mathbb{R})$ (bezüglich des Standardskalarproduktes) aus Eigenvektoren von A. Ist T eine Matrix, deren Spalten eine solche ON-Basis bilden, so ist $T^{-1} = T^{tr}$ (dann heißt T orthogonal), und es ist

$$T^{tr} \cdot A \cdot T = T^{-1} \cdot A \cdot T = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Beweis: (i) Als Polynom in $\mathbb{C}[t]$ zerfällt das charakteristische Polynom $P_A(t)$ in Linearfaktoren, $P_A(t) = \prod_{i=1}^{n}(t - \lambda_i)$. Sei $\lambda = \lambda_i$ für ein i, und sei $v = (v_1 ... v_n)^{tr} \in M(n \times 1, \mathbb{C}) - \{0\}$ ein Eigenvektor von A zum Eigenwert λ, $A \cdot v = \lambda \cdot v$. Eine Notation: Bei Matrizen $B = (b_{ij}) \in M(n_1 \times n_2, \mathbb{C})$ bezeichnet $\overline{B}$ die Matrix $\overline{B} = (\overline{b}_{ij})$. Es ist

$$0 < \sum_{j=1}^{n} |v_j|^2 = \sum_{j=1}^{n} v_j \overline{v}_j = v^{tr} \cdot \overline{v} \qquad \text{wegen } v \neq 0.$$

Und es ist wegen $A = \overline{A}$ (A reell) und $A = A^{tr}$ (A symmetrisch)

$$\begin{aligned} \lambda \cdot (v^{tr} \cdot \overline{v}) &= (\lambda \cdot v)^{tr} \cdot \overline{v} = (A \cdot v)^{tr} \cdot \overline{v} \\ &= v^{tr} \cdot A^{tr} \cdot \overline{v} = v^{tr} \cdot \overline{A} \cdot \overline{v} = v^{tr} \cdot \overline{A \cdot v} \\ &= v^{tr} \cdot \overline{\lambda \cdot v} = \overline{\lambda} \cdot (v^{tr} \cdot \overline{v}). \end{aligned}$$

Daher ist $\lambda = \overline{\lambda}$, also $\lambda \in \mathbb{R}$.

(ii) Wegen (i) und Satz 12.15 (a) gibt es eine Basis $b_1, , , .b_n$ von $M(n \times 1, \mathbb{R})$, so dass mit $B := (b_1 ... b_n) \in GL(n, \mathbb{R})$ die Matrix $B^{-1}AB$ in Jordannormalform ist.

Nun Indirekter Beweis. **Annahme:** *A ist nicht diagonalisierbar.*

Dann gibt es wegen Bemerkung 12.14 (ii) einen Eigenwert λ, so dass die Matrix $B^{-1}AB$ einen Jordanblock der Größe ≥ 2 zum Eigenwert λ hat. Also gibt es Vektoren b_j und b_{j+1} in der Basis oben mit

$$A \cdot b_j = \lambda \cdot b_j \quad \text{und} \quad A \cdot b_{j+1} = \lambda \cdot b_{j+1} + b_j.$$

Aber dann ist (wieder wegen $A = A^{tr}$ symmetrisch)

$$\begin{aligned} \lambda \cdot b_j^{tr} \cdot b_{j+1} + b_j^{tr} \cdot b_j &= b_j^{tr} \cdot (\lambda \cdot b_{j+1} + b_j) = b_j^{tr} \cdot (A \cdot b_{j+1}) \\ &= (b_j^{tr} \cdot A^{tr}) \cdot b_{j+1} = (A \cdot b_j)^{tr} \cdot b_{j+1} = \lambda \cdot b_j^{tr} \cdot b_{j+1} \\ \text{also} \qquad b_j^{tr} \cdot b_j &= 0. \end{aligned}$$

Aber das ist wegen $b_j \neq 0$ unmöglich. Also war die Annahme oben falsch.

(iii) Seien v_1 und $v_2 \in M(n \times 1, \mathbb{R}) - \{0\}$ Eigenvektoren zu verschiedenen Eigenwerten λ_1 und λ_2. Dann ist (wieder wegen $A = A^{tr}$ symmetrisch)

$$\begin{aligned} \lambda_1 \cdot v_1^{tr} \cdot v_2 &= (A \cdot v_1)^{tr} \cdot v_2 = v_1^{tr} \cdot A^{tr} \cdot v_2 \\ &= v_1^{tr} \cdot (A \cdot v_2) = v_1^{tr} \cdot (\lambda_2 \cdot v_2) = \lambda_2 \cdot v_1^{tr} \cdot v_2. \end{aligned}$$

Wegen $\lambda_1 \neq \lambda_2$ ist $v_1^{tr} \cdot v_2 = 0$, also $v_1 \perp v_2$.

(iv) Man muß nur auf jedem Eigenraum $\mathrm{Eig}(A, \lambda_i)$ eine ON-Basis wählen. Die ON-Basen für alle Eigenräume setzen sich zu einer ON-Basis von $M(n \times 1, \mathbb{R})$ zusammen, die aus Eigenvektoren besteht. Sei T eine Matrix, deren Spalten so eine ON-Basis aus Eigenvektoren sind. Weil die Spalten eine ON-Basis sind, ist $T^{tr} \cdot T = E_n$, also $T^{-1} = T^{tr}$. Weil die Spalten Eigenvektoren von A sind, ist

$$A \cdot T = T \cdot \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}. \qquad \square$$

Im folgenden werden einige allgemeine Punkte zu Bilinearformen über beliebigen Körpern zusammengestellt. Zuerst wird die Beziehung zwischen Matrizen und Bilinearformen studiert. Das ist analog zu Sätzen in Kapitel 5 (zur Beziehung zwischen Matrizen und linearen Abbildungen). Danach kommen wichtige Begriffe, die im Fall eines Euklidischen Vektorraums trivial werden, aber in anderen Fällen nicht (*Radikal, isotrope Vektoren*).
Schließlich gibt es einen Ausblick auf *hermitesche Sesquilinearformen*. Das ist ein Analogon über $\mathbb{C}$ zu symmetrischen Bilinearformen über $\mathbb{R}$. Da gibt es auch ein Analogon zu Euklidischen Vektorräumen, die *unitären Vektorräume*.

Beispiel 13.2 ($\sim$ Beispiel 9.2 (iv)) Die folgende Definition von Bil_A und das Lemma 13.3 geben eine erste Verbindung zwischen Bilinearformen und Matrizen, für den Fall von Spaltenvektorräumen.

Eine $(m \times n)$-Matrix $A = (a_{ij}) \in M(m \times n, K)$ definiert (offenbar) eine Bilinearform auf $M(m \times 1, K) \times M(n \times 1, K)$ durch

$$\begin{aligned} \mathrm{Bil}_A : M(m \times 1, K) \times M(n \times 1, K) &\to K \\ (x, y) &\mapsto x^{tr} \cdot A \cdot y = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i \cdot a_{ij} \cdot y_j. \end{aligned}$$

Lemma 13.3 *($\sim$ Lemma 9.3) (Matrizen und Bilinearformen, 1.Teil)*
(a) Zu jeder Bilinearform ϕ auf $M(m \times 1, K) \times M(n \times 1, K)$ gibt es eine eindeutige Matrix $A \in M(m \times n, K)$ mit $\phi = \mathrm{Bil}_A$.
(b) Die Menge der Bilinearformen auf $M(m \times 1, K) \times M(n \times 1, K)$ ist ein K-Vektorraum, und die Abbildung $A \mapsto \mathrm{Bil}_A$ ist ein Vektorraumisomorphismus von $M(m \times n, K)$ in diese Menge.
(c) Im Fall von $m = n$ ist Bil_A genau dann symmetrisch [bzw. schiefsymmetrisch], wenn A symmetrisch [bzw. schiefsymmetrisch] ist.

Beweis: Übung. □

Notationen 13.4 (a) (Eine Variante der Verallgemeinerung der Matrizenmultiplikation in Notation 5.7)
Sei V ein K-Vektorraum und $m, n \in \mathbb{N}$. Die Abbildung

$$\begin{aligned} M(n \times m, K) \times M(m \times 1, V) &\to M(n \times 1, V) \\ ((a_{ij}), \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}) &\mapsto \begin{pmatrix} \sum_{j=1}^{m} a_{1j} b_j \\ \vdots \\ \sum_{j=1}^{m} a_{nj} b_j \end{pmatrix} =: (a_{ij}) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \end{aligned}$$

wird als eine Verallgemeinerung der Matrizenmultiplikation aufgefaßt. Hier wird die Körpermultiplikation durch die skalare Multiplikation ersetzt. Offenbar ist

$$\left((a_{ij}) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \right)^{tr} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}^{tr} \cdot (a_{ij})^{tr} = (b_1 \cdots b_m) \cdot (a_{ij})^{tr}.$$

(b) (Matrizen und Bilinearformen, 2. Teil)
Seien V und W K-Vektorräume, sei $\phi : V \times W \to K$ eine Bilinearform, sei $\mathcal{A} = (a_1, ..., a_m) \in M(1 \times m, V)$, $\mathcal{B} = (b_1, ..., b_n) \in M(1 \times n, W)$. Dann bezeichnet $\phi(\mathcal{A}^{tr}, \mathcal{B})$ die Matrix

$$\phi(\mathcal{A}^{tr}, \mathcal{B}) := (\phi(a_i, b_j)) \in M(m \times n, K).$$

Das kann man als Notationsmißbrauch ansehen, oder man kann es als Definition einer Erweiterung von ϕ zu einer Abbildung $\phi : M(m \times 1, V) \times M(1 \times n, W) \to M(m \times n, K)$ auffassen.

Lemma 13.5 *(Matrizen und Bilinearformen, 3. Teil)*
(a) In der Situation von Notation 13.4 (b) sei $C \in M(k\times m, K)$ und $D \in M(n\times l, K)$. Dann ist

$$\phi(C \cdot \mathcal{A}^{tr}, \mathcal{B} \cdot D) = C \cdot \phi(\mathcal{A}^{tr}, \mathcal{B}) \cdot D.$$

(Das kann man als Verallgemeinerung der Bilinearität von ϕ im Fall von $\phi : M(m \times 1, V) \times M(1 \times n, W) \to M(m \times n, K)$ deuten.)

(b) (Transformation von Matrizen zu Bilinearformen bezüglich Basiswechsel)
In der Situation von Notation 13.4 (b) seien $\mathcal{A}$ und $\mathcal{A}' = (a'_1, ..., a'_m)$ Basen von V und $\mathcal{B}$ und $\mathcal{B}' = (b'_1, ..., b'_n)$ Basen von W. Dann gilt

$$\phi((\mathcal{A}')^{tr}, \mathcal{B}') = M(\mathcal{A}, \mathcal{A}')^{tr} \cdot \phi(\mathcal{A}^{tr}, \mathcal{B}) \cdot M(\mathcal{B}, \mathcal{B}').$$

Beweis: (a) Sei $C = (c_{ij})$ und $D = (d_{ij})$. Die beiden Seiten in der Matrizengleichung werden koeffizientenweise verglichen. Die dritte Gleichung benutzt die Bilinearität von $\phi : V \times W \to K$.

$$\begin{aligned}
\big(\phi(C \cdot \mathcal{A}^{tr}, \mathcal{B} \cdot D)\big)_{pq} &= \phi(\big(C \cdot \mathcal{A}^{tr}\big)_p, (\mathcal{B} \cdot D)_q) \\
&= \phi(\sum_{i=1}^{m} c_{pi} a_i, \sum_{j=1}^{n} d_{jq} b_j) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} c_{pi}\phi(a_i, b_j) d_{jq} \\
&= \big(C \cdot \phi(\mathcal{A}^{tr}, \mathcal{B}) \cdot D\big)_{pq}.
\end{aligned}$$

(b) Das folgt unmittelbar aus (a) und $\mathcal{A}' = \mathcal{A} \cdot M(\mathcal{A}, \mathcal{A}')$ und $\mathcal{B}' = \mathcal{B} \cdot M(\mathcal{B}, \mathcal{B}')$. □

Definition/Satz 13.6 *(Matrizen und Bilinearformen, 4. Teil)*
Seien V und W endlich-dimensionale K-Vektorräume. Sei $\mathcal{A} = (a_1, ..., a_m)$ eine Basis von V und $\mathcal{B} = (b_1, ..., b_n)$ eine Basis von W.

(a) (Definition) Eine $(m \times n)$-Matrix $C = (c_{ij}) \in M(m \times n, K)$ definiert offenbar eine Bilinearform auf $V \times W$ durch

$$\begin{aligned}
\mathrm{Bil}_{C,\mathcal{A},\mathcal{B}} : V \times W &\to K \\
(\sum_{i=1}^{m} x_i a_i, \sum_{j=1}^{n} y_j b_j) &\mapsto \big(x_1 \cdots x_m\big) \cdot C \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{m}\sum_{j=1}^{n} x_i \cdot c_{ij} \cdot y_j.
\end{aligned}$$

(b) (Satz) Die Menge

$$\mathrm{Bil}(V, W) := \{\phi : V \times W \to K \text{ Bilinearform}\}$$

ist ein K-Vektorraum, und die Abbildung

$$M(m \times n, K) \to \mathrm{Bil}(V, W), \quad C \mapsto \mathrm{Bil}_{C,\mathcal{A},\mathcal{B}},$$

ist ein Vektorraumisomorphismus. Der inverse Vektorraumisomorphismus ist

$$\mathrm{Bil}(V,W) \to M(m \times n, K), \quad \phi \mapsto \phi(\mathcal{A}^{tr}, \mathcal{B}).$$

Beweis: (a) Definition. $\mathrm{Bil}_{C,\mathcal{A},\mathcal{B}}$ ist bilinear, weil in jedem Summanden der Summe $\sum_{i=1}^{m}\sum_{j=1}^{n} x_i \cdot c_{ij} \cdot y_j$ die x_k linear auftreten und die y_l linear auftreten.
(b) Die Aussagen sind zu denen von Satz 5.5 (b) und Satz 5.11 (b) verwandt.
Man muß zuerst zeigen, dass $\mathrm{Bil}(V,W)$ ein K-Vektorraum ist. Tatsächlich ist es ein Untervektorraum von $\mathrm{Abb}(V \times W, K)$, der ein Vektorraum mit der punktweisen Addition und skalaren Multiplikation ist (Beispiel 3.3 (c)).
Danach muß man zeigen, daß die beiden Abbildungen $\phi \mapsto \phi(\mathcal{A}^{tr}, \mathcal{B})$ und $C \mapsto \mathrm{Bil}_{C,\mathcal{A},\mathcal{B}}$ invers zueinander sind und daß eine der beiden linear ist.
Die Details: Übung. □

Beispiele 13.7 (i) Polynome geben auf $[0,1]$ stetige Abbildungen. Daher ist $\mathbb{R}[t] \subset \mathcal{C}^0([0,1],\mathbb{R})$. Der Untervektorraum $\mathbb{R}[t]_{\leq 3}$ hat die Basis $\mathcal{B} = (1, t, t^2, t^3)$. In Beispiel 9.2 (ii) war die Bilinearform ϕ_{int} mit

$$\phi_{int}(f,g) := \int_{[0,1]} f(x)g(x)dx$$

auf $\mathcal{C}^0([0,1],\mathbb{R})$ betrachtet worden. In Beispiel 9.4 (ii) war festgestellt worden, dass sie positiv definit ist, so daß $\mathcal{C}^0([0,1],\mathbb{R})$ ein Euklidischer Vektorraum mit ϕ_{int} als Skalarprodukt ist. Wegen $\int_0^1 t^k \mathrm{d}t = \frac{1}{k+1}$ hat ϕ_{int} auf $\mathbb{R}[t]_{\leq 3}$ bezüglich der Basis $\mathcal{B}$ die Matrix

$$\phi_{int}(\mathcal{B}^{tr}, \mathcal{B}) = (\int_0^1 t^{i+j-2}\mathrm{d}t)_{i,j=1,\ldots,4} = (\frac{1}{i+j-1})_{i,j=1,\ldots,4} = \begin{pmatrix} \frac{1}{1} & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}.$$

(ii) Auf $M(2 \times 1, K)$ hat man die Standardbasis $\mathcal{B} = (e_1, e_2) = (\binom{1}{0}, \binom{0}{1})$ und die Basis $\mathcal{B}' = (\binom{1}{2}, \binom{1}{-1})$. Die Basiswechselmatrix ist natürlich $T := M(\mathcal{B}, \mathcal{B}') = \left(\begin{smallmatrix} 1 & 1 \\ 2 & -1 \end{smallmatrix}\right)$. Die symmetrische Bilinearform ϕ_2 mit

$$\phi_2(\mathcal{B}^{tr}, \mathcal{B}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

erfüllt

$$\begin{aligned} \phi_2((\mathcal{B}')^{tr}, \mathcal{B}') &= T^{tr} \cdot \phi_2(\mathcal{B}^{tr}, \mathcal{B}) \cdot T = T^{tr} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot T \\ &= \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -3 & 3 \\ 3 & 0 \end{pmatrix}. \end{aligned}$$

(iii) Die Bilinearform ϕ_3 auf $M(2 \times 1, K)$ mit

$$\phi_3(\mathcal{B}^{tr}, \mathcal{B}) = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$$

ist nicht symmetrisch, wegen $\phi_3(e_2, e_1) = 2 \neq 0 = \phi_3(e_1, e_2)$. Es ist $\phi_3(v, e_2) = 0$ für alle $v \in M(2 \times 1, K)$.

Bemerkungen 13.8 (i) Häufig ist eine Bilinearform $\phi : V \times W \to K$ durch eine Matrix $\phi(\mathcal{A}^{tr}, \mathcal{B})$ gegeben, wobei $\mathcal{A}$ und $\mathcal{B}$ irgendwelche (bekannten) Basen von V und W sind. Dann sagen Satz 13.6 (b) und die Formel in 13.6 (a) mit $C = \phi(\mathcal{A}^{tr}, \mathcal{B})$, wie man $\phi(X, Y)$ für $X \in V$ und $Y \in W$ ausrechnet:
Man bestimmt die Koeffizienten x_i und y_j in $X = \sum_{i=1}^{m} x_i a_i$ und $Y = \sum_{j=1}^{n} y_j b_j$ und berechnet das Matrizenprodukt $x^{tr} \cdot C \cdot y$, wobei $x = (x_1 \cdots x_m)^{tr}$ und $y = (y_1 \cdots y_n)^{tr}$ ist.

(ii) Das Transformationsverhalten in Lemma 13.5 (b) von Matrizen $\phi(\mathcal{A}^{tr}, \mathcal{B})$ zu Bilinearformen $\phi : V \times W \to K$ bei Basiswechseln ist anders als das von Matrizen $M(\mathcal{A}, f, \mathcal{B})$ zu Homomorphismen $f : W \to V$. Erinnerung an Bemerkung 5.12 (i):

$$M(\mathcal{A}', f, \mathcal{B}') = M(\mathcal{A}, \mathcal{A}')^{-1} \cdot M(\mathcal{A}, f, \mathcal{B}) \cdot M(\mathcal{B}, \mathcal{B}').$$

Bei Bilinearformen hat man links $(..)^{tr}$ statt $(..)^{-1}$.

(iii) (Matrizen und Bilinearformen, 5. Teil)
Sei $\phi : V \times V \to K$ eine Bilinearform auf V mit Matrix $T := \phi(\mathcal{B}^{tr}, \mathcal{B})$ bezüglich einer Basis $\mathcal{B} = (b_1, ..., b_n)$ von V. Es gilt offenbar:

$$\begin{aligned} \phi \text{ ist symmetrisch} &\iff T \text{ ist symmetrisch, d.h. } T^{tr} = T; \\ \phi \text{ ist schiefsymmetrisch} &\iff T \text{ ist schiefsymmetrisch, d.h. } T^{tr} = -T. \end{aligned}$$

Definition 13.9 Sei $\phi : V \times V \to K$ eine symmetrische oder schiefsymmetrische Bilinearform.

(a) (Definition) Zwei Elemente $x, y \in V$ heißen orthogonal, falls $\phi(x, y) = 0$ ist. Notation: $x \perp y$.

(b) (Triviales Lemma) Weil ϕ symmetrisch oder schiefsymmetrisch ist, ist $x \perp y \iff y \perp x$. [Sonst müßte man rechts-orthogonal und links-orthogonal unterscheiden.]

(c) (Definition) Sei $U \subset V$ ein Untervektorraum. Der orthogonale Untervektorraum $U^{\perp} \subset V$ ist

$$U^{\perp} := \{x \in V \mid \text{für alle } \ y \in U \ \text{ ist } x \perp y\}.$$

(d) (Definition) Das Radikal von ϕ ist $\mathrm{Rad}(\phi) := V^{\perp} \subset V$.

(e) (Definition) ϕ heißt nichtentartet, falls $\mathrm{Rad}(\phi) = \{0\}$ ist. Falls $\mathrm{Rad}(\phi) \neq \{0\}$ ist, heißt ϕ entartet.

(f) (Definition) Ein Element $v \in V$ heißt isotrop, falls es erfüllt:

$$\phi(v,v) = 0 \quad \text{und} \quad v \neq 0.$$

(g) (Triviales Lemma) Offenbar ist jedes Element von $\mathrm{Rad}(\phi) - \{0\}$ isotrop.

Beispiele 13.10 (i) Es gibt in $\mathcal{C}^0([0,1],\mathbb{R})$ bezüglich ϕ_{int} (Beispiel 13.7 (i)) kein isotropes Element, denn $f^2(x) \geq 0$ für alle $x \in [0,1]$, und $f \in \mathcal{C}^0([0,1],\mathbb{R})$ erfüllt (Beweis: Analysis)

$$\int_0^1 f^2(x)\mathrm{d}x = 0 \iff f = 0.$$

Insbesondere ist $\mathrm{Rad}(\phi_{int}) = \{0\}$. Also ist ϕ_{int} nichtentartet.

(ii) Die Bilinearform ϕ_2 von Beispiel 13.7 (ii) ist nichtentartet (Übung oder Lemma 13.11 (c) und $\det \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = -1 \neq 0$). Aber es gibt isotrope Elemente:

$$\begin{aligned} x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq 0 \text{ ist isotrop} \quad &\iff \quad \begin{pmatrix} x_1 & x_2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \\ &\iff \quad x_1^2 - x_2^2 = 0. \end{aligned}$$

Also sind zum Beispiel $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ isotrop. [Bei $K = \mathbb{F}_2$ sind sie natürlich gleich.] Bei $U := K \cdot \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ ist

$$U^\perp = U.$$

Lemma 13.11 *Sei V ein endlich-dimensionaler K-Vektorraum, $\phi : V \times V \to K$ eine symmetrische oder schiefsymmetrische Bilinearform, $\mathcal{B} = (b_1, ..., b_n)$ eine Basis von V und $A := \phi(\mathcal{B}^{tr}, \mathcal{B})$.*

(a) Sei $U \subset V$ ein Untervektorraum. Sei $\mathcal{C} = (c_1, ..., c_k)$ eine Basis von U, und sei $\Gamma = (\gamma_{ij})$ die $(n \times k)$-Matrix mit $c_j = \sum_{i=1}^n \gamma_{ij} b_i$ (also $\mathcal{C} = \mathcal{B} \cdot \Gamma$). Dann ist

$$\begin{aligned} &\mathrm{rang}\,\Gamma = k \geq \mathrm{rang}\left(\Gamma^{tr} \cdot A\right), \\ &U^\perp = \{\sum_{i=1}^n x_i b_i \mid \Gamma^{tr} \cdot A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0\}, \\ &\dim U^\perp = n - \mathrm{rang}(\Gamma^{tr} \cdot A) \geq n - k = n - \dim U, \\ &\dim U^\perp + \dim U \geq n. \end{aligned}$$

(b) Insbesondere ist

$$\begin{aligned} &\mathrm{Rad}(\phi) = \{\sum_{i=1}^n x_i b_i \mid A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0\}, \\ &\dim \mathrm{Rad}(\phi) = n - \mathrm{rang}\, A. \end{aligned}$$

(c) ϕ ist nichtentartet $\iff \det A \neq 0$.

(d) Wenn ϕ nichtentartet ist, ist

$$\dim U^\perp + \dim U = n.$$

e) Wenn es keine isotropen Vektoren gibt ($\Rightarrow \mathrm{Rad}(\phi) = \{0\} \Rightarrow \phi$ ist nichtentartet), ist $U \cap U^\perp = \{0\}$ und

$$U \oplus U^\perp = V.$$

Beweis: (a) Die Ungleichung $\operatorname{rang} \Gamma \geq \operatorname{rang}(\Gamma^{tr} \cdot A)$ ist klar. Es ist

$$\phi(c_j, \sum_{i=1}^{n} x_i b_i) = (\gamma_{1j} \cdots \gamma_{nj}) \cdot A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

wegen Bemerkung 13.8 (i) bzw. Satz 13.6. Daraus folgt die Beschreibung von $U^\perp$. Der Rest ist klar.
(b) Das folgt aus (a) und der Definition des Radikals.
(c) Das folgt aus (b).
(d) In dem Fall ist $\operatorname{rang} \Gamma = \operatorname{rang}(\Gamma^{tr} \cdot A)$ und $\dim U^\perp = n - k = n - \dim U$.
(e) Jeder Vektor in $U \cap U^\perp - \{0\}$ wäre isotrop. Wenn es keine isotropen Vektoren gibt, ist $U \cap U^\perp = \{0\}$. Wegen (d) ist dann $U \oplus U^\perp = V$. □

Bemerkung 13.12 Ist $A \in M(n \times n, K)$ (schief)symmetrisch und $T \in M(n \times n, K)$ beliebig, so ist auch $T^{tr} \cdot A \cdot T$ (schief)symmetrisch. Denn wegen $(B \cdot C)^{tr} = C^{tr} \cdot B^{tr}$ (Lemma 4.15 b)) ist

$$(T^{tr} \cdot A \cdot T)^{tr} = T^{tr} \cdot A^{tr} \cdot T = (-)T^{tr} \cdot A \cdot T.$$

Als letzter Punkt im Kapitel 13 kommt ein Ausblick auf eine Variante von Bilinearformen, die über dem Körper $\mathbb{C}$ zu einem Analogon von Euklidischen Vektorräumen führt, den *unitären Vektorräumen*.

Definition 13.13 Sei V ein $\mathbb{C}$-Vektorraum.
(a) (Erinnerung, vgl. Beispiel 9.2 (iii)) Eine Abbildung $f : V \to \mathbb{C}$ ist *linear*, also eine *Linearform*, wenn sie ein Vektorraumhomomorphismus ist, d.h. wenn sie erfüllt:

$$f(v + w) = f(v) + f(w) \text{ (Additivität)},$$
$$f(\lambda \cdot v) = \lambda \cdot f(v) \text{ (Linearität bezüglich skalarer Multiplikation)}.$$

Hier sind $v, w \in V$ und $\lambda \in \mathbb{C}$.

(b) Eine Abbildung $f : V \to \mathbb{C}$ ist *semilinear*, wenn sie erfüllt:

$$f(v + w) = f(v) + f(w) \text{ (Additivität)},$$
$$f(\lambda \cdot v) = \overline{\lambda} \cdot f(v) \text{ (Semilinearität bezüglich skalarer Multiplikation)}.$$

[*semi* $\sim$ ein halb.]

(c) Eine Abbildung $\phi : V \times V \to \mathbb{C}$ ist *sesquilinear* wenn sie im linken Argument linear und im rechten semilinear ist, d.h. wenn für jedes $y \in V$ die Abbildung

$$\phi(.,y) : V \to \mathbb{C},\ x \mapsto \phi(x,y)$$

linear ist und wenn für jedes $x \in V$ die Abbildung

$$\phi(x,.) : V \to \mathbb{C},\ y \mapsto \phi(x,y)$$

semilinear ist. Eine sesquilineare Abbildung heißt auch *Sesquilinearform*.
[*sesqui* $\sim$ anderthalb.]

(d) Eine Sesquilinearform $\phi : V \times V \to \mathbb{C}$ heißt *hermitesch*, falls gilt:

$$\phi(y,x) = \overline{\phi(x,y)} \text{ für alle } x, y \in V$$

(wegen $\phi(x,x) = \overline{\phi(x,x)}$ ist in jedem Fall $\phi(x,x) \in \mathbb{R}$).
Eine Matrix $M \in M(n \times n, \mathbb{C})$ heißt *hermitesch*, falls

$$A = \overline{A}^{tr} \quad (= \overline{A^{tr}})$$

ist.

(e) (Leichtes Lemma, Beweis Übung) Sei $\mathcal{B} = (b_1, ..., b_n)$ eine Basis eines $\mathbb{C}$-Vektorraums, und sei $\phi : V \times V \to \mathbb{C}$ eine Sesquilinearform. Dann gilt

$$\phi \text{ ist hermitesch } \iff \phi(\mathcal{B}^{tr}, \mathcal{B}) := (\phi(b_i, b_j)) \text{ ist hermitesch.}$$

(f) Eine hermitesche Sesquilinearform $\phi : V \times V \to \mathbb{C}$ heißt *positiv definit* [bzw. *negativ definit*], falls gilt:

$$\phi(x,x) > 0 \quad [\text{ bzw. } < 0] \quad \text{ für alle } x \in V - \{0\}.$$

(g) Ein $\mathbb{C}$-Vektorraum V zusammen mit einer positiv definiten hermiteschen Sesquilinearform $\phi : V \times V \to \mathbb{C}$ heißt *unitärer Vektorraum*. Die Sesquilinearform ϕ ist sein *Skalarprodukt*. Die *Norm* (oder *Länge*) eines Vektors v ist

$$\|v\| := \sqrt{\phi(v,v)} \geq 0.$$

Bemerkungen 13.14 (i) Sesquilinearformen sind eng verwandt mit Bilinearformen über dem Körper $\mathbb{R}$. *Hermitesch* bei Sesquilinearformen ist das Analogon zu *symmetrisch* bei Bilinearformen. Es gilt der

1. Meta-Satz: Alle Konstruktionen und Resultate, die in Kapitel 9 und hier in Kapitel 13 für symmetrische Bilinearformen ausgeführt wurden, lassen sich mit kleinen Modifikationen auch für hermitesche Sesquilinearformen ausführen.

(ii) *Unitärer Vektorraum* ist das Analogon bei $K = \mathbb{C}$ zum *Euklidischen Vektorraum* bei $K = \mathbb{R}$. Es gilt der

2. Meta-Satz: Abgesehen von der Definition des Winkels lassen sich alle Konstruktionen und alle Resultate, die in Kapitel 9 für Euklidische Vektorräume ausgeführt wurden, mit kleinen Modifikationen auch für unitäre Vektorräume ausführen.

(iii) Diese beiden Meta-Sätze müßten natürlich bewiesen werden, indem man alle Konstruktionen und Beweise durchgeht und gegebenenfalls modifiziert. Doch dafür ist in dieser Vorlesung keine Zeit mehr. Nur einige Bemerkungen und vier Beispiele kommen noch.

(iv) Bei einer hermiteschen Sesquilinearform ϕ ist $\phi(v,v) = \overline{\phi(v,v)}$, also $\phi(v,v) \in \mathbb{R}$. Nur deshalb hat ϕ eine Chance, positiv definit zu sein.

(v) Bei der Definition des Winkels in Kapitel 9 (Definition 9.10) war $\phi(x,y) \in \mathbb{R}$ für alle $x, y \in V$ wichtig gewesen. Das ist bei hermiteschen Sesquilinearformen nicht erfüllt.

(vi) Das Transformationsverhalten von Matrizen zu Sesquilinearformen bezüglich Basiswechseln ist fast gleich zum Transformationsverhalten von Matrizen zu Bilinearformen (Lemma 13.5 (b)). Nur bei der rechten Basiswechselmatrix braucht man nun eine komplexe Konjugation.

(vii) Bei Sesquilinearformen sind die Begriffe orthogonaler Unterraum, Radikal, nichtentartet und isotrop ganz analog zu Definition 13.9 definiert. Lemma 13.11 hat ein fast gleichlautendes Analogon bei Sesquilinearformen.

(viii) Auch die Begriffe Orthogonalbasis und Orthonormalbasis sind bei einem unitären Vektorraum ganz analog wie bei einem Euklidischen Vektorraum definiert. Und das Gram-Schmidtsche Orthogonalisierungsverfahren funktioniert mit den gleichen Formeln wie im Fall eines Euklidischen Vektorraums.

(ix) In einem unitären Vektorraum V mit Skalarprodukt ϕ gilt auch die Cauchy-Schwarzsche Ungleichung:

$$|\phi(x,y)| \leq \|x\| \cdot \|y\| \qquad \text{für } x, y \in V,$$

und Gleichheit gilt genau dann, wenn x und y linear abhängig sind. Allerdings ist der Beweis etwas schwerer als im Fall eines Euklidischen Vektorraums.

Beispiele 13.15 (i) Der wichtigste unitäre Vektorraum ist der $\mathbb{C}^n$ mit dem Standardskalarprodukt $\phi_{st,\mathbb{C}}$,

$$\phi_{st,\mathbb{C}} : ((x_1,...,x_n),(y_1,...,y_n)) \overset{\phi_{st,\mathbb{C}}}{\mapsto} (x_1 \cdots x_n) \cdot \begin{pmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{pmatrix} = \sum_{i=1}^{n} x_i \overline{y_i}.$$

(ii) Der $\mathbb{C}$-Vektorraum $\mathcal{C}^0([0,1],\mathbb{C})$ mit der Sesquilinearform $\phi_{int,\mathbb{C}}$,

$$\phi_{int,\mathbb{C}} : (f,g) \mapsto \int_0^1 f(x)\overline{g(x)}\mathrm{d}x,$$

ist ein unitärer Vektorraum wegen (Beweis: Analysis)

$$\int_0^1 |f(x)|^2\mathrm{d}x > 0, \text{ falls } f \neq 0.$$

(iii) Eine Matrix $A = (a_{ij}) \in M(n \times n, \mathbb{C})$ definiert (offenbar) eine Sesquilinearform Sesq_A auf $M(n \times 1, \mathbb{C})$ durch

$$(x,y) \;\mapsto\; x^{tr} \cdot A \cdot \overline{y} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i \cdot a_{ij} \cdot \overline{y_j}.$$

(iv) (Gram-Schmidtsches Orthogonalisierungsverfahren) Sei V ein unitärer Vektorraum mit Skalarprodukt ϕ und einer Basis $\mathcal{A} = (a_1, a_2)$ mit

$$\begin{aligned} M(\mathcal{A},\phi,\mathcal{A}) &= \begin{pmatrix} 2 & -3i \\ 3i & 5 \end{pmatrix}, \\ \text{d.h.} \qquad & \phi(a_1,a_1) = 2,\ \phi(a_2,a_2) = 5,\ \phi(a_1,a_2) = -3i,\ \phi(a_2,a_1) = 3i. \end{aligned}$$

Mit dem Gram-Schmidt-Verfahren erhält man folgende Orthogonalbasis $\mathcal{B} = (b_1, b_2)$:

$$\begin{aligned} b_1 &:= a_1, \\ b_2 &:= a_2 - \frac{\phi(a_2,b_1)}{\phi(b_1,b_1)} \cdot b_1 = a_2 - \frac{\phi(a_2,a_1)}{\phi(a_1,a_1)} \cdot a_1 = a_2 - \frac{3i}{2} \cdot a_1. \end{aligned}$$

Normieren:

$$\begin{aligned} \|b_1\|^2 &= \phi(b_1,b_1) = \phi(a_1,a_1) = 2, \\ c_1 &:= \frac{b_1}{\|b_1\|} = \frac{b_1}{\sqrt{2}}, \\ \|b_2\|^2 &= \phi(b_2,b_2) = \phi(a_2 - \frac{3i}{2}a_1, a_2 - \frac{3i}{2}a_1) \\ &= \phi(a_2,a_2) - \frac{3i}{2}\phi(a_1,a_2) - \frac{-3i}{2}\phi(a_2,a_1) + \frac{-3i}{2}\frac{3i}{2}\phi(a_1,a_1) \\ &= 5 - \frac{3i}{2}(-3i) - \frac{-3i}{2}(3i) + \frac{9}{4} \cdot 2 \\ &= 5 - \frac{9}{2} - \frac{9}{2} + \frac{9}{4} \cdot 2 = \frac{1}{2}, \\ c_2 &:= \frac{b_2}{\|b_2\|} = \sqrt{2} \cdot b_2. \end{aligned}$$