Teil III

Algebra und Geometrie
Vorlesung
Wintersemester 1998/99

Volker Mehrmann

Übungen: Matthias Bollhöfer
Matthias Pester

Seminar: Uwe Schrader

Kapitel 1

Ringe, Ideale und Restklassenringe

In den ersten beiden Teilen der Vorlesung haben wir schon einige algebraische Grundbegriffe, wie Gruppe, Körper, Ring kennengelernt. Wir wollen uns mit diesen algebraischen Strukturen nun vertieft beschäftigen.
Sei dazu R ein kommutativer Ring mit Einselement 1, siehe Teil I, Definition 1.1. Als Beispiele können wir z.B. $R = \mathbb{Z}$ oder $R = K[x] = \left\{ \sum_{i=0}^{n} a_i x^i \mid a_i \in K \right\}$, den Ring der Polynome in einer Variablen über dem Körper K, nehmen.

Definition 1.1 *Eine Teilmenge $I \subset R$ eines kommutativen Rings mit Einselement heißt* Ideal in R, *wenn*

(i) $(I, +) \subset (R, +)$ *eine Untergruppe ist und* $x \cdot r \in I$ *für alle* $x \in I$ *und* $r \in R$.

(ii) Ein Ideal heißt echtes Ideal, *falls* $I \neq R$.

(iii) Ein Ideal heißt maximales Ideal, *wenn* $I \neq R$ *und es kein Ideal* J *gibt mit* $I \subsetneq J \subsetneq R$.

(iv) Ein Ideal I *in* R *heißt* Hauptideal, *falls* $I = xR = \{xr \mid r \in R\}$ *für ein* $x \in R$. *Man sagt dann auch, daß* I von x erzeugt *ist.*

Wir betrachten einige Beispiele:

Beispiel 1.2

(a) Natürlich sind $\{0\} \subset R$ und R Ideale.

(b) Sei $R = \mathbb{Z}$ und I die Menge der geraden Zahlen, so ist I ein Hauptideal und $I = 2 \cdot \mathbb{Z}$.

(c) Ist $R = K$ ein Körper, so sind $\{0\}$ und R die einzigen Ideale in R, denn sei $\{0\} \neq I \subset R$ und $0 \neq x \in I$. Da K ein Körper ist, so gibt es ein $\tilde{x} \in K$ mit $x\tilde{x} = 1$. Also ist $1 \in I$ und damit $R \subset I$ und damit $R = I$.

Wir können auf einem Ring eine Äquivalenzrelation einführen durch die folgende Relation: Sei $I \subset R$ ein Ideal. Für $r, s \in R$ setzen wir $r \sim s$, falls $r - s \in I$. Dies ist eine Äquivalenzrelation, denn falls $r, s, t \in R$ und

(i) $r - s \in I$, so ist $s - r = (-r) - (-s) \in I$,

(ii) $r - r = 0 \in I$,

(iii) $r - s \in I,\ s - t \in I \implies (r - s) + (s - t) \in I$
$\implies r - t \in I.$

Für diese Äquivalenzrelation „$\sim$“ können wir nun die Äquivalenzklassen definieren mittels $[r] = \{s \in R \mid r - s \in I\}$. Die Menge dieser Äquivalenzklassen bezeichnen wir wie üblich mit R/I und nennen sie den *Restklassenring* nach I oder *Faktorring*. Wir können auf dieser Menge die Operationen „$+, \cdot$“ definieren über

$$\begin{aligned} [r_1] + [r_2] &= [r_1 + r_2] \\ [r_1] \cdot [r_2] &= [r_1 \cdot r_2] \end{aligned} \tag{1.3}$$

Im folgenden lassen wir die Klammern oft weg, d.h., wir schreiben nur den Repräsentanten r für die Klasse $[r]$.

Lemma 1.4 *Die Menge der Äquivalenzklassen R/I mit den Operationen $+, \cdot$ wie in (1.3) ist ein kommutativer Ring mit Eins.*

Beweis: Übungsaufgabe □

Definition 1.5 *Sei X eine Menge. Eine Teilmenge $H \subset X \times X$ heißt* Halbordnung *auf X falls*

(i) $(x, x) \in H\ \forall x \in X$

(ii) $(x, y) \in H$ *und* $(y, x) \in H$, *so gilt* $x = y$.

(iii) $(x, y) \in H$ *und* $(y, z) \in H$, *so gilt* $(x, z) \in H$.

Beispiel: „$\leq$“ oder Teilmengeninklusion.
Wir brauchen im folgenden ein wichtiges Lemma, das Lemma von Zorn, das wir aber nicht beweisen werden. Dazu zuerst eine Definition.

Definition 1.6 *Eine nichtleere Teilmenge A einer Menge X mit einer Halbordnung „$\leq$“ heißt* vollständig geordnet, *falls für alle $a, b \in A$ stets $a \leq b$ oder $b \leq a$ gilt.*
Ein Element $s(A) \in X$ heißt obere Schranke *von A, falls $a \leq s(A)$, $\forall a \in A$.*
Ein Element $m(A) \in A$ heißt maximales Element *von A, wenn aus $a \in A$ und $m(A) \leq a$ stets $m(A) = a$ folgt.*
Eine Menge X mit Halbordnung heißt induktiv geordnet, *wenn jede vollständig geordnete Teilmenge von X eine obere Schranke in X besitzt.*

Lemma 1.7 (Lemma von Zorn) *Ist X eine nichtleere Menge, die bezüglich einer Halbordnung „$\leq$“ induktiv geordnet ist, so gibt es zu jedem $a \in X$ ein maximales Element $m \in X$ mit $a \leq m$.*

Beweis: Sehr komplizierter Beweis mit Methoden der Mengenlehre. Diesen Beweis machen wir hier nicht. □

Nun kommen wir zu den Ringen zurück.

Lemma 1.8 *Sei R ein kommutativer Ring mit Eins.*

(i) Jedes echte Ideal A in R ist Teilmenge eines maximalen Ideals.

(ii) Ist I ein maximales Ideal in R, so ist $K = R/I$ ein Körper.

Beweis:

(i) Sei A ein echtes Ideal.
Betrachte die Menge X der von R verschiedenen Ideale von R, die das echte Ideal A enthalten, d.h.

$$X = \{B \mid B \text{ Ideal in } R,\ B \neq R,\ A \subset B\}$$

Nach Voraussetzung ist $X \neq \emptyset$.
Ein maximales Ideal ist ein maximales Element von X bezüglich der durch Inklusion gegebenen Halbordnung. Wir zeigen jetzt, daß X induktiv geordnet ist und wenden das Lemma von Zorn an.

Da $A \in X$, ist X nicht leer. Sei $Y \neq \emptyset$ eine vollständig geordnete Teilmenge von X. Zeige nun, daß $s(Y) = \bigcup_{B \in Y} B$ obere Schranke von Y in X ist.
Da $B \in s(Y),\ \forall B \in Y$, müssen wir nur noch zeigen, daß $s(Y)$ in X liegt. Seien $a, b \in s(Y)$ und sei $a \in B_a \in Y,\ b \in B_b \in Y$. Da Y vollständig geordnet ist, folgt $B_a \subset B_b$ oder $B_b \subset B_a$. Falls $B_a \subset B_b$, so folgt $a, b \in B_b$ (der andere Fall ist analog), also auch $b - a \in B_b \subset s(Y)$ und $ra, ar \in B_b \subset s(Y)$ für alle $r \in R$. Also ist $s(Y)$ ein Ideal.
Falls $1 \notin B$, so ist $B \neq R$. Wenn alle $B \in Y$ die 1 nicht enthalten, so auch $s(Y)$ nicht, und damit ist $s(Y) \neq R$ und da $s(Y) \in X$, folgt mit dem Zorn'schen Lemma, daß $s(Y)$ obere Schranke ist.

(ii) Wir müssen zeigen, daß jedes $[x] \in (R/I) \setminus \{0\}$ invertierbar ist.

Betrachte $I + xR$. Dies ist ein Ideal in R, welches I echt enthält. Da I maximal ist, folgt $I + xR = R$ und damit gibt es $m \in I$, $s \in R$ mit $m + xs = 1$. Also ist $[x] \cdot [s] = 1$ oder $[s] = [x]^{-1}$.

□

Beachte, die Umkehrung in Teil (ii) von Lemma 1.8 gilt auch.

Definition 1.9 *Seien R_1, R_2 Ringe.*
Eine Abbildung $\varphi : R_1 \to R_2$ heißt Ringhomomorphismus, *wenn für alle $x, y \in R_1$ gilt*

$$\begin{cases} \varphi(x+y) &= \varphi(x) + \varphi(y) \\ \varphi(x \cdot y) &= \varphi(x) \cdot \varphi(y) \\ \varphi(1_{R_1}) &= 1_{R_2} \end{cases} \tag{1.10}$$

$$\begin{aligned} \text{Bild}\,\varphi &= \varphi(R_1) = \{\varphi(x) \in R_2 \mid x \in R_1\} \\ \text{Kern}\,\varphi &= \{x \in R_1 \mid \varphi(x) = 0\}. \end{aligned}$$

Beispiel 1.11 Sei $\varphi : \mathbb{Z} \to \mathbb{Z}$ ein Ringhomomorphismus. Dann gilt

$$\begin{aligned} \varphi(n) &= \varphi(1 + \ldots + 1) = \varphi(1) + \ldots + \varphi(1) = n\varphi(1) \text{ und} \\ \varphi(n) &= \varphi(n \cdot 1) = \varphi(n) \cdot \varphi(1). \end{aligned}$$

und damit $\varphi(n)\varphi(1) = \varphi(n) = n\varphi(1)$.
Wäre $\varphi(1) = 0$, so folgt $\varphi(n) = 0\ \forall n \in \mathbb{Z}$. Ansonsten kann man kürzen durch $\varphi(1)$ und erhält $\varphi(n) = n$ also $\varphi = id$.

Beispiel 1.12

(i) Sei $R = \mathbb{Z}$ und $n\mathbb{Z} = \{nx \mid x \in \mathbb{Z}\}$.
Dann ist $\mathbb{Z}/n\mathbb{Z}$ ein Ring mit n Elementen und es gilt, daß $n\mathbb{Z}$ genau dann ein maximales Ideal ist, wenn n eine Primzahl ist.

Beweis: Sei $I = p\mathbb{Z}$ für eine Primzahl p und sei Y mit $I \subset Y \subset \mathbb{Z}$ ein Ideal. Sei $x \in Y$ mit $|x| \neq 0$ minimal und sei $y \in Y \setminus \{0\}$. Teile y durch x mit Rest, also $y = qx + r$ mit $0 \leq r < |x|$. Da x minimal, so ist $r = 0$ oder, anders gesagt, Y ist von x erzeugt. Da $p \in Y$, so gibt es $z \in R$ mit $p = xz$ und da p eine Primzahl ist, folgt $x = 1$ oder $z = 1$, also $Y = R$ oder $Y = I$ und d.h., daß $I = p\mathbb{Z}$ maximal ist und mit Lemma 1.8 (ii) ist damit $\mathbb{Z}/p\mathbb{Z} =: \mathbb{F}_p$ ein Körper. □

(ii) Sei $R = K[x]$ und $f = \sum_{i=0}^{n} \alpha_i x^i \in R$. Sei I das Ideal, welches von $x \in R$ erzeugt wird, so ist

$$\begin{aligned} R/I &\to K \\ [f] &\to \alpha_0 \end{aligned}$$

ein Ringisomorphismus.

Definition 1.13 *Ein nullteilerfreier, kommutativer Ring, der wenigstens ein von Null verschiedenes Element enthält, heißt* Integritätsring.

Beispiel 1.14 $\mathbb{Z}$ und $n \cdot \mathbb{Z}$ sind Integritätsringe und natürlich sind auch alle Körper Integritätsringe. Das Bild eines Integritätsringes unter einem Homomorphismus ist jedoch i.a. kein Integritätsring, denn $\mathbb{Z}/4\mathbb{Z}$ ist homomorphes Bild von $\mathbb{Z}$, hat jedoch wegen $2 \cdot 2 = 0$ Nullteiler.

Definition 1.15 *Sei R ein kommutativer Ring. Ein Ideal P von R heißt* Primideal, *falls*

(i) $P \neq R$.

(ii) Aus $a, b \in R$ und $a \cdot b \in P$ folgt stets $a \in P$ oder $b \in P$.

Beispiel 1.16 In einem kommutativen Ring $R \neq \{0\}$ ist $\{0\}$ ein Primideal genau dann, wenn R ein Integritätsring ist, denn falls $a \cdot b = 0$, so folgt aus der Nullteilerfreiheit $a = 0$ oder $b = 0$.
In $\mathbb{Z}$ ist $\{0\}$ ein Primideal und $\{p\}$ mit $p > 0$ ein Primideal genau dann, wenn p Primzahl ist.

Satz 1.17 *Es sei R ein kommutativer Ring und $P \neq R$ ein Ideal von R. Dann sind die folgenden Aussagen äquivalent:*

(a) P ist Primideal.

(b) $a, b \in R$ mit $a \notin P$ und $b \notin P \Longrightarrow ab \notin P$.

(c) $R \setminus P$ ist mit der Multiplikation aus R eine Halbgruppe.

(d) R/P ist Integritätsring.

(e) Es gibt einen Homomorphismus $f : R \to S$ (S Integritätsring) mit $P = \text{Kern } f$.

Beweis:
(a) $\Longrightarrow$ (b)
Wäre $a \cdot b \in P$, so auch $a \in P$ oder $b \in P$ nach Def. 1.15 (ii) und das ist ein Widerspruch.

(b) $\Longrightarrow$ (c)
Aus (b) folgt, daß $R \setminus P$ abgeschlossen bezüglich der Multiplikation in R ist und da das Assoziativgesetz in R gilt, folgt (c).

(c) $\Longrightarrow$ (d)
Da $R/P \neq \{0\}$ und kommutativ ist, folgt für $[a], [b] \in R/P$ mit $[a], [b] \neq 0$, i.e. $a \notin P, b \notin P$ auch, daß $ab \notin P$ also $[a] \cdot [b] \neq 0$. Also gibt es in R/P keine Nullteiler und es ist ein Integritätsring.

(d) $\Longrightarrow$ (e)
Die kanonische Abbildung (vgl. Def. 12.10, Teil I)

$$\Pi : R \to R/P\,,$$

die jedem Element aus R die Äquivalenzklasse zuordnet, zu der es gehört, ist der gewünschte Homomorphismus. ($\Pi(a) = [0]\ \forall a \in P$)

(e) $\Longrightarrow$ (a)
Sei $f : R \to S$ Homomorphismus in den Integritätsring S und $P = \text{Kern } f$. Da $P \neq R$, so gilt falls $a \cdot b \in \text{Kern } f$, $0 = f(ab) = f(a) \cdot f(b)$ und da S nullteilerfrei ist, folgt $f(a) = 0$ oder $f(b) = 0 \implies a \in P$ und $b \in P$. □

Korollar 1.18 *In einem kommutativen Ring mit $1 \neq 0$ ist jedes maximale Ideal auch Primideal.*

Beweis: Es sei M ein maximales Ideal des Ringes R. Nach Lemma 1.8 ist R/M ein Körper, insbesondere ein Integritätsring, also nach Satz 1.17 ein Primideal. □

Beachte, man kann nicht darauf verzichten, daß der Ring eine 1 hat, denn mit $R = 2\mathbb{Z}$ und $M = 4\mathbb{Z}$ ist $R/M = \{0, 2\}$ kein Integritätsring.

Die ganzen Zahlen $\mathbb{Z}$ sind Unterring von $\mathbb{Q} = \left\{\frac{m}{n} \mid m, n \in \mathbb{Z}\right\}$. So eine Konstruktion können wir auch für jeden anderen Integritätsring machen.

Sei R ein Integritätsring. Betrachte alle Paare (x, y) mit $x, y \in R,\ y \neq 0$, und setze

$$F(R) = \left\{\frac{x}{y} \mid x, y \in R, y \neq 0\right\},$$

dann ist $F(R)$ nicht leer, da R mindestens ein von Null verschiedenes Element enthält, und wir definieren auf $F(R)$ eine Äquivalenzrelation durch $\frac{x}{y} \sim \frac{a}{b}$ genau dann, wenn $xb = ay$.

Die Menge der Äquivalenzklassen von $\sim$ bezeichnen wir mit $Q(R)$ und definieren dann Addition und Multiplikation durch

$$\begin{cases} \dfrac{x}{y} + \dfrac{a}{b} &= \dfrac{xb + ay}{yb} \\ \dfrac{x}{y} \cdot \dfrac{a}{b} &= \dfrac{xa}{yb} \end{cases} \tag{1.19}$$

Damit erhalten wir

Satz 1.20 *Sei R ein Integritätsring, so ist $(Q(R), +, \cdot)$ ein Körper.*

Beweis: Übung □

$Q(R, +, \cdot)$ heißt *Quotientenkörper von R*.
Wir übertragen nun die Grundlagen der Arithmetik auf Integritätsringe.

Definition 1.21 *Sei R ein Integritätsring mit 1 und seien $x, y \in R$. x heißt* Teiler *von y (oder y ein* Vielfaches *von x), wenn es $z \in R$ gibt mit $y = xz$. Ist x Teiler von y, so schreiben wir $x|y$.*

Satz 1.22 *Sei R ein Integritätsring mit 1. Es gelten die folgenden Rechenregeln.*

(a) $1|x, x|0$ *und* $x|x$ *für alle* $x \in R$.

(b) x *ist Teiler von* 1 *genau dann, wenn x eine Einheit (d.h. invertierbar) in R ist.*

(c) Aus $x|y$ *folgt* $xz|yz$ *für alle* $z \in R$.

(d) Aus $x|y_i\ (1 \leq i \leq n)$ *folgt* $x|z_1y_1 + \ldots + z_ny_n$ *für alle* $z_i \in R$.

(e) Aus $x|y$ *und* $y|z$ *folgt* $x|z$.

(f) $x|y \iff y \in xR \iff yR \subset xR$.

(g) $x|y$ *und* $y|x \iff xR = yR \iff x = y \cdot z$ *mit z Einheit aus der Menge der invertierbaren Elemente von R.*

Beweis:

(a) - (e) trivial.

(f) $x|y \implies \exists r \in R$ mit $y = rx$, d.h. $y \in xR$ also $yR \subset xR$.

(g) $x|y$ und $y|x \iff yR \subset xR$ und $xR \subset yR \iff xR = yR$, also gibt es $u, v \in R$ mit $x = uy$, $y = vx \implies x = uvx$ und durch Kürzen $1 = uv$. Also ist u eine Einheit.
Ist umgekehrt $x = uy$ mit u invertierbar und Einheit, so ist y Teiler von x und da $y = xu^{-1}$, so auch x Teiler von y.

□

Definition 1.23 *Zwei Elemente x, y eines Integritätsringes R heißen* assoziiert, *falls es eine invertierbare Einheit u gibt mit $x = yu$ (wir schreiben $x \sim y$).*
Es seien $p, q \neq 0$ Nichteinheiten eines Integritätsringes R mit 1. p heißt Primelement *von R, wenn für $x, y \in R$ aus $p|xy$ stets $p|x$ oder $p|y$ folgt.*
q heißt unzerlegbar *oder* irreduzibel, *wenn $q = xy$ nur für x oder y invertierbar möglich ist.*

Ein Teiler x von y heißt echt, *falls x weder eine Einheit, noch mit y assoziiert ist.*

Beachte, Primelement und irreduzibles Element sind in $\mathbb{Z}$ dasselbe, im allgemeinen nicht.

Satz 1.24 *Sei R ein Integritätsring mit 1 und $p \in R \setminus \{0\}$ keine Einheit von R. Dann gilt*

(i) Jedes Primelement in R ist irreduzibel.

(ii) p ist ein Primelement genau dann, wenn pR ein Primideal ist.

(iii) p ist irreduzibel genau dann, wenn pR in der Menge der Hauptideale von R (die nicht gleich R sind) bezüglich Inklusion maximal ist.

Beweis:

(i) Sei p ein Primelement und $p = xy$. Da $p|p = xy$, so folgt nach Definition 1.23, daß $p|x$ oder $p|y$.
Falls $x = pr$, so ergibt sich $p = xy = pry \implies y$ ist Einheit.
Für $y = pr$ analog.

(ii) Sei p ein Primelement. Da p nicht invertierbar, so folgt $pR \neq R$. Sei $x, y \in pR$, d.h. $xy = rp$ mit $r \in R$. Dann gilt $p|xy \implies p|x$ oder $p|y \implies x \in pR$ oder $y \in pR$. Damit ist pR ein Primideal. Umgekehrt sei $pR \neq R$ Primideal, dann ist p keine Einheit. Falls $p = xy$, so folgt $xy \in pR$ also $x \in pR$ oder $y \in pR$, d.h. $p|x$ oder $p|y$.

(iii) Sei p irreduzibel und $p = xr$ mit $r \in R$. Da eine solche Zerlegung nur für invertierbare x und r gilt, folgt im ersten Fall $xR = R$ und im zweiten Fall, daß p und x assoziiert sind, also $pR = xR$. Ist umgekehrt $pR \neq R$ maximal unter den Hauptidealen von R und $p = xy$ mit x nicht invertierbar. Dann ist $xR \neq R$ und $pR \subset xR$. Aus der Maximalität folgt $pR = xR$, also ist p assoziiert mit x, d.h. $p = xy = x \cdot u$ mit u invertierbar. Kürzen ergibt $y = u$, also ist p irreduzibel.

□

Definition 1.25 *Ein Ring R heißt* Hauptidealring, *falls R ein Integritätsring mit 1 ist und jedes Ideal von R ein Hauptideal ist.*

Natürlich ist $\mathbb{Z}$ ein Hauptidealring.

Satz 1.26 *Sei R ein Hauptidealring und $p \in R \setminus \{0\}$.*

(i) p ist Primelement genau dann, wenn p irreduzibel.

(ii) pR ist Primideal genau dann, wenn pR maximales Ideal ist.

Beweis: Mit Satz 1.24 und Korollar 1.18 folgt:
p Primelement $\Longrightarrow$ p irreduzibel $\Longrightarrow$ pR maximal $\Longrightarrow$ pR Primideal $\Longrightarrow$ p Primelement. □

Damit können wir nun Begriffe wie ggT oder kgV einführen.

Definition 1.27 *Sei R ein Integritätsring mit 1.*
Das Element $d \in R$ heißt größter gemeinsamer Teiler (ggT) *von $a_1, \ldots, a_n \in R$, wenn d jedes a_i, $1 \leq i \leq n$, teilt und wenn jeder gemeinsame Teiler von allen a_i, $1 \leq i \leq n$, auch Teiler von d ist. Wir schreiben $d = (a_1, \ldots, a_n)$.*
Wenn jeder gemeinsame Teiler von $a_1, \ldots, a_n$ eine Einheit ist, so heißen $a_1, \ldots, a_n$ teilerfremd *oder* relativ prim.

Beachte: Ein ggT ist nur bis auf Einheiten eindeutig bestimmt.

Satz 1.28 *Sei R ein Hauptidealring und $d, a_1, \ldots a_n \in R$. Dann ist d ein ggT von $a_1, \ldots, a_n$ genau dann, wenn das von $a_1, \ldots, a_n$ erzeugte Ideal gerade dR ist.*

Beweis: Falls das von $a_1, \ldots, a_n$ erzeugte Ideal dR ist, so folgt $a_i = d\,r_i$ für $i = 1, \ldots, n$, d.h. $d | a_i$. Ist c ein weiterer gemeinsamer Teiler der a_i, dann teilt c jedes Element von $r_1 a_1 + \ldots + r_n a_n$ (dies folgt aus Satz 1.22), also auch d. Also ist d ggT. Ist umgekehrt d ein ggT von $a_1, \ldots, a_n$ und wird das von $a_1, \ldots, a_n$ erzeugte Ideal, welches ein Hauptideal ist, von einem $\tilde{d}$ erzeugt, dann ist $\tilde{d}$ ein ggT von $a_1, \ldots, a_n$, also sind $d, \tilde{d}$ assoziiert. □

Korollar 1.29 (Satz von Bezout) *Sei R ein Hauptidealring und seien $a_1, \ldots, a_n \in R$. Dann gibt es den ggT von $a_1, \ldots, a_n$ und dieser hat die Form $d = r_1 a_1 + \ldots + r_n a_n$, mit $r_i \in R$.*

Korollar 1.30 *Sei R ein Hauptidealring. Dann sind $a_1, \ldots, a_n \in R$ teilerfremd genau dann, wenn es $r_1, \ldots, r_n \in R$ gibt, mit $1 = r_1 a_1 + \ldots + r_n a_n$.*

Beweis: Aus Korollar 1.29 folgt, daß 1 als ggT von teilerfremden $a_1, \ldots, a_n$ sich entsprechend darstellen läßt. Ist $1 = r_1 a_1 + \ldots + r_n a_n$ und wäre a ein gemeinsamer Teiler $a_1, \ldots, a_n$ so ist a Teiler von 1 und damit eine Einheit. □

Korollar 1.31 *Sei R ein Hauptidealring und seien $a_1, a_2 \in R$ teilerfremd.*

(i) Falls $a_1 | a_2 c$, so folgt $a_1 | c$.

(ii) Falls $a_1|c$ und $a_2|c$, so folgt $a_1a_2|c$.

Beweis: Nach Korollar 1.30 gibt es $r_1, r_2 \in R$ mit $1 = r_1a_1 + r_2a_2$, also folgt $c = r_1a_1c + r_2a_2c$.

(i) Falls $a_1|a_2c$, so teilt a_1 jeden Summanden, also auch $a_1|c$.

(ii) Falls $a_1|c$ und $a_2|c$, dann teilt a_1a_2 jeden Summanden, also auch c.

□

Satz 1.32 *Sei R ein Integritätsring mit 1 und $a \in R$, so gilt entweder $p|a$ oder p und a sind teilerfremd.*

Beweis: Falls a und p nicht teilerfremd sind, so gibt es eine Nichteinheit d, so daß $d|a$ und $d|p$. Da aber p irreduzibel ist und damit keine echten Teiler hat, so gilt $d \sim p$. Mit jedem Teiler sind aber auch die Assoziierten Teiler, also $p|a$. □

Insbesondere impliziert Satz 1.32, daß 2 Primelemente entweder assoziiert oder teilerfremd sind. Damit erhalten wir ein Resultat, das wir von $\mathbb{Z}$ her sehr gut kennen.

Satz 1.33 *Sei R ein Hauptidealring und $x \in R \setminus \{0\}$ keine Einheit. Dann läßt sich x als Produkt von endlich vielen Primelementen darstellen.*

Beweis: Betrachte die Menge $X \subset R\setminus\{0\}$ von Nichteinheiten in R, die sich nicht als Produkt von endlich vielen Primelementen darstellen lassen. Angenommen, X ist nicht leer. Da R Hauptidealring ist, so gibt es ein maximales Element z in X (Lemma 1.7). z ist kein Primelement und nach Satz 1.24 nicht irreduzibel. Also gibt es Nichteinheiten $x, y \in R$ mit $z = xy$. Da x, y echte Teiler von z sind, gilt

$$zR \subsetneqq xR \text{ und } zR \subsetneqq yR.$$

Da z maximal ist, gehören x, y nicht zu X und können daher als Produkt von endlich vielen Primelementen dargestellt werden, also ist $z = x \cdot y$ auch Produkt endlich vieler Primelemente. (Widerspruch) □

Definition 1.34 *Ein Integritätsring mit 1 heißt* ZPE–Ring, *falls die folgenden Bedingungen gelten:*

Sei $a \in R \setminus \{0\}$ Nichteinheit, so ist a endliches Produkt irreduzibler Elemente, d.h. $a = p_1 \dots p_s$ mit p_i irreduzibel und diese Zerlegung ist bis auf die Reihenfolge und Multiplikation mit Einheiten eindeutig, d.h., aus $a = \tilde{p}_1 \dots \tilde{p}_r$ folgt $r = s$ und $p_i \sim \tilde{p}_{\pi(i)}$ für eine Permutation π.

(ZPE: Zerlegung in Primelemente ist eindeutig.)

Satz 1.35 *Sei R ein Integritätsring mit 1. Dann sind folgende Aussagen äquivalent:*

(a) R ist ZPE–Ring.

(b) Jede Nichteinheit in $R \setminus \{0\}$ ist endliches Produkt von irreduziblen Elementen und jedes irreduzible Element ist Primelement.

(c) Jede Nichteinheit in $R \setminus \{0\}$ ist endliches Produkt von Primelementen.

Beweis:
(a) $\Longrightarrow$ (b)
Nach Definition 1.34 müssen wir nur zeigen, daß jedes irreduzible Element eines ZPE-Ringes Primelement ist.
Sei $a \in R$ irreduzibel und a ein Teiler des Produkts xy, d.h. $xy = ra$ mit $r \in R$. Ist x (oder y) eine Einheit, so teilt a den anderen Faktor y (oder x). Sind x, y Nichteinheiten, so ist auch r eine Nichteinheit (denn sonst wäre a nicht irreduzibel). Damit können wir x, y, r als Produkt von irreduziblen Elementen darstellen:

$$x = x_1 \dots x_t,\ y = y_1 \dots y_s,\ r = r_1 \dots r_u \quad \text{mit } x_i, y_j, r_k \text{ irreduzibel. Es folgt}$$
$$x_1 \dots x_t y_1 \dots y_s = a\, r_1 \dots r_u.$$

Auf beiden Seiten steht ein Produkt irreduzibler Elemente und da die Darstellung nach Defintion 1.34 im wesentlichen eindeutig ist, folgt $a = \varepsilon x_i$ oder $a = \tilde{\varepsilon} y_j$ mit Einheiten $\varepsilon, \tilde{\varepsilon}$.
$\Longrightarrow \quad a|x$ oder $a|y \quad \Longrightarrow \quad a$ ist Primelement.

(b) $\Longrightarrow$ (c)
ist trivial.

(c) $\Longrightarrow$ (b)
Da jedes Primelement irreduzibel ist, existiert für jedes $x \in R \setminus \{0\}$ eine endliche Zerlegung in irreduzible Elemente. Ist $q \in R \setminus \{0\}$ irreduzibel und ein Produkt $q = p_1 \dots p_k$ von Primelementen, so folgt $k = 1 \Longrightarrow q$ ist Primelement.

(b) $\Longrightarrow$ (a)
Es reicht, die wesentliche Eindeutigkeit der Zerlegung in irreduzible Faktoren zu zeigen.
Seien $p_1 \dots p_k = q_1 \dots q_l$ mit p_i, q_j irreduzibel und sei o.B.d.A. $k \leq l$. Nach Voraussetzung sind alle p_i, q_j auch Primelemente.
Da $p_1 \mid q_1 \dots q_l$, so teilt p_1 mindestens eines der q_j, o.B.d.A. $p_1|q_1$. Da nach Korollar 1.32 je zwei Primelemente assoziiert oder teilerfremd sind, folgt $q_1 = \varepsilon_1 p_1$ mit invertierbarer Einheit ε_1. Kürzen von p_1 auf beiden Seiten und Wiederholung ergibt

$$q_i = \varepsilon_i p_{\pi(i)} \quad \text{und} \quad 1 = \varepsilon_1 \dots \varepsilon_k\, q_{k+1} \dots q_l$$

$\Longrightarrow \quad k = l$, denn die q_i sind keine Einheiten. □

Wir schließen damit sofort, daß jeder Hauptidealring ein ZPE–Ring ist.

Definition 1.36 *Ein Integritätsring R heißt* euklidischer Ring, *wenn es eine Abbildung*

$$\delta : R \setminus \{0\} \to \mathbb{N}_0$$

gibt, so daß zu $a, b \in R, b \neq 0$ es Elemente $q, r \in R$ gibt mit $a = qb + r$ (Division mit Rest, wobei entweder $r = 0$ oder $\delta(r) < \delta(b)$.) Diese Abbildung δ heißt Gradfunktion *auf R.*

Beispiel 1.37 *Für $k \in \mathbb{N}$ ist $\mathbb{Z}$ mit $\delta(m) = |m|^k$ ein euklidischer Ring.*

Satz 1.38 *Jeder euklidische Ring ist ein Hauptidealring und damit auch ZPE–Ring*

Beweis: Sei I ein Ideal des euklidischen Rings R und sei $I \neq 0 \cdot R$. Dann ist die Menge $\{\delta(i) \mid i \in I, i \neq 0\}$ nicht leer und hat als Teilmenge von $\mathbb{N}_0$ ein kleinstes Element $\delta(j)$. Sei $v \in I$ mit $v \neq j$. Nach Definition 1.36 gibt es $q, r \in R$ mit

$$v = qj + r \text{ wobei } r = 0 \text{ oder } \delta(r) < \delta(j).$$

Mit v und j liegt auch $r = v - qj$ in I. Also würde $\delta(r) < \delta(j)$ einen Widerspruch zur Minimalität von j bilden,

$$\Longrightarrow\ r = 0 \ \Longrightarrow\ v = qj \ \Longrightarrow\ I = Rj.$$

Damit ist jedes Ideal I in R Hauptideal. Wir müssen also nur noch zeigen, daß R eine 1 hat.
Dazu betrachten wir $I = R$. Da $R = Rc$ mit $c \in R$, so gibt es $e \in R$ mit $c = ec$ und zu jedem $x \in R$ gibt es $y \in R$ mit $x = yc$.

$$\Longrightarrow\ xe = (yc)e = y(ce) = yc = x \ \Longrightarrow\ e = 1.$$

$\Longrightarrow$ R ist Hauptidealring. Wir erhalten damit natürlich sofort (nach Satz 1.33), daß jeder euklidische Ring ein ZPE–Ring ist. □

Beispiel 1.39 In euklidischen Ringen können wir den euklidischen Algorithmus zur Bestimmung des ggT verwenden. Seien $a_1, a_2 \in R \setminus \{0\}$ und δ die Gradfunktion in R.

$$\begin{aligned}
\delta(a_2) &\leq \delta(a_1) \\
a_1 &= q_1 a_2 + a_3\,, \quad \text{wobei } \begin{cases} a_3 = 0 \text{ oder} \\ \delta(a_3) < \delta(a_2), \end{cases} \\
&\qquad\qquad \text{falls } a_3 \neq 0 \\
a_2 &= q_2 a_3 + a_4\,, \quad \text{wobei } \begin{cases} a_4 = 0 \text{ oder} \\ \delta(a_4) < \delta(a_3) \end{cases} \\
&\ \ \vdots
\end{aligned}$$

Da die Folge der $\delta(a_i)$ monoton fallend ist, folgt irgendwann $a_n = q_n a_{n+1} + a_{n+2}$ mit $a_{n+2} = 0$, aber $a_{n+1} \neq 0$.
Dann ist $d = a_{n+1}$ der ggT von a_1, a_2.

Wir können diesen Algorithmus auch in Matrixschreibweise schreiben:

$$\begin{bmatrix} q_i & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{i+1} \\ a_{i+2} \end{bmatrix} = \begin{bmatrix} a_i \\ a_{i+1} \end{bmatrix}, \quad i \geq 1.$$

Dies ist eine Differenzengleichung und es folgt, daß

$$\begin{aligned}
\begin{bmatrix} a_{i+1} \\ a_{i+2} \end{bmatrix} &= \begin{bmatrix} q_i & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} a_i \\ a_{i+1} \end{bmatrix} \\
&= \begin{bmatrix} q_i & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} q_{i-1} & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} a_{i-1} \\ a_i \end{bmatrix} \\
&\vdots \\
&= \begin{bmatrix} q_i & 1 \\ 1 & 0 \end{bmatrix}^{-1} \cdots \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 \\ 1 & -q_i \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -q_{i-1} \end{bmatrix} \cdots \begin{bmatrix} 0 & 1 \\ 1 & -q_1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}
\end{aligned}$$

und da $a_{n+2} = 0$ erhalten wir die Darstellung von $\begin{bmatrix} a_{n+1} \\ 0 \end{bmatrix}$ mittels $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ und den q_i.

Um den ggT von n Zahlen $a_1, \ldots, a_n$ zu bestimmen, geht man sukzessive vor und bestimmt den ggT d_1 von a_1, a_2, dann den ggT d_2 von d_1 und a_3 usw.

Korollar 1.40 *In $\mathbb{Z}$ gibt es unendlich viele Primzahlen.*

Beweis: Es seien $p_1, \ldots, p_n$ Primzahlen in $\mathbb{Z}$. Da $\mathbb{Z}$ ein ZPE–Ring ist, gibt es eine Primzahl p, die $(p_1 p_2 \ldots p_n) + 1$ teilt, und p ist verschieden von $p_1, \ldots, p_n$, da sonst p Teiler von 1 wäre. Also folgt mit Induktion, daß es zu je endlich vielen Primzahlen eine weitere gibt. □

Beispiel 1.41 Sei $n \in \mathbb{Z} \setminus \{0, 1\}$ eine Zahl, in deren Faktorisierung kein Quadrat einer Primzahl vorkommt.
Setze $\mathbb{Q}(\sqrt{n}) = \{x + y\sqrt{n} \mid x, y \in \mathbb{Q}\}$ und für $z = x + y\sqrt{n} \in \mathbb{Q}(\sqrt{n})$ setze $\bar{z} = x - y\sqrt{n}$, so gilt für $z \neq 0$, daß

$$z^{-1} = \frac{\bar{z}}{z\bar{z}}$$

und damit ist $\mathbb{Q}(\sqrt{n})$ ein Körper (die restlichen Axiome folgen trivialerweise). Setze $\mathbb{Z}(\sqrt{n}) = \{x + y\sqrt{n} \mid x, y \in \mathbb{Z}\}$. So ist $\mathbb{Z}(\sqrt{n})$ ein Unterring von $\mathbb{Q}(\sqrt{n})$ und zwar ein Integritätsring mit 1.

Satz 1.42 *Für $n = -1, -2, 2, 3$ ist $\mathbb{Z}(\sqrt{n})$ mit $\delta : z \to |z\bar{z}|$ für $z \neq 0$ ein euklidischer Ring.*

Beweis: $\delta : \mathbb{Z}(\sqrt{n}) \setminus \{0\} \to \mathbb{N}$. Seien $u = u_1 + u_2\sqrt{n}$, $v = v_1 + v_2\sqrt{n} \in \mathbb{Z}(\sqrt{n}) \setminus \{0\}$. Da v in $\mathbb{Q}(\sqrt{n})$ invertierbar ist, sei $uv^{-1} = w_1 + w_2\sqrt{n}$ mit $w_1, w_2 \in \mathbb{Q}$. Seien $x_1, x_2 \in \mathbb{Z}$ die Zahlen, die am nächsten an w_1, w_2 liegen, d.h.
$|w_1 - x_1| \leq \frac{1}{2}$ und $|w_2 - x_2| \leq \frac{1}{2}$. Setze $q = x_1 + x_2\sqrt{n}$ und $r = v(uv^{-1} - q)$. Dann ist

$$u = qv + r \quad \text{mit } r = 0$$

oder

$$\begin{aligned}\delta(r) &= \delta\left(v(uv^{-1}-q)\right)\\ &= |v\bar{v}|\left|(uv^{-1}-q)(\overline{uv^{-1}-q})\right|\\ &= \delta(v)\cdot\left|(w_1-x_1)^2-(w_2-x_2)^2n\right|\end{aligned}$$

für $n=-1,-2,2,3$ folgt, da

$$|w_1-x_1|\leq\frac{1}{2},\quad |w_2-x_2|\leq\frac{1}{2},$$

daß $|(w_1-x_1)^2-(w_2-x_2)^2n|<1$, also $\delta(r)<\delta(v)$. □

Beispiel 1.43 $\mathbb{Z}(i)=\{x+yi\in\mathbb{C}\mid x,y\in\mathbb{Z}\}$ ist ein euklidischer Ring.

Kapitel 2

Polynomringe

In diesem Kapitel wollen wir die Konzepte aus Kapitel 1 auf Polynomringe anwenden.

Sei R ein kommutativer Ring mit $1 \neq 0$. Eine Reihe $a(z) = \sum_{i=0}^{\infty} a_i z^i$ mit $a_i \in R$ heißt *formale Potenzreihe*. Die Menge der formalen Potenzreihen mit Koeffizienten in R und unbekannter z bezeichnen wir mit $R[[z]]$.
Seien $f, g \in R[[z]]$, so definiere

$$(f+g)(z) = f(z)+g(z) = \sum_{i=0}^{\infty}(f_i+g_i)z^i \tag{2.1}$$

$$(f \cdot g)(z) = \sum_{i=0}^{\infty} l_i z^i, \quad \text{mit } l_i = \sum_{s+t=i} f_s \cdot g_t \tag{2.2}$$

$$e(z) = 1 + \sum_{i=1}^{\infty} 0 \cdot z^i \tag{2.3}$$

Satz 2.4 *R sei kommutativer Ring mit 1. Dann ist $R[[z]]$ mit Addition (2.1), Multiplikation (2.2) und Einselement (2.3) ein kommutativer Ring mit 1.*

Beweis: Übung □

Satz 2.5 *Die Menge $R[z] = \{a \in R[[z]]$ mit $a = \sum_{i=0}^{\infty} a_i z^i$ und $a_i = 0$ für fast alle $i \in \mathbb{N}_0\}$ ist Unterring von $R[[z]]$.*

Beweis: $a, b \in R[z] \implies a = \sum_{i=0}^{r} a_i z^i, \; b = \sum_{j=0}^{s} b_j z^j \implies$

$$a+b = \sum_{i=0}^{\max(s,r)} (a_i+b_i)z^i,$$

$$a \cdot b = \left(\sum_{i=0}^{r} a_i z^i\right)\left(\sum_{j=0}^{s} b_j z^j\right) = \sum_{k=0}^{r+s}\left(\sum_{i+j=k}(a_i b_j)\right) z^k$$

haben nur endlich viele von 0 verschiedene Koeffizienten. Die anderen Axiome sind klar. □

Definition 2.6 *$R[z]$ heißt* Polynomring über R in z.
Die Abbildung

$$\begin{aligned} &\textit{Grad: } R[z]\backslash\{0\} \to \mathbb{N}_0 \\ &\textit{Grad }(a) = \max\{i \in \mathbb{N}_0 \mid a_i \neq 0\} \end{aligned}$$

heißt Grad von a.
a_n mit $n = Grad(a)$ heißt führender Koeffizient, *und a heißt* normiert *falls $a_n = 1$.*
Das Nullpolynom hat keinen Grad. Das Einselement ist

$$1 = 1 + 0z^1 + \ldots + 0z^n.$$

Beispiel 2.7 Sei $R = \mathbb{F}_5$, so ist $1z^5 + 2z^2 + 3 \in \mathbb{F}_5[z]$ normiertes Polynom und mittels Äquivalenz ist auch $6z^5 + 3z^2 + 1 \in \mathbb{F}_5[z]$ normiert, da in $\mathbb{F}_5$ gilt $6 = 1$.

Satz 2.8 *Seien $a, b \in R[z]\backslash\{0\}$, so gilt*

(a) $a \cdot b = 0$ oder

$$Grad(a \cdot b) \leq Grad(a) + Grad(b).$$

(b) Falls R ein Integritätsring ist, so gilt

$$Grad(a \cdot b) = Grad(a) + Grad(b).$$

(c) $R[z]$ ist Integritätsring genau dann, wenn R ein Integritätsring ist.

(d) Ist R ein Integritätsring mit 1, so ist $a \in R[z]$ in $R[z]$ invertierbar genau dann, wenn a eine invertierbare Konstante ist.

Beweis:

(a), (b) trivial.

(c) Wir müssen nur zeigen, daß $R[z]$ nullteilerfrei ist und ein Element $\neq 0$ enthält. Seien $a, b \in R[z]\backslash\{0\}$ mit führenden Koeffizienten a_s, b_t (gibt es, da R Elemente $\neq 0$ enthält), so folgt, daß $a_s \cdot b_t$ der führende Koeffizient von $a \cdot b$ ist und da R nullteilerfrei, so ist $a_s \cdot b_t \neq 0$. Damit ist auch $R[z]$ nullteilerfrei, denn für $a \cdot b = 0$ folgt

$$\sum_{i+j=m} a_i \cdot b_j = 0 \ \ \forall m = 0, \ldots, s+t \implies a_s \cdot b_t = 0 \implies a_s = 0 \ \text{ oder } \ b_t = 0.$$

Angenommen, $b_t \neq 0, a_s = 0$. Dann folgt

$$\begin{aligned} 0 = \sum_{i+j=t+s-1} a_i b_j &= a_s b_{t-1} + a_{s-1} b_t = a_{s-1} b_t \\ &\implies a_{s-1} = 0 \end{aligned}$$

$$\begin{aligned} 0 = \sum_{i+j=t+s-2} a_i b_j &= a_s b_{t-2} + a_{s-1} b_{t-1} + a_{s-2} b_t = a_{s-2} b_t \\ &\implies a_{s-2} = 0 \\ &\vdots \end{aligned}$$

$\implies a = 0.$

(d) Sei $a \in R[z]$ invertierbar und $a \cdot b = 1$ mit $b \in R[z]$.
$\Longrightarrow$ Grad $1 = 0 =$ Grad $(a \cdot b) =$ Grad $(a)+$ Grad (b), also folgt
Grad $(a) =$ Grad $(b) = 0 \implies a, b$ sind konstante Polynome.

□

Beispiel 2.9 Es muß nicht gelten, daß Grad $(a \cdot b) =$ Grad $(a)+$ Grad (b), denn in $\mathbb{Z}_8[z]$ gilt

$$\begin{aligned}(1+2z)(4+z+4z^2) &= 4+z+4z^2+8z+2z^2+8z^3 \\ &= 4+9z+6z^2+8z^3 \\ &\quad \| \qquad\quad \| \\ &\quad 1 \qquad\quad 0\end{aligned}$$

Beispiel 2.10 Da für einen Integritätsring R auch $R[z]$ ein Integritätsring ist, können wir den Quotientenkörper $R(z) = Q(R[z]) = \{a/b \mid a, b \in R[z],\ b \neq 0\}$ bilden.
$R(z)$ heißt *Körper der rationalen Funktionen* über R in z.

Definition 2.11 *Sei S ein (nicht notwendigerweise kommutativer) Ring und R ein kommutativer Unterring von S, so heißt die Abbildung, die für $s \in S$ durch*

$$\begin{aligned}\Phi_s \ &: \ R[z] \to S \\ &\quad\ a \mapsto a(s)\end{aligned}$$

gegeben ist, Einsetzhomomorphismus.

Daß dieser Name gerechtfertigt ist, folgt aus $(a+b)(s) = a(s)+b(s)$ und $(a \cdot b)(s) = a(s) \cdot b(s)$.

Beispiel 2.12 Sei K ein Körper, so ist für $x \in K$ $x \to xI_n$ eine Abbildung von $K \to K^{n,n}$ und für $a \in K[z]$ und $S \in K^{n,n}$ ist

$$a(S) = \sum_{i=0}^{n} a_i S^i$$

ein Homomorphismus von $K[z] \to K^{n,n}$.

Satz 2.13 *Sei R ein Integritätsring mit 1, so ist der für $g \in R[z]$ definierte Einsetzhomomorphismus*

$$\begin{aligned}\Phi_g \ &: \ R[z] \to R[z] \\ &\quad\ a \mapsto a(g)\end{aligned}$$

ein Isomorphismus genau dann, wenn $g = g_2 z + g_1$ mit $g_1 \in R$ und g_2 invertierbar in R.

Beweis: Sei $g = g_2 z + g_1$ mit g_2 invertierbar in R und sei $\tilde{g} = g_2^{-1}(z - g_1)$, so gilt

$$\Phi_g \Phi_{\tilde{g}} = \Phi_{\tilde{g}} \Phi_g = Id, \quad \text{denn}$$

$$\begin{aligned} \Phi_g \Phi_{\tilde{g}}(a) &= \Phi_g g_2^{-1}(a - g_1) \\ &= g_2 \left(g_2^{-1}(a - g_1) \right) + g_1 \\ &= a - g_1 + g_1 = a. \end{aligned}$$

Also ist Φ_g bijektiv und damit ein Isomorphismus.
Wenn wir die Grade vergleichen, so folgt für $a, g \neq 0$ $\operatorname{grad} a(g) = \operatorname{grad} a \cdot \operatorname{grad} g$. Ist $\Phi_g(a)$ surjektiv, so ist das Polynom z im Bild von Φ_g.
Da $a(g) = z$, ergibt sich $1 = \operatorname{grad}(a)\operatorname{grad}(g) \Longrightarrow \operatorname{grad}(a) = \operatorname{grad}(g) = 1$.

$$\begin{aligned} \Longrightarrow \quad & g = g_2 z + g_1 \\ & a = a_2 z + a_1, \quad \text{mit } g_2, a_2 \neq 0 \\ & z = a(g) = a_2(g_2 z + g_1) + a_1 \\ \Longrightarrow \quad & a_2 g_2 = 1 \quad \Longrightarrow \quad g_2 \text{ invertierbar .} \end{aligned}$$

□

Wir können nun in Polynomringen Division mit Rest machen.

Satz 2.14 *Sei R ein kommutativer Ring mit $1 \neq 0$, $a, b \in R[z]$, $a = \sum_{i=0}^{n} a_i z^i$ mit a_n invertierbar. Dann gibt es eindeutig bestimmte Polynome $q, r \in R[z]$, so daß $b = qa + r$ mit $r = 0$ oder* $\operatorname{grad}(r) < \operatorname{grad}(a)$

Beweis: Mit Induktion nach $\operatorname{grad}(b)$.
O.B.d.A. sei $b \neq 0$, ansonsten sind wir fertig mit $q = 0, r = 0$.

I.A. Sei nun $\operatorname{grad}(b) = 0 \implies b \in R$.
Ist auch $\operatorname{grad}(a) = 0$ also $a = a_0 \in R$ invertierbar, so folgt die Behauptung mit $q = b a_0^{-1}$ und $r = 0$. Ist $\operatorname{grad}(a) > 0$, so folgt die Behauptung mit $q = 0$ und $r = b$.

I.V. Sei Behauptung richtig für $\operatorname{grad}(b) = m - 1 \geq 0$.

I. S. Sei $\operatorname{grad}(b) = m$, $\operatorname{grad}(a) = n$ und $m \geq n \geq 1, b = \sum_{i=0}^{m} b_i z^i$.
Setze $B_1 = b - b_m a_n^{-1} z^{m-n} a$, so ist der m–te Koeffizient von B_1 gerade

$$b_m - b_m a_n^{-1} a_n = 0 \quad \Longrightarrow \quad B_1 = 0 \quad \text{oder} \quad \operatorname{grad}(B_1) < \operatorname{grad}(b).$$

In beiden Fällen folgt aus I.A. bzw. I.V., daß es $q_1, r_1 \in R[z]$ gibt mit $B_1 = q_1 a + r_1$, wobei $r_1 = 0$ oder $\operatorname{grad}(r_1) < \operatorname{grad}(a) \Longrightarrow b = (b_m a_n^{-1} z^{m-n} + q_1)a + r_1$, wie gewünscht.
Mit $b = qa + r = \tilde{q}a + \tilde{r}$ folgt $(q - \tilde{q})a = \tilde{r} - r$.
Wäre $r \neq \tilde{r}$, so folgt $(q - \tilde{q})a \neq 0$ und da $\operatorname{grad}(\tilde{q} - q) + \operatorname{grad} a = \operatorname{grad}(\tilde{r} - r)$ ergibt sich ein Widerspruch, da $\operatorname{grad}(\tilde{r} - r) \leq \operatorname{grad} r < \operatorname{grad} a$. Also $r = \tilde{r}$ und da a kein Nullteiler ist, auch $\tilde{q} = q$.

□

Beispiel 2.15 Seien $a(x) = 6x^3 + 3x^2 + 2x + 1, \quad b(x) = x^2 + 2x + 1 \in \mathbb{Z}[x]$.

$$\begin{array}{rrrrcl}
(6x^3 & +3x^2 & +2x & +1) & : & (x^2+2x+1) = 6x-9 \\
-(6x^3 & +12x^2 & +6x &) & & \\
\hline
 & -9x^2 & -4x & +1 & = & B_1(x) \\
- & (-9x^2 & -18x & -9) & & \\
\hline
 & & 14x & +10 & = & B_2(x)
\end{array}$$

$$a(x) : b(x) = 6x - 9 + \frac{14x + 10}{x^2 + 2x + 1}$$

Satz 2.16 *Der Polynomring $K[z]$ über einem Körper K ist ein euklidischer Ring.*

Beweis: Da K Körper ist, so sind alle führenden Koeffizienten invertierbar, damit kann Satz 2.14 angewendet werden für alle $a, b \in K[x]$ mit $b \neq 0$ und die Funktion grad ist wie gewünscht in Def. 1.36 und mit Satz 2.8 (c) folgt, daß $K[z]$ ein Integritätsring mit 1 und damit ein euklidischer Ring ist. □

Korollar 2.17 *Sei K ein Körper.*

(a) $K[z]$ ist ein Hauptidealring.

(b) Ist $a \in K[z]$ nicht konstantes Polynom, so läßt sich a eindeutig (bis auf Reihenfolge) als Produkt $a = a_0 p_1 p_2 \dots p_r$ schreiben, wobei $p_i \in K[z]$ irreduzible Polynome in $K[z]$ und $K \ni a_0 \neq 0$.

(c) Ist $p \in K[z]$ irreduzibel, dann ist der Restklassenring $K[z]/\{p(z)\}$ ein Körper.

Beispiel 2.18 *Sei $K = \mathbb{F}_2, a = 1x^2 + 0x + 1 = 1 \cdot (1x+1)(1x+1)$.*
$P = 1x+1$ ist irreduzibel, denn wenn $P = a(x) \cdot b(x)$, so folgt, daß die führenden Koeffizienten alle invertierbar sind und damit ist mindestens eines der beiden $a(x)$ oder $b(x)$ ein konstantes invertierbares Polynom und damit a und b invertierbar.

Satz 2.19 *Sei R ein kommutativer Ring mit 1. Dann sind folgende Aussagen äquivalent:*

(i) R ist Körper.

(ii) $R[z]$ ist euklidischer Ring.

(iii) $R[z]$ ist Hauptidealring.

Beweis: (i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii) haben wir bereits bewiesen. Für (iii) $\Longrightarrow$ (i) betrachten wir den Einsetzhomomorphismus

$$\begin{array}{rcl}
\Phi & : & R[z] \to R, \\
 & & a(z) \to a(0) = a_0.
\end{array}$$

Dann ist $R[z]/\text{Kern}\,\Phi$ isomorph zu R (siehe Übung).

Da $R[z]$ Hauptidealring, ist $R[z]$ per Definition auch ein Integritätsring, so ist R als Unterring von $R[z]$ auch Integritätsring. Damit ist wegen der Isomorphie $R[z]/\text{Kern}\,\Phi$ ein Integritätsring. Nach Satz 1.17 ist Kern Φ Primideal, welches nach Satz 1.24 in $R[z]$ maximal ist. Also ist $R \cong R[z]/\text{Kern}\,\Phi$ ein Körper nach Lemma 1.8. □

Beispiel 2.20 Aus Satz 2.19 folgt sofort, daß $\mathbb{Z}[z]$ kein Hauptidealring ist, jedoch $\mathbb{F}_p[z]$.

Wir betrachten nun Einsetzhomomorphismen und Erweiterungen von Ringen.

Definition 2.21 *Sei S ein kommutativer Ring mit 1, der den kommutativen Ring R mit 1 als Unterring hat. S heißt* Erweiterung von R. *Zu $a \in R[z]$ betrachten wir dann $\tilde{a} : S \to S$, definiert durch $\tilde{a}(s) = a(s)$ für alle $s \in S$.*
$\tilde{a}$ heißt die durch a induzierte Polynomabbildung auf S *mit Koeffizienten aus R.*

Beispiel 2.22 $R = \mathbb{Z}, S = \mathbb{Q}$

$$a = \sum_{i=0}^{n} a_i z^i \in R[z], \qquad \begin{aligned} \tilde{a} : \quad & S \to S \\ & s \mapsto \sum_{i=0}^{n} a_i s^i. \end{aligned}$$

Man beachte, die klassischen Polynome, wie wir sie in der Schule kennengelernt haben, sind eigentlich die Polynomabbildungen.

Beispiel 2.23 Die Polynomabbildung und die Elemente aus $R[z]$ sind wesentlich unterschiedliche Objekte.
Sei z.B. $R = \mathbb{F}_2[z]$ und $a = z + z^2 \in \mathbb{F}_2[z]$. Da $a(0) = 0$ und $a(1) = 1 + 1^2 = 1 + 1 = 0$, so folgt, daß die durch a induzierte Abbildung

$$\tilde{a} : \mathbb{F}_2 \to \mathbb{F}_2$$

die Nullabbildung ist, obwohl a nicht das Nullpolynom ist.
Dies kann aber nur bei Ringen passieren, bei denen für $n \in \mathbb{N}, n \cdot r = 0$ gelten kann, ohne daß $r = 0$ oder $n = 0$ ist.

Definition 2.24 *Sei R ein Ring, so heißt $q = \min\{n \in \mathbb{N} \mid \exists a \in R, a \neq 0, na = 0\}$ die* Charakteristik *von R, und wird als* Char $(R) = q$ *bezeichnet. Falls dieses Minimum nicht existiert, so sagen wir* Char $(R) = 0$.

Beispiel 2.25 $\mathbb{F}_p$ hat Charakteristik p; $\mathbb{R}, \mathbb{Q}, \mathbb{Z}$ haben Charakteristik 0.
In $\mathbb{F}_p$ wissen wir, daß $\underbrace{1 + 1 + \ldots + 1}_{p\text{-mal}} = p\,1 = p = 0$ und für $0 < r < p$, $ra \neq 0\ \forall a \in \mathbb{F}_p$.
In $\mathbb{R}, \mathbb{Q}, \mathbb{Z}$ folgt aus $na = 0$ sofort, daß $a = 0$ oder $n = 0$.

Definition 2.26 *Sei S kommutativer Ring und Erweiterung des kommutativen Rings R. Sei $a \in R[z]$. $s \in S$ heißt* Nullstelle von a *wenn $a(s) = 0$.*

Beispiel 2.27 $\sqrt{-1} = i \in \mathbb{C}$ ist Nullstelle von $z^2 + 1 \in \mathbb{Z}[z]$.

Satz 2.28 *Sei R ein Integritätsring mit 1 und $a \in R[z]$ habe* grad $(a) = n$. *Dann hat a höchstens n Nullstellen in R und für jede Nullstelle $\lambda \in R$ ist $z - \lambda$ Teiler von a.*

Beweis: Sei $\lambda \in R$ und $a(\lambda) = 0$, dann gibt es nach Satz 2.14 zu a und $z - \lambda$ Polynome q und r mit $a(z) = q(z)(z - \lambda) + r$, wobei entweder $r = 0$ oder grad $(r) <$ grad $(z - \lambda) = 1$, d.h., $r \in R$. Da $0 = a(\lambda) = q(\lambda)(\lambda - \lambda) + r$, so folgt $r = 0$ und $a(z) = q(z)(z - \lambda)$.

Ist $\tilde{\lambda} \neq \lambda$ eine weitere Nullstelle, so folgt aus $0 = a(\tilde{\lambda}) = q(\tilde{\lambda})(\tilde{\lambda} - \lambda)$, daß $\tilde{\lambda}$ Nullstelle von q ist, da es in R keine Nullteiler gibt.
Wir wenden dann Induktion an und zeigen mittels Gradvergleich, daß es höchstens n Nullstellen geben kann. □

Satz 2.29 *Sei R Integritätsring mit 1. R habe unendlich viele Elemente, dann ist $R[z]$ isomorph zum Ring aller Polynomabbildungen von R.*

Beweis: Es ist klar, daß die Menge aller Polynomabbildungen von R als Unterring aller Abbildungen von $R \to R$ ein Ring ist. Sei $a \in R[z]$ und $\tilde{a}$ die induzierte Polynomabbildung. Dann ist die Abbildung φ, die jedem a die induzierte Abbildung $\tilde{a}$ zuweist, ein injektiver Homomorphismus (auch Epimorphismus genannt).
Sei $a \in$ Kern φ, d.h. $\tilde{a} = 0$ bzw. $\tilde{a}(s) = 0 \ \forall s \in R$. Ist $a \neq 0$, so hat a nach Satz 2.28 höchstens grad (a) viele Nullstellen, also R höchstens grad (a) viele Elemente, dies ist ein Widerspruch, also folgt $a = 0$. □

Damit können wir einen wichtigen Satz beweisen, der in der Numerischen Mathematik eine große Rolle spielt.

Satz 2.30 (Interpolationsformel nach Lagrange) *Sei K ein Körper mit unendlich vielen Elementen, seien $x_0, x_1, \ldots, x_n \in K$ paarweise verschieden und $y_0, y_1, \ldots, y_n \in K$ beliebig. Dann gibt es ein eindeutiges Polynom $a(z) \in K[z]$ mit* grad $(a) \leq n$, *für welches die Interpolationsbedingungen*

$$a(x_i) = y_i, \quad i = 0, \ldots, n \tag{2.31}$$

erfüllt sind. Dieses Polynom ist gegeben durch

$$a(z) = \sum_{i=0}^{n} y_i L_i(z), \tag{2.32}$$

wobei $L_i(z)$ die Lagrange-Polynome sind mit

$$L_i(z) = \prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{z - x_j}{x_i - x_j} \tag{2.33}$$

Beweis: Einsetzen der x_j ergibt sofort, daß (2.32) die Interpolationsbedingungen (2.31) erfüllt und grad $(a) = n$. Sei b weiteres Polynom mit grad $(b) \leq n$, welches die Bedingungen (2.31) erfüllt, so folgt, daß das Polynom $b - a$ immer grad $(b - a) \leq n$ hat, aber $n + 1$ Nullstellen $x_0, \ldots, x_n$. Nach Satz 2.28 folgt damit, daß $b - a$ das Nullpolynom ist. □

Damit hat man die Möglichkeit, zu $n + 1$ Punkten (x_i, y_i) ein Polynom n–ten Grades anzugeben, welches diese Punkte interpoliert.

Beispiel 2.34 Betrachte die Wertetabelle

x_i	y_i
1	0
2	1
3	2

in $\mathbb{R}$. So interpoliert das Polynom $a(z) = \sum_{i=0}^{2} y_i L_i(z)$ mit

$$L_0(z) = \frac{(z-2)(z-3)}{(1-2)(1-3)} = \frac{1}{2}(z-2)(z-3) = \frac{1}{2}(z^2 - 5z + 6)$$

$$L_1(z) = \frac{(z-1)(z-3)}{(2-1)(2-3)} = -(z-1)(z-3) = -(z^2 - 4z + 3)$$

$$L_2(z) = \frac{(z-1)(z-2)}{(3-1)(3-2)} = \frac{1}{2}(z-1)(z-2) = \frac{1}{2}(z^2 - 3z + 2)$$

also

$$\begin{aligned} a(z) &= 0L_0 + 1L_1 + 2L_2 \\ &= -(z^2 - 4z + 3) + (z^2 - 3z + 2) \\ &= z - 1. \end{aligned}$$

Sei nun $R = \mathbb{F}_3$ und betrachte die Tabelle

x	y
0	1
1	1
2	2.

Natürlich folgt mit

$$\begin{aligned} L_0(z) &= \frac{(z-1)(z-2)}{(0-1)(0-2)} = 2^{-1}(z^2 - 3z + 2) = 2z^2 + 1, \\ L_1(z) &= \frac{(z-0)(z-2)}{(1-0)(1-2)} = \frac{1}{(-1)}(z^2 - 2z) = 2z^2 - 1z = 2z^2 + 2z, \\ L_2(z) &= \frac{(z-0)(z-1)}{(2-0)(2-1)} = \frac{1}{2}(z^2 - z) = 2z^2 - 2z = 2z^2 + 1z, \end{aligned}$$

daß

$$\begin{aligned} a(z) &= (2z^2+1)+(2z^2+2z)+2(2z^2+z) \\ &= 2z^2+1+2z^2+2z+4z^2+2z \\ &= 2z^2+z+1 \end{aligned}$$

ein Interpolationspolynom ist.

In Satz 2.29 haben wir gezeigt, daß $R[z]$ isomorph zum Ring der Polynomabbildungen ist, falls R unendlich viele Elemente hat. Wir werden nun zeigen, daß für ZPE–Ringe R auch $R[z]$ wieder ein ZPE–Ring ist. Dazu brauchen wir einige Vorbereitungen.

Lemma 2.35 *Sei P ein Primideal eines kommutativen Ringes R, so ist $P[z]$ ein Primideal in $R[z]$.*

Beweis: Seien $a = \sum a_i z^i,\ b = \sum b_i z^i$ nicht in $P[z]$, d.h., es gibt Koeffizienten, die nicht in P liegen, und seien a_r, b_s unter diesen Koeffizienten diejenigen mit dem kleinsten Index. Betrachte $c = ab = \sum c_i z^i$, so gilt

$$c_{r+s} = (a_0 b_{s+r} + a_1 b_{s+r-1} + \ldots + a_{r-1} b_{s+1}) + a_r b_s + (a_{r+1} b_{s-1} + \ldots + a_{r+s} b_0).$$

Alle Terme in den Klammern sind in $P \implies c_{r+1} = a_r \cdot b_s$ modulo P. Da nach Satz 1.17(b) $a_r b_s \neq P$, folgt $c_{r+s} \notin P \implies a(z) \cdot b(z) \notin P[z]$. $\overset{\text{Satz 1.17(b)}}{\Longrightarrow}$ $P[z]$ ist Primideal. $\square$

Definition 2.36 *Sei R ein ZPE–Ring und $0 \neq a = \sum_{i=0}^{n} a_i z^i \in R[z]$. Dann heißt a <u>primitiv</u>, falls der ggT von $a_0, \ldots, a_n$ eine Einheit in R ist. Dieser ggT wird mit $I(a)$ bezeichnet.*

Beispiel 2.37 Über einem Körper K ist jedes Polynom $\neq 0$ primitiv und natürlich ist für ZPE–Ringe jedes Polynom mit höchstem Koeffizient 1 primitiv.

$3z^2 + 2z + z$ ist über $\mathbb{Z}$ primitiv.

Satz 2.38 *Sei R ZPE–Ring und $a, b \in R[z]$ so ist $I(a \cdot b)$ assoziert mit $I(a) \cdot I(b)$. Anders ausgedrückt heißt dies, daß das Produkt primitiver Polynome primitiv ist.*

Beweis:
Da wir immer $a = I(a)\tilde{a},\ b = I(b)\tilde{b}$ schreiben können, mit $\tilde{a}, \tilde{b}$ primitiv, so folgt

$$I(a \cdot b) = I(a) I(b) I(\tilde{a} \cdot \tilde{b}).$$

Es reicht also zu zeigen, daß $I(\tilde{a} \cdot \tilde{b}) \sim 1$.

Ist für $a \cdot b \in R[z],\ a, b \neq 0,\ a \cdot b$ nicht primitiv, d.h. $I(a \cdot b)$ nicht eine Einheit, so gibt es nach Satz 1.35 ein Primelement $p \in R$, welches $I(ab)$ teilt. Also liegen alle Koeffizienten von ab in pR und dies ist nach Satz 1.24 ein Primideal. Nach Lemma 2.35 ist dann $(pR)[z]$ Primideal in $R[z]$ und da $a \cdot b \in (pR)[z]$, folgt, daß alle Koeffizienten von a oder alle Koeffizienten von b in pR liegen $\implies p \mid I(a)$ oder $p \mid I(b)$. Dann können aber nicht beide a und b primitiv sein. $\square$

Satz 2.39 *Sei R ein ZPE–Ring, $K = Q(R)$ der Quotientenkörper von R und $a \neq 0$ aus $R[z]$. Sei $a = b \cdot c$ eine Zerlegung von a in $K[z]$. So gibt es $\beta, \gamma \in K$ und $\delta \in R$, so daß $a = \delta \cdot b_1 c_1$, mit $b_1 = \beta \cdot b$, $c_1 = \gamma \cdot c$ primitiv in $R[z]$.*

Beweis: Sei m_b Produkt der Nenner der Koeffizienten von b und m_c Produkt der Nenner der Koeffizienten von c. Dann sind $m_b \cdot b$, $m_c \cdot c \in R[z]$. Es gilt $m_b b = I(m_b b) \cdot b_1$ und $m_c c = I(m_c c) c_1$ mit b_1, c_1 primitiv in $R[z]$. Dann ist

$$r \cdot a = r \cdot b \cdot c = s \cdot b_1 \cdot c_1, \quad \text{mit } r = m_b m_c,\ s = I(m_b b) I(m_c c).$$

Mit Satz 2.38 folgt $rI(a) \sim sI(b_1 c_1) \implies r$ Teiler von s und $\dfrac{s}{r} \in R$.

Dann folgt mit $\beta = \dfrac{m_b}{I(m_b b)}$, $\gamma = \dfrac{m_c}{I(m_c c)}$, $\delta = \dfrac{s}{r}$ die Behauptung. □

Korollar 2.40 *Sei R ein ZPE–Ring und $K = Q(R)$.*

(i) Sei $a \in R[z] \setminus R$ irreduzibel, so ist a irreduzibel in $K[z]$.

(ii) Sei $a \in R[z] \setminus R$ reduzibel in $K[z]$, so ist a bereits in $R[z]$ reduzibel.

(iii) Seien $a, b \in R[z]$, $a \neq 0$, b primitiv und $b \mid a$ in $K[z]$. Dann gilt bereits $b \mid a$ in $R[z]$.

Beweis: (i) und (ii) folgen sofort aus Satz 2.39.
(iii) Sei $a = c \cdot b \overset{\text{Satz 2.39}}{\implies} a = \gamma c_1 \cdot b$, da $b = b_1$, also $a \mid b$ in $R[z]$. □

Beispiel 2.41 Ist $n \in \mathbb{N}$ und kein Quadrat, so ist $z^2 - n$ über $\mathbb{Z}$ irreduzibel, also auch über $\mathbb{Q}$, also ist $\sqrt{n} \notin \mathbb{Q}$ und damit irrational.

Satz 2.42 (Satz von Gauß)
Sei R ein ZPE–Ring, so ist auch $R[z]$ ein ZPE–Ring.

Beweis: Wir müssen die Bedingungen von Definition 1.34 nachweisen. Da R ein ZPE–Ring ist, so ist $R[z]$ ein Integritätsring mit 1. Da für $a \in R \setminus \{0\}$ $a = uv$ in $R[z]$ nur mit $u, v \in R$ möglich ist (vergleiche die Grade), so ist $a \in R$ in R irreduzibel $\iff a$ in $R[z]$ irreduzibel.

Sei $0 \neq a \in R[z]$ eine Nichteinheit. Ist a konstant, so ist a als Produkt irreduzibler Elemente darstellbar, die auch irreduzibel in $R[z]$ sind. Sei daher grad $(a) > 0$. Dann zerlegen wir a in $K[z]$ mit $K = Q(R)$ als Produkt irreduzibler Faktoren $a = p_1 \cdot \ldots \cdot p_r$.

Aus Satz 2.39 folgt $a = \delta q_1 \ldots q_r$ mit $\delta \in R$ und $q_i = \alpha_i p_i \in R[z]$ primitiv. Da mit p_i auch q_i über K irreduzibel, kann man über R höchstens noch einen konstanten Faktor abspalten, der (da q_i primitiv) nur eine Einheit sein kann $\implies q_i$ irreduzibel in $R[z]$.
Wenn wir nun δ noch in irreduzible Elemente zerlegen, so haben wir die gewünschte Zerlegung. Seien $a = \delta q_1 \ldots q_r = \tilde{\delta} \tilde{q}_1 \ldots \tilde{q}_s$ zwei solche Zerlegungen, so sind $\tilde{q}_1 \ldots \tilde{q}_s$ auch irreduzibel in $K[z]$ und $\delta, \tilde{\delta}$ Einheiten in $K[z]$. In $K[z]$ (ZPE–Ring) folgt $r = s$ und (eventuell nach

Umordnung) $q_i = \beta_i \tilde{q}_i$ und $\beta_i \in K$. Aus Korollar 2.40 (iii) folgt $q_i \sim \tilde{q}_i$ in $R[z] \Longrightarrow \delta \sim \tilde{\delta}$ und dies hat eine eindeutige Zerlegung nach Voraussetzung. □

Ähnlich wie in $\mathbb{Z}$ können wir jetzt für Polynome eine Zerlegung in irreduzible Polynome durchführen. Im allgemeinen ist es aber nicht so einfach, diese Faktorisierung zu bestimmen. Aber es gibt ein paar Kriterien.

Satz 2.43
Ist $b_1 z + b_0$ Teiler von $a(z) = \sum_{i=0}^{n} a_i z^i$, so ist b_1 Teiler von a_n und b_0 Teiler von a_0.

Beweis:
$a_n z^n + \ldots + a_0 = (b_1 z + b_0)(c_{n-1} z^{n-1} + \ldots + c_0) \quad \Longrightarrow \quad a_n = b_1 c_{n-1},\ a_0 = b_0 c_0.$ □

Satz 2.44 (Satz von Eisenstein)
Sei R ZPE–Ring, $K = Q(R)$ und $a = \sum_{i=0}^{n} a_i z^i$ nicht konstantes primitives Polynom aus $R[z]$ vom Grad n. Falls es ein Primelement $p \in R$ gibt mit $p \mid a_i$, $0 \le i \le n-1$, aber $p \nmid a_n$ und $p^2 \nmid a_0$, so ist a irreduzibel in $R[z]$ und auch in $K[z]$.

Beweis: Sei $a = b \cdot c$

$$b = \sum_{i=0}^{k} b_i z^i, \quad c = \sum_{i=0}^{l} c_i z^i, \quad b_k c_l \neq 0$$
$$\Longrightarrow \quad a_0 = b_0 c_0, \quad a_n = b_k c_l, \quad a_r = \sum_{i+j=r} b_i c_j \,.$$

Da p Primelement ist, so folgt aus $p \mid a_0 = b_0 c_0$, daß $p \mid b_0$ oder $p \mid c_0$.
O.B.d.A. $p \mid b_0 \Longrightarrow p \nmid c_0$, da $p^2 \nmid a_0$. $p \nmid a_n = b_k c_l \Longrightarrow p \nmid b_k$.
Sei m der kleinste Koeffizient mit $p \nmid b_m$, so gilt $0 < m \le k \le n$
$a_m = b_m c_0 + b_{m-1} c_1 + \ldots$.
Da $p \nmid b_m c_0$, aber $p \mid b_{m-j} c_j$, so folgt $p \nmid a_m$. Nach Voraussetzung ist dies nur für $m = n$ möglich, also folgt aus $m \le k \le n$, daß grad $(b) = k = n =$ grad (a), und damit ist $a = b \cdot c$ keine echte Zerlegung. Da $1 = I(a) \sim I(b \cdot c)$, so folgt, daß c in R invertierbar ist. Die Behauptung folgt dann aus Korollar 2.40 (i). □

Beispiel 2.45

(a) Zeige, daß $a(z) = z^5 + 4z^3 + 2z + 2 \in \mathbb{Z}[z]$ über $\mathbb{Q}$ irreduzibel ist.
Mit Satz von Eisenstein folgt, daß $a(z)$ irreduzibel, da

$$\begin{aligned} p &= 2 \mid a_0, \ldots, a_4 \\ p &= 2 \nmid a_5 \\ p^2 &= 4 \nmid a_0. \end{aligned}$$

(b) Sei $n \in \mathbb{Z}$, $n \neq 0$, $n \neq q^2$ für $q \in \mathbb{Z}$. So ist $z^k - n \in \mathbb{Z}[z]$ für $k \ge 1$ irreduzibel über $\mathbb{Q}$. Berechne Primfaktorzerlegung von n und wende den Satz von Eisenstein für alle Primfaktoren an $\Longrightarrow \sqrt[k]{n}$ ist irrational.

Seien R und S kommutative Ringe mit $1 \neq 0$ und $\Phi : R \to S$ ein Ringhomomorphismus mit $\Phi(1) = 1$. Dann induziert Φ durch

$$\tilde{\Phi}\left(\sum a_i z^i\right) = \sum \Phi\left(a_i\right) z^i$$

einen Homomorphismus $\tilde{\Phi} : R[z] \to S[z]$. Dabei gilt:

Entweder ist $\tilde{\Phi}(a) = 0$ oder grad $\tilde{\Phi}(a) \leq$ grad (a). Da

$$\begin{aligned}
\tilde{\Phi}\left(\sum a_i z^i\right) = 0 &\iff \sum \Phi(a_i) z^i = 0 \\
&\iff \Phi(a_i) = 0 \quad \text{für alle } i \\
&\iff a_i \in \text{Kern } \Phi \quad \text{für alle } i. \\
&\implies \text{Kern } \tilde{\Phi} = (\text{Kern } \Phi)[z].
\end{aligned}$$

Wenn wir als Φ den kanonischen Homomorphismus $\Pi : R \to R/I$ für ein Ideal I verwenden, erhalten wir $\tilde{\Pi} : R[z] \to (R/I)[z]$ und den Kern $(\tilde{\Pi}) = I[z]$.

Satz 2.46 *Seien R, S Integritätsringe mit 1 und $\Phi : R \to S$ ein Homomorphismus mit $\Phi(1) = 1$. Sei $a \in R[z]$, so daß $a \notin$ Kern $\tilde{\Phi}$ und sei* grad $\tilde{\Phi}(a) =$ grad a.
Wenn $\tilde{\Phi}(a)$ irreduzibel in $S[z]$ ist, so ist auch a in $R[z]$ nicht echt zerlegbar, d.h. aus $a = bc$ folgt $b \in R$ oder $c \in R$.

Beweis: Wäre $a = bc$ eine echte Zerlegung, d.h. grad (b), grad $(c) \geq 1$, so folgte

$$\text{grad } (b) + \text{grad } (c) = \text{grad } (a) = \text{grad } (\tilde{\Phi}(a)) = \text{grad } (\tilde{\Phi}(b)) + \text{grad } (\tilde{\Phi}(c)).$$

Da grad $(\tilde{\Phi}(b)) \leq$ grad (b) und grad $(\tilde{\Phi}(c)) \leq$ grad (c), muß in beiden Ungleichungen Gleichheit gelten. Damit ist $\tilde{\Phi}(a) = \tilde{\Phi}(b)\tilde{\Phi}(c)$ ebenfalls eine echte Zerlegung, dies aber ist ein Widerspruch zur Irreduziblität von $\tilde{\Phi}$. □

Korollar 2.47 *Sei p eine Primzahl. $\Pi : \mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ und $a = \sum a_i z^i \in \mathbb{Z}[z]$ mit* grad $(a) \geq 1$ *und p kein Teiler des führenden Koeffizienten.*
Falls $\tilde{\Pi}(a) = \sum \Pi(a_i) z^i$ irreduzibel über $\mathbb{Z}/p\mathbb{Z}$, so ist a auch irreduzibel über $\mathbb{Q}$.

Beispiel 2.48 Sei $a = z^3 + 6z^2 + 8z + 4 \in \mathbb{Z}[z]$. Sei $p = 3$ und

$$\Pi : \mathbb{Z} \to \mathbb{Z}/3\mathbb{Z}, \qquad \tilde{\Pi}(a) = z^3 - 1z + 1 = z^3 + 2z + 1.$$

In einer echten Zerlegung hätten wir mindestens einen Faktor vom Grad 1.

$$\tilde{\Pi}(a)(r) = 1 \qquad \forall r \in \mathbb{Z}/3\mathbb{Z},$$

so hat $\tilde{\Pi}(a)$ keine Nullstellen und damit keinen Faktor vom Grad 1. Also ist a irreduzibel über $\mathbb{Q}$.

Algorithmus 2.49 (Faktorisierungsalgorithmus von Kronecker)
Sei $a \in \mathbb{Z}[z]$ ein Polynom vom Grad $n > 1$. Falls a in $\mathbb{Z}[z]$ zerlegbar ist, d.h.,

$$a = b \cdot c \text{ mit } b, c \in \mathbb{Z}[z] \text{ und } \operatorname{grad}(b), \operatorname{grad}(c) \geq 1,$$

so gilt für alle $r \in \mathbb{Z}$

$$a(r) = b(r) \cdot c(r) \quad \Longrightarrow \quad b(r) \mid a(r).$$

$b(z)$ ist also unter den Polynomen p zu suchen, für die $p(r) \mid a(r)$.

Wir geben uns daher Zahlen $r_0, \ldots, r_s$ vor und bestimmen alle Teiler von $a(r_i)$, $i = 0, \ldots, s$. Dann wendet man Lagrange–Interpolation an, so daß diese Werte interpoliert werden, für alle Möglichkeiten und prüft dann, welche der Polynome a teilen.

Beispiel 2.50 Sei $a = z^4 - z^2 - 2 \ \in \mathbb{Z}[z]$.
Da $a_4 = 1$ und $a_0 = -2$, so folgt mit Satz 2.43, daß mögliche Linearfaktoren nur

$$(x-1), (x+1), (x+2), (x-2)$$

sein können. Man prüft nach, daß diese keine Faktoren sind.

Um Polynome vom Grad 2 zu untersuchen, wähle $r_0 = 0,\ r_1 = 1,\ r_2 = -1$.

$$a(0) = -2, \quad a(1) = -2, \quad a(-1) = -2.$$

Es bleiben als Faktoren $\{\pm 1, \pm 2\}$ und damit $4 \cdot 4 \cdot 4$ Möglichkeiten für Polynome p, für die $p(r_i) = \pm 1, \pm 2$. Man sieht dann sofort durch Überprüfen, daß

$$z^2 + 1 \text{ und } z^2 - 2$$

Lösungen sind.

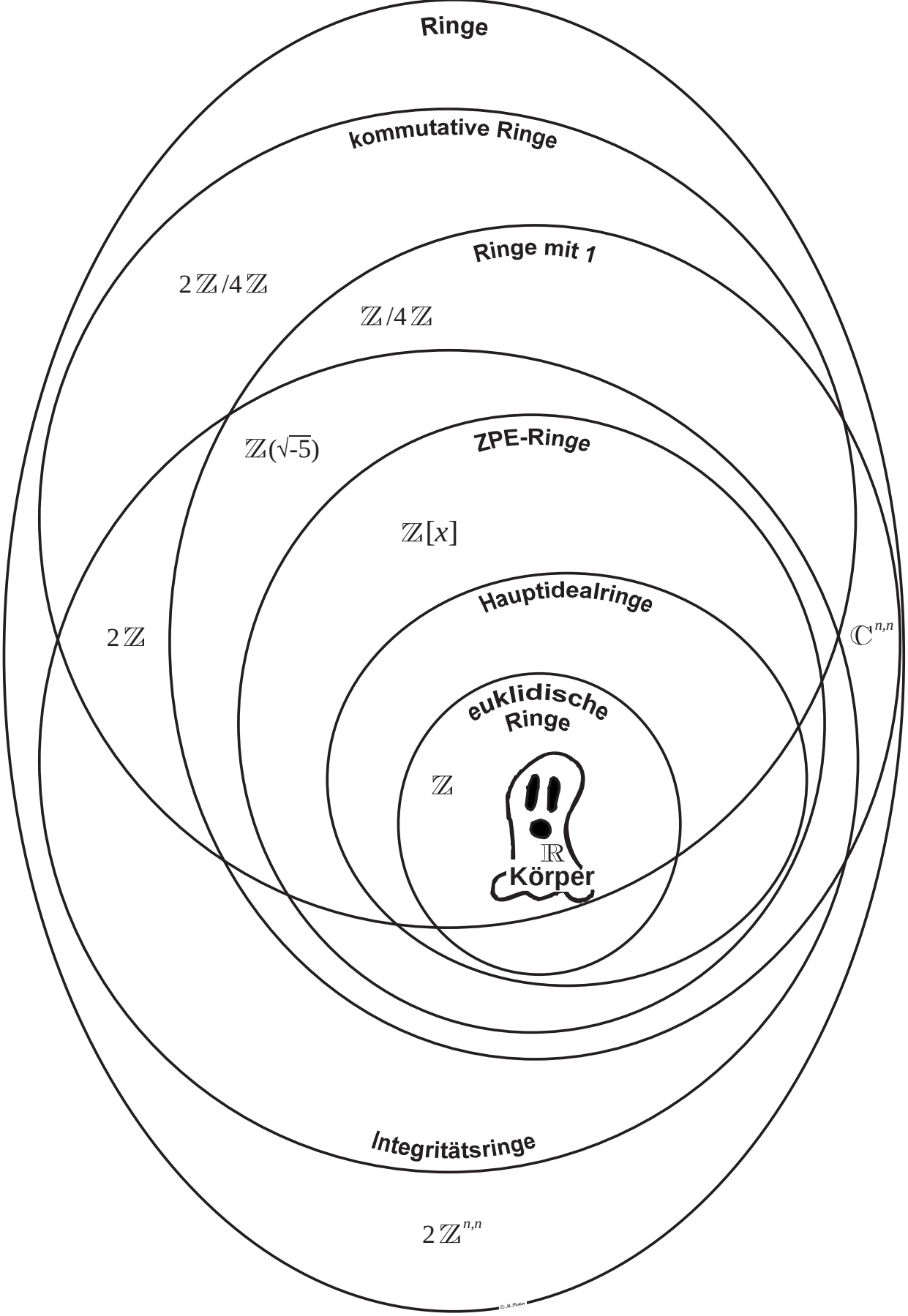
Zusammenfassung einiger algebraischer Strukturen und Eigenschaften

Ring	$(R,+,\cdot)$ mit Assoz., Distr., Null, Inv.(+)
kommut. Ring mit Eins	$(R,+,\cdot)$ Ring mit $+,\cdot$ kommutativ, Eins
Einheit	$e \in R$ mit $\exists\, r \in R:\ e \cdot r = 1$, ($e$ invertierbar)
Ideal	additive Untergruppe $I \subset R$ mit $I \cdot R = I$ $(I \cdot R := \{ar \mid a \in I, r \in R\})$
Hauptideal	$I = x \cdot R$ für ein $x \in R$
maximales Ideal	echtes Ideal I: es gibt kein Ideal J mit $I \subsetneq J \subsetneq R$
Primideal	$I \subset R$: für $a, b \in R$ und $a \cdot b \in I \ \Rightarrow\ a \in I$ oder $b \in I$
Restklassenring	$R/I = (\{[r] : [r] = \{s \in R : r - s \in I\}\}, +, \cdot)$
Integritätsring	$R \neq \{0\}, \quad a \cdot b = 0 \ \Rightarrow\ a = 0$ oder $b = 0$
Primelement	Nichteinheit p: $p \mid x \cdot y \ \Rightarrow\ p \mid x$ oder $p \mid y$
irreduzibel	Nichteinheit q: $q = x \cdot y \ \Rightarrow\ x$ oder y Einheit (invertierbar)
ZPE-Ring	$\forall\, a \in R \setminus \{0\}: \quad a = p_1 \cdot \ldots \cdot p_s$, mit p_i irreduzibel, (eindeutig bis auf Einheiten und Reihenfolge)
Hauptidealring	Integritätsring mit 1, in dem jedes Ideal Hauptideal ist
euklidischer Ring	Integritätsring mit Gradfunktion $\delta : R \setminus \{0\} \to N_0$, Division mit Rest

Beachte: In nichtkommutativen Ringen gelten manche Eigenschaften auch jeweils nur „von links“ oder „von rechts“.

einige Zusammenhänge:

- euklidischer Ring $\Rightarrow$ Hauptidealring $\Rightarrow$ ZPE-Ring
- Integritätsring mit 1: Primelement $\Rightarrow$ irreduzibel
- p Primelement in $R \Leftrightarrow pR$ ist Primideal
- q irreduzibles Element in $R \Leftrightarrow qR$ ist maximales Ideal
- Kommutativer Ring mit $1 \neq 0$: Primideal $\Rightarrow$ maximales Ideal
- I Primideal in R: R/I Integritätsring
- I maximales Ideal in R: R/I Körper


Ringe
kommutative Ringe
Ringe mit 1
$2\mathbb{Z}/4\mathbb{Z}$
$\mathbb{Z}/4\mathbb{Z}$
$\mathbb{Z}(\sqrt{-5})$
ZPE-Ringe
$\mathbb{Z}[x]$
Hauptidealringe
$2\mathbb{Z}$
$\mathbb{C}^{n,n}$
euklidische Ringe
$\mathbb{Z}$
$\mathbb{R}$
Körper
Integritätsringe
$2\mathbb{Z}^{n,n}$

Kapitel 3

Systeme von Differentialgleichungen höherer Ordnung

Im II. Teil haben wir uns mit der Lösung von gewöhnlichen Differentialgleichungen

$$\dot{x} = Ax + f\,, \quad x(0) = x^0 \tag{3.1}$$

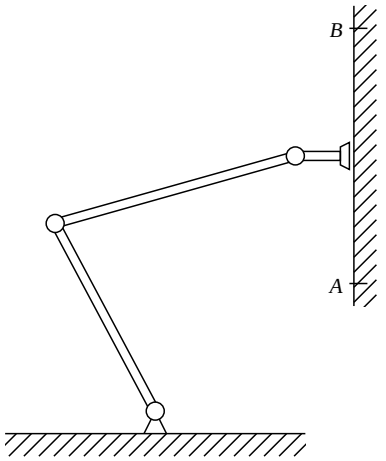
beschäftigt und gesehen, daß wir die Lösungen mit Hilfe der Jordan'schen Normalform vollständig beschreiben können.

In diesem Kapitel wollen wir diese Theorie nun auf Systeme höherer Ordnung übertragen. Diese haben die Form

$$L_l x^{(l)}(t) + L_{(l-1)} x^{l-1}(t) + \ldots + L_1 x^{(1)}(t) + L_0 x(t) = f(t) \tag{3.2}$$

wobei $L_l, \ldots, L_0 \in \mathbb{C}^{n,n}$, und im allgemeinen fordern wir auch noch, daß L_l invertierbar ist, dies ist aber nicht unbedingt nötig.

Beispiel 3.3 Mechanisches Mehrkörpersystem



Fensterputzroboter oder Schweißroboter
Modell:

$$M(\Theta)\ddot{\Theta} + C\left(\Theta, \dot{\Theta}\right) + G(\Theta) = f(u) \tag{3.4}$$

$\Theta = \begin{bmatrix} \Theta_1 \\ \Theta_2 \\ \Theta_3 \end{bmatrix}$ ist der Vektor der Verschiebung der Gelenke. $u \in \mathbb{R}^3$ ist der Steuervektor, der die Drehung der Gelenke beschreibt. $M \in \mathbb{R}^{3,3}$ ist die Massenmatrix, $C(\Theta, \dot{\Theta}) \in \mathbb{R}^3$ ist der Vektor der Zentrifugal- und Coriolis-Kräfte, $G \in \mathbb{R}^3$ ist der Gravitationsvektor, f verallgemeinerte Kräfte.

Man linearisiert dieses Modell, d.h. man nimmt an, man hat eine angenäherte Lösung $x_0(t)$ und setzt $\Theta = x_0 + x$, dann macht man Taylor-Entwicklung und erhält näherungsweise für x eine Gleichung

$$M_0\ddot{x} + C_0\dot{x} + G_0 x = f_0 \tag{3.5}$$

Man kann ein System (3.2) von Differentialgleichungen höherer Ordnung auf viele Arten lösen, z.B., indem man daraus ein System erster Ordnung macht. Man setzt

$$z_l = x^{(l-1)},\ z_{l-1} = x^{(l-2)},\ \ldots,\ z_2 = x^{(1)},\ z_1 = x \tag{3.6}$$

und erhält dann aus (3.2)

$$\left\{ \begin{array}{rcl} \dot{z}_i & = & z_{i+1}\,, \quad i = 1, \ldots, l-1\,, \\ L_l\dot{z}_l + L_{l-1}\dot{z}_{l-1} + \ldots + L_1\dot{z}_1 + L_0 z_1 & = & f \end{array} \right. \tag{3.7}$$

oder in Matrixschreibweise

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f(t) \end{bmatrix} + \begin{bmatrix} 0 & I_n & & & \\ & 0 & I_n & & \\ & & \ddots & \ddots & \\ & & & & I_n \\ -L_0 & -L_1 & \ldots & & -L_{l-1} \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \\ \vdots \\ x^{(l-2)} \\ x^{(l-1)} \end{bmatrix} = \begin{bmatrix} \dot{x} \\ \ddot{x} \\ \vdots \\ x^{(l-1)} \\ L_l x^{(l)} \end{bmatrix} = \tag{3.8}$$

$$= \begin{bmatrix} I_n & & & & \\ & I_n & & & \\ & & \ddots & & \\ & & & I_n & \\ & & & & L_l \end{bmatrix} \begin{bmatrix} \dot{x} \\ \ddot{x} \\ \vdots \\ x^{(l-1)} \\ x^{(l)} \end{bmatrix}$$

oder

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \\ \vdots \\ \dot{z}_{l-1} \\ \dot{z}_l \end{bmatrix} = \begin{bmatrix} 0 & I_n & & & & \\ 0 & 0 & I_n & & & \\ 0 & 0 & 0 & I_n & & \\ & & & & \ddots & \\ & & & & & I_n \\ -\hat{L}_0 & -\hat{L}_1 & \ldots & & & -\hat{L}_{l-1} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_{l-2} \\ z_{l-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \hat{f}(t) \end{bmatrix} \tag{3.9}$$

wobei $\hat{L}_i = L_l^{-1} L_i,\ i = 1, \ldots, l$ und $\hat{f} = L_l^{-1} f$. Dies ist jetzt wieder ein System der Form

$$\dot{z} = Az + g \tag{3.10}$$

während (3.8) ein System der Form

$$E\dot{z} = Az + g \tag{3.11}$$

ist.

Damit können wir die Methoden aus Teil II Kapitel 1 und 2 anwenden.

Definition 3.12 *Ein Polynom der Form*

$$L_l\lambda^l + L_{l-1}\lambda^{l-1} + \ldots + L_0\lambda =: L(\lambda)$$

wobei: $L_i \in \mathbb{C}^{n,n}$ *für* $i = 0, \ldots, l$ $\lambda \in \mathbb{C}$, *heißt* Matrix-Polynom. *Das zum System (3.8) bzw. (3.11) gehörende Matrixbüschel*

$$(E_L, A_L) = \left(\begin{bmatrix} I_n & & & \\ & \ddots & & \\ & & I_n & \\ & & & L_l \end{bmatrix}, \begin{bmatrix} 0 & I_n & & & \\ & 0 & I_n & & \\ & & \ddots & \ddots & \\ & & & 0 & I_n \\ -L_0 & -L_1 & \cdots & -L_{l-2} & -L_{l-1} \end{bmatrix} \right)$$

heißt Begleitmatrixbüschel *und die zu (3.9) bzw. (3.10) gehörende Matrix*

$$A_{\hat{L}} = \begin{bmatrix} 0 & I_n & & & \\ & 0 & I_n & & \\ & & \ddots & \ddots & \\ & & & 0 & I_n \\ -\hat{L}_0 & -\hat{L}_1 & \ldots & -\hat{L}_{l-2} & -\hat{L}_{l-1} \end{bmatrix}$$

heißt Begleitmatrix des Matrixpolynoms.

Beachte: $A_{\hat{L}}$ ist nur definiert, falls L_l^{-1} existiert.

Der klassische Ansatz zur Lösung von (3.2) ist natürlich auch möglich. Wir lösen zuerst die homogene Gleichung

$$L_l x^{(l)} + \ldots + L_1 x^{(1)} + L_0 = 0 \tag{3.13}$$

mit dem Ansatz $x = e^{\lambda t}$ und erhalten

$$L_l\lambda^l x + \ldots + L_1\lambda x + L_0 = 0 \tag{3.14}$$

oder

$$L(\lambda)x := \left(\sum_{i=0}^{l} L_i\lambda^i \right) x = 0\,. \tag{3.15}$$

Also muß λ eine Nullstelle von

$$\det \left(\sum_{i=0}^{l} L_i\lambda^i \right) = 0 \tag{3.16}$$

sein und für so ein λ dann

$$x \in \text{Kern} \left(\sum_{i=0}^{l} L_i\lambda^i \right)\,. \tag{3.17}$$

Wenn wir wieder alle Nullstellen von (3.16) bestimmen und zugehörige Vektoren x, die (3.17) erfüllen, so können wir wieder die Gesamtlösung der homogenen Gleichung bestimmen.

Lemma 3.18 *Es gilt*

$$\det L(\lambda) = \det(\lambda I_{l \cdot n} - A_{\hat{L}}) \det L_l$$

Beweis: Übung □

Man beachte, daß für $L_k = [v_{ij}^{(k)}]$ gilt

$$L(\lambda) = \begin{bmatrix} a_{11}(\lambda) & \dots & a_{1n}(\lambda) \\ \vdots & & \vdots \\ a_{nn}(\lambda) & \dots & a_{nn}(\lambda) \end{bmatrix} \tag{3.19}$$

mit

$$a_{ij}(x) = \sum_{k=0}^{l} v_{ij}^{(k)} \lambda^k \,, \tag{3.20}$$

d.h., jeder Eintrag der Matrix $L(\lambda)$ ist ein Polynom vom Grad $\leq l$ in λ. Für die Analyse der Nullstellen von $\det L(\lambda)$ ist es daher notwendig, Matrizen zu studieren, deren Einträge Polynome sind. Solche Matrizen heißen *λ-Matrizen* oder *Matrixpolynome*. Die Koeffizientenmatrix L_l der höchsten Potenz in λ heißt *führender Koeffizient*.

Kapitel 4

Matrixpolynome

In diesem Kapitel werden wir nun Matrixpolynome genauer untersuchen. Dabei kommen uns die allgemeinen theoretischen Grundlagen aus den ersten beiden Kapiteln zu Hilfe.

Lemma 4.1 *Die Menge $K^{n,n}$ der $n \times n$-Matrizen über einem Körper K bildet einen (nicht kommutativen) Ring mit Eins. Ebenso bildet die Menge $K^{n,n}[\lambda]$ der Matrixpolynome einen nicht kommutativen Ring mit Eins als Teilmenge von $K^{n,n}[[\lambda]]$ der formalen Potenzreihen mit Koeffizienten in $K^{n,n}$.*

Beweis: Übung □

Wir werden nun die Ergebnisse von Kapitel 1 und 2 verwenden, müssen dabei aber immer darauf achten, daß die Multiplikation nicht kommutativ ist.

Definition 4.2 *Sei $A(\lambda) \in K^{n,n}[\lambda]$. Dann heißt $A(\lambda)$ regulär, falls $\det(A(\lambda))$ nicht identisch verschwindet. Der Grad von $A(\lambda)$ ist der maximale Polynomgrad in den Koeffizienten, d.h. Grad $(\lambda I - M) = 1$ für alle $M \in K^{n,n}$ und Grad $(M) = 0$ für alle $M \in K^{n,n} \setminus \{0\}$. $A(\lambda)$ heißt invertierbar, falls es $B(\lambda) \in K^{n,n}[\lambda]$ gibt mit*

$$A(\lambda)B(\lambda) = B(\lambda)A(\lambda) = I_n.$$

Lemma 4.3 *Ein Matrixpolynom $A(\lambda) \in K^{n,n}[\lambda]$ ist invertierbar genau dann, wenn $\det(A(\lambda)) \in K \setminus \{0\}$ konstant ist.*

Beweis: Falls $\det(A(\lambda)) = c \neq 0$, $\forall \lambda \in K$, so können wir bilden

$$B(\lambda) = (A(\lambda))^{-1} = \frac{1}{c} \cdot \text{adj } (A(\lambda)),$$

und damit sind alle Elemente $b_{ij}(\lambda)$ Polynome in λ, denn die Elemente von adj $(A(\lambda))$ sind Minoren und damit Summen und Produkte von Elementen von $A(\lambda)$, also $B(\lambda) \in K^{n,n}[\lambda]$.

Für die Umkehrung sei $A(\lambda)$ invertierbar in $K^{n,n}[\lambda]$, so gibt es $B(\lambda)$ mit

$$B(\lambda)A(\lambda) = A(\lambda)B(\lambda) = I_n.$$

Also $\det(B(\lambda)) \cdot \det(A(\lambda)) = 1$. Da aber $\det(B(\lambda))$ und $\det(A(\lambda))$ Polynome in $K[\lambda]$ sind, so folgt aus Satz 2.28, daß $\det(A(\lambda)) = c \neq 0$. □

Definition 4.4 *Ein Matrixpolynom $A(\lambda) \in K^{n,n}[\lambda]$ heißt* unimodular, *falls $A(\lambda)$ invertierbar ist.*

Beispiel 4.5 Die Matrixpolynome in $\mathbb{Z}^{n,n}[\lambda]$

$$A_1(\lambda) = \begin{bmatrix} 1 & \lambda & -2\lambda^2 \\ 0 & 1 & \lambda^4 \\ 0 & 0 & 1 \end{bmatrix}, \qquad A_2(\lambda) = \begin{bmatrix} (\lambda-1)^2 & \lambda \\ \lambda - 2 & 1 \end{bmatrix}$$

sind unimodular, denn

$$\begin{aligned} \det A_1(\lambda) \equiv 1, \qquad \det A_2\lambda &= (\lambda-1)^2 - \lambda(\lambda-2) \\ &= \lambda^2 - 2\lambda + 1 - \lambda^2 + 2\lambda \equiv 1. \end{aligned}$$

Wir wollen nun Division mit Rest in $K^{n,n}[\lambda]$ einführen.

Definition 4.6 *Seien $A(\lambda), B(\lambda) \in K^{n,n}[\lambda]$ Matrixpolynome mit* grad $(A(\lambda)) = l$, grad $B(\lambda) = m$ *und $B(\lambda)$ hat in $K^{n,n}$ invertierbaren führenden Koeffizienten. Angenommen, es gibt $Q(\lambda), \tilde{Q}(\lambda), R(\lambda), \tilde{R}(\lambda) \in K^{n,n}[\lambda]$*

$$\begin{array}{llllll} \textit{mit} & \text{grad}\,(R(\lambda)) & < & \text{grad}\,(B(\lambda)) & \textit{oder} & R(\lambda) \equiv 0 \\ \textit{und} & \text{grad}\,(\tilde{R}(\lambda)) & < & \text{grad}\,(B(\lambda)) & \textit{oder} & \tilde{R}(\lambda) \equiv 0\,, \end{array}$$

so daß

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda) \tag{4.7}$$

bzw.

$$A(\lambda) = B(\lambda)\tilde{Q}(\lambda) + \tilde{R}(\lambda) \tag{4.8}$$

so heißen $Q(\lambda)$ rechter Quotient *und $\tilde{Q}(\lambda)$* linker Quotient *unter Division von $A(\lambda)$ durch $B(\lambda)$ und $R(\lambda)$ bzw. $\tilde{R}(\lambda)$ heißen* rechter Rest *bzw.* linker Rest.

Satz 4.9 *Seien $A(\lambda) = \sum_{i=0}^{l} \lambda^i A_i$, $B(\lambda) = \sum_{i=0}^{m} \lambda^i B_i \in K^{n,n}[\lambda]$ mit*

$$\text{grad}\ (A(\lambda)) = l,\ \text{grad}\ (B(\lambda)) = m \quad \textit{und} \quad \det B_m \neq 0.$$

Dann gibt es rechte und linke Quotienten und Reste bei Division von $A(\lambda)$ durch $B(\lambda)$.

Beweis: Falls $l < m$, so ist $\tilde{Q}(\lambda) = Q(\lambda) = 0$ und $R(\lambda) = \tilde{R}(\lambda) = A(\lambda)$.

Sei also nun $l \geq m$. „Dividiere“ zuerst durch $B_m\lambda^m$, d.h. bilde

$$A(\lambda) = A_l B_m^{-1} \lambda^{l-m} B(\lambda) + A^{(1)}(\lambda),$$

so hat $A^{(1)}(\lambda) = A^{(1)}_{l_1}\lambda^{l_1} + \ldots + A^{(1)}_0$ den Grad $l_1 \leq l-1$ und es gilt $A^{(1)}_{l_1} \neq 0$.

Falls $l_1 \geq m$, so wiederholen wir „Division“ durch $B(\lambda)$ für $A^{(1)}(\lambda)$ und erhalten

$$A^{(2)}(\lambda) = A^{(2)}_{l_2}\lambda^{l_2} + \ldots + A^{(2)}_0$$

mit $A^{(2)}_{l_2} \neq 0$ und $l_2 < l_1$.
Auf diese Weise erzeugen wir induktiv eine Folge $A^{(1)}(\lambda), A^{(2)}(\lambda), \ldots$ von Matrixpolynomen mit monoton fallendem Grad so lange, bis wir bei einem Matrixpolynom $A^{(r)}(\lambda)$ ankommen, für welches grad $A^{(r)}(\lambda) = l_r < m$. Dann setzen wir mit $A(\lambda) = A^{(0)}(\lambda)$

$$A^{(r-1)}(\lambda) = A^{(r-1)}_{l_{r-1}} B_m^{-1} \lambda^{l_{r-1}-m} B(\lambda) + A^{(r)}(\lambda)$$

und erhalten

$$\begin{aligned} A(\lambda) &= \left(A_l B_m^{-1}\lambda^{l-m} + A^{(1)}_{l_1} B_m^{-1}\lambda^{l_1-m} + \ldots + A^{(r-1)}_{l_{r-1}} B_m^{-1}\lambda^{l_{r-1}-m}\right) B(\lambda) + A^{(r)}(\lambda) \\ &=: Q(\lambda)B(\lambda) + R(\lambda). \end{aligned}$$

Division von der anderen Seite geht analog. □

Satz 4.10 *Unter den Voraussetzungen von Satz 4.9 sind rechte und linke Quotienten und Reste eindeutig bestimmt.*

Beweis: Angenommen es gibt $Q(\lambda), R(\lambda), \hat{Q}(\lambda), \hat{R}(\lambda)$, so daß

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda) = \hat{Q}(\lambda)B(\lambda) + \hat{R}(\lambda),$$

wobei $R(\lambda), \hat{R}(\lambda) \equiv 0$ oder vom Grad $< m$. Dann gilt $\left[Q(\lambda) - \hat{Q}(\lambda)\right] B(\lambda) = \hat{R}(\lambda) - R(\lambda)$.

Falls $Q(\lambda) \neq \hat{Q}(\lambda)$, so hat die linke Seite einen Grad $\geq m$, die rechte Seite aber einen Grad $< m$. Das ist ein Widerspruch, also $Q(\lambda) = \hat{Q}(\lambda)$ und damit $R(\lambda) = \hat{R}(\lambda)$.
Für die linken Quotienten ist der Beweis analog. □

Korollar 4.11 *(entfällt)*

Beispiel 4.12 Seien

$$A(\lambda) = \begin{bmatrix} \lambda^3 + \lambda^2 + 2\lambda - 1 & \lambda^3 + \lambda^2 + \lambda - 1 \\ \lambda^3 - 2\lambda - 2 & \lambda^2 + \lambda \end{bmatrix}, \; B(\lambda) = \begin{bmatrix} \lambda + 1 & 1 \\ \lambda + 1 & \lambda + 1 \end{bmatrix} \in \mathbb{R}^{2,2}[\lambda].$$

Da $B_1 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ invertierbar ist, $B_1^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$, können wir Division mit Rest machen.

$l = 3, m = 1$.

$$
\begin{aligned}
A(\lambda) &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \lambda^3 + \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \lambda^2 + \begin{bmatrix} 2 & 1 \\ -2 & 1 \end{bmatrix} \lambda + \begin{bmatrix} -1 & -1 \\ -2 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \lambda^2 B(\lambda) + A^{(1)}(\lambda) \\
&= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda^3 + \lambda^2 & \lambda^2 \\ \lambda^3 + \lambda^2 & \lambda^3 + \lambda^2 \end{bmatrix} + A^{(1)}(\lambda) \\
&= \begin{bmatrix} \lambda^3 + \lambda^2 & \lambda^3 + \lambda^2 \\ \lambda^3 + \lambda^2 & \lambda^2 \end{bmatrix} + A^{(1)}(\lambda) \qquad \Longrightarrow \\
A^{(1)}(\lambda) &= \begin{bmatrix} 2\lambda - 1 & \lambda - 1 \\ -\lambda^2 - 2\lambda - 2 & \lambda \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} \lambda^2 + \begin{bmatrix} 2 & 1 \\ -2 & 1 \end{bmatrix} \lambda + \begin{bmatrix} -1 & -1 \\ -2 & 0 \end{bmatrix} \\
A^{(1)}(\lambda) &= \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \lambda B(\lambda) + A^{(2)}(\lambda) \\
&= \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \lambda^2 + \lambda & \lambda \\ \lambda^2 + \lambda & \lambda^2 + \lambda \end{bmatrix} + A^{(2)}(\lambda) \\
&= \begin{bmatrix} 0 & 0 \\ -(\lambda^2 + \lambda) & -\lambda \end{bmatrix} + A^{(2)}(\lambda) \qquad \Longrightarrow \\
A^{(2)}(\lambda) &= \begin{bmatrix} 2\lambda - 1 & \lambda - 1 \\ -\lambda - 2 & 2\lambda \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \lambda + \begin{bmatrix} -1 & -1 \\ -2 & 0 \end{bmatrix} \\
A^{(2)}(\lambda) &= \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} B(\lambda) + A^{(3)}(\lambda) \\
&= \begin{bmatrix} 1 & 1 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} \lambda + 1 & 1 \\ \lambda + 1 & \lambda + 1 \end{bmatrix} + A^{(3)}(\lambda) \\
&= \begin{bmatrix} 2\lambda + 2 & \lambda + 2 \\ -\lambda - 1 & 2\lambda - 1 \end{bmatrix} + A^{(3)}(\lambda) \qquad \Longrightarrow \\
A^{(3)}(\lambda) &= \begin{bmatrix} -3 & -3 \\ -1 & 1 \end{bmatrix}
\end{aligned}
$$

Also ist

$$
A(\lambda) = \underbrace{\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} \lambda + \begin{bmatrix} 1 & 1 \\ -3 & 2 \end{bmatrix} \right)}_{Q(\lambda)} B(\lambda) + \underbrace{\begin{bmatrix} -3 & -3 \\ -1 & 1 \end{bmatrix}}_{R(\lambda)} .
$$

mit Grad $R(\lambda) <$ Grad $B(\lambda)$.

Betrachte nun den Einsetzhomomorphismus:

Für $A(\lambda) \in K^{n,n}[\lambda]$ und $B \in K^{n,n}$ ist der rechte (bzw. linke) Funktionswert gerade

$$
A(B) = A_l B^l + A_{l-1} B^{l-1} + \ldots + A_0 \tag{4.13}
$$

bzw.

$$\tilde{A}(B) = B^l A_l + B^{l-1} A_{l-1} + \ldots + B^1 A_1 + A_0. \tag{4.14}$$

Für skalare Polynome wissen wir, daß der Restsatz gilt, d.h., für

$$p(\lambda) = q(\lambda)(\lambda - b) + r \quad \text{gilt} \quad r = p(b), \tag{4.15}$$

denn $p(b) = q(b)(b-b) + r = r$. Ein ähnlicher Satz gilt auch für Matrixpolynome.

Satz 4.16 *Die rechten und linken Reste bei der Division von* $A(\lambda) = \sum_{i=0}^{l} A_i \lambda^i$ *durch* $\lambda I - B$ *sind* $A(B)$ *bzw.* $\tilde{A}(B)$.

Beweis: Es gilt

$$\lambda^j I - B^j = \left(\lambda^{j-1} I + \lambda^{j-2} B + \ldots + \lambda B^{j-2} + B^{j-1}\right)(\lambda I - B).$$

Multiplikation mit A_j und Aufsummieren aller rechten und linken Seiten für $j = 1, \ldots, l$ ergibt

$$\sum_{j=1}^{l} A_j \lambda^j - \sum_{j=1}^{l} A_j B^j = A(\lambda) - A(B) = C(\lambda) \cdot (\lambda I - B)$$

mit $C(\lambda) \in K^{n,n}[\lambda]$. Also gilt $A(\lambda) = C(\lambda)(\lambda I - B) + A(B)$.
Die Eindeutigkeit der Division mit Rest ergibt die Behauptung. Für die linken Reste ist der Beweis analog. □

Definition 4.17 *Eine Matrix* $X \in K^{n,n}$ *mit* $A(X) = 0$ *für* $A \in K^{n,n}[\lambda]$ *heißt* rechte Solvente *von* $A(\lambda)$. Linke Solvente *ist analog definiert als Matrix, für die* $\tilde{A}(X) = 0$.

Korollar 4.18 *Ein Matrixpolynom* $A(\lambda) \in K^{n,n}[\lambda]$ *ist rechts teilbar (links teilbar) durch* $\lambda I - B$ *mit Rest* 0 *genau dann, wenn* B *rechte (linke) Solvente von* $A(\lambda)$ *ist.*

Korollar 4.19 (Satz von Cayley–Hamilton)
Sei $A \in K^{n,n}$ *und* $P_A(\lambda) = \det(\lambda I - A)$, *so gilt* $P_A(A) = 0$.

(vgl. Teil I, Satz 6.10)

Beweis: Sei $A \in K^{n,n}$ und sei $B(\lambda) = adj(\lambda I - A) \in K^{n,n}[\lambda]$. ($B(\lambda)$ hat Grad $n-1$). Es gilt

$$(\lambda I - A)B(\lambda) = B(\lambda)(\lambda I - A) = \det(\lambda I - A) = P_A(\lambda).$$

$\overset{\text{Korollar 4.18}}{\Longrightarrow}$ Behauptung. □

Wir wollen nun Äquivalenztransformation für Matrixpolynome definieren und entsprechende Normalformen bestimmen. Wie schon in Teil 1, bei der Treppennormalform, führen wir Elementaroperationen ein, wie Multiplikation von Zeilen und Spalten mit $c \in K \setminus \{0\}$, vertauschen von Zeilen oder Spalten bzw. Addition eines polynomialen Vielfachen einer Zeile (Spalte) zu einer anderen.

Es ist klar, daß wir solche Elementaroperationen durch Multiplikation mit Matrizen von links oder rechts beschreiben können. Für $c \in K, b(\lambda) \in K[\lambda]$ seien Elementarmatrizen

$$M_i(c) = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & c & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}, \quad P_{ij} = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \end{matrix}$$

$$G_{ji}(\lambda) = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & b(\lambda) & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ i \\ \\ \\ \\ \end{matrix} \quad .$$
$$\begin{matrix} & & & & j & & \end{matrix}$$

Definition 4.20 *Seien $A(\lambda), B(\lambda) \in K^{n,n}[\lambda]$. $A(\lambda)$ heißt* äquivalent *zu $B(\lambda)$ ($A(\lambda) \sim B(\lambda)$), falls es eine endliche Anzahl Elementarmatrizen $E_1, \ldots, E_s$ gibt, so daß*

$$B(\lambda) = E_k(\lambda) \ldots E_1(\lambda) A(\lambda) E_{k+1}(\lambda) \ldots E_s(\lambda), \tag{4.21}$$

oder

$$B(\lambda) = P(\lambda) A(\lambda) Q(\lambda) \tag{4.22}$$

mit $P(\lambda)$, $Q(\lambda)$ unimodular.

Definition 4.23 *Ein Matrixpolynom $A(\lambda) \in K^{n,n}[\lambda]$ heißt* kanonisches Matrixpolynom, *falls es die folgende Eigenschaft hat:*

$$A = \begin{bmatrix} a_1(\lambda) & & \\ & \ddots & \\ & & a_n(\lambda) \end{bmatrix}$$

wobei $a_j(\lambda)$ entweder das Nullpolynom ist oder normiert, d.h. führender Koeffizient ist 1, *und weiterhin teilt das Polynom $a_{j-1}(\lambda)$ das Polynom $a_j(\lambda)$ für $j = 2, \ldots, n$.*

Es ist klar, daß dann alle Nullpolynome in den unteren Diagonalpositionen stehen müssen.

Damit können wir nun das Analogon zur Treppennormalform für Matrixpolynome herleiten.

Satz 4.24 *Jedes $A(\lambda) \in K^{n,n}[\lambda]$ ist äquivalent zu einem kanonischen Matrixpolynom.*

Beweis: Der Beweis ist konstruktiv mittels einer Folge von Transformationen mit Elementarmatrizen. O.B.d.A. sei $A(\lambda) \neq 0$.
Sei nun $a_{ij}(\lambda)$ das Element von $A(\lambda)$ vom kleinsten Grad (beachte $a_{ij} \neq 0$, da das Nullpolynom keinen Grad hat). Wende elementare Operationen P_{1i} von links und P_{1j} von rechts an. Setze

$$\tilde{A}(\lambda) := P_{1i}A(\lambda)P_{1j} = [\tilde{a}_{ij}(\lambda)].$$

Für jedes Element der ersten Zeile und Spalte mache Division durch $\tilde{a}_{11}(\lambda)$ mit Rest.

$$\begin{aligned} \tilde{a}_{1j}(\lambda) &= \tilde{a}_{11}(\lambda)q_{1j}(\lambda) + r_{1j}(\lambda), \\ \tilde{a}_{i1}(\lambda) &= \tilde{a}_{11}(\lambda)q_{i1}(\lambda) + r_{i1}(\lambda), \qquad i,j = 2,\ldots,n. \end{aligned}$$

Für jedes i,j verwende Transformationen mit Matrizen vom Typ $G_{ji}(\lambda)$, um das $q_{1j}(\lambda)$-fache der ersten Spalte von der j-ten Spalte abzuziehen und das $q_{i1}(\lambda)$-fache der ersten Zeile von der i-ten Zeile. Dadurch werden alle Elemente $\tilde{a}_{1j}(\lambda)$ durch $r_{1j}(\lambda)$ und alle Elemente $\tilde{a}_{i1}(\lambda)$ durch $r_{i1}(\lambda)$ ersetzt. Dabei sind alle r_{i1}, r_{1j} vom Grad kleiner als $a_{11}(\lambda)$ oder 0.

Falls nicht alle r_{1j}, r_{i1} Null sind, verwende eine weitere Permutation P_{i1} von links oder P_{1j} von rechts um das Polynom vom kleinsten Grad in die (1,1)-Position zu bringen, und teile dann wieder alles durch das neue Element in der $(1,1)$-Position mit Rest. Da der Grad monoton fällt, haben wir irgendwann ein Matrixpolynom der Form

$$\begin{bmatrix} a_{11}(\lambda) & 0 & \cdots & 0 \\ 0 & a_{22}(\lambda) & \cdots & a_{2n}(\lambda) \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}(\lambda) & \cdots & a_{nn}(\lambda) \end{bmatrix} \tag{4.25}$$

Es kann nun sein, daß es Elemente $a_{ij}(\lambda),\ i,j = 2,\ldots,n$ gibt, die einen Grad kleiner als $a_{11}(\lambda)$ haben, dann wiederholen wir den bisherigen Prozeß so lange, bis wir ein Matrixpolynom der Form (4.25) haben, wobei grad $(a_{11}(\lambda))$ der minimale vorkommende Grad ist.

Falls es nun ein Element $a_{ij}(\lambda)$ gibt, das nicht durch $a_{11}(\lambda)$ teilbar ist, so addieren wir Spalte j zur ersten dazu und wiederholen von Anfang an. Damit erhalten wir dann ein neues $a_{11}(\lambda)$ von kleinerem Grad. Dies machen wir solange, bis wir die Form

$$\begin{bmatrix} a_{11}(\lambda) & 0 & \ldots & 0 \\ 0 & b_{22}(\lambda) & \ldots & b_{2n}(\lambda) \\ \vdots & \vdots & & \vdots \\ 0 & b_{n2}(\lambda) & \ldots & b_{nn}(\lambda) \end{bmatrix}$$

haben, wobei $a_{11}(\lambda)$ alle $b_{ij}(\lambda)$ teilt, $i,j = 2,\ldots,n$. Verwende eine elementare Operation vom Typ $M_1(c)$ um $a_{11}(\lambda)$ zu normieren. Dieses normierte Polynom nennen wir $a_1(\lambda)$. Der Rest folgt dann mit Induktion. □

Definition 4.26 *Der* Rang *eines Matrixpolynoms $A(\lambda) \in K^{n,n}[\lambda]$ ist die maximale Anzahl von Polynomen ungleich 0 des zu $A(\lambda)$ äquivalenten kanonischen Matrixpolynoms.*

Beachte, das Matrixpolynom $\begin{bmatrix} 1-\lambda^2 & 0 \\ 0 & 1-\lambda^2 \end{bmatrix}$ hat Rang 2, obwohl für $\lambda = \pm 1$ der Rang der Matrix Null ist.

Das aus Satz 4.24 folgende kanonische Matrixpolynom zu einem beliebigen Matrixpolynom $A(\lambda)$ heißt *Smith–Normalform* von $A(\lambda)$. Die Polynome $a_j(\lambda)$ in der Smith-Normalform heißen *invariante Polynome* von $A(\lambda)$.

Die Smith–Normalform entspricht der Treppennormalform $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ einer Matrix und der Äquivalenz PAQ mit P, Q invertierbar. Als nächstes betrachten wir nun die Ähnlichkeit von Matrizen unter diesem Gesichtspunkt.

Satz 4.27 *Zwei Matrizen $A, B \in K^{n,n}$ sind ähnlich, d.h. $\exists\, P \in K^{n,n}$ invertierbar, so daß $B = PAP^{-1}$, genau dann, wenn die Matrixpolynome $\lambda I - A$ und $\lambda I - B$ äquivalent sind, d.h. dieselben invarianten Polynome besitzen.*

Beweis: Falls A, B ähnlich, so folgt $\lambda I - B = \lambda I - PAP^{-1} = P(\lambda I - A)P^{-1}$
$\Longrightarrow \lambda I - A, \lambda I - B$ sind äquivalent.

Angenommen, $\lambda I - A$, $\lambda I - B$ sind äquivalent $\Longrightarrow$ $\lambda I - A, \lambda I - B$ haben die gleichen invarianten Polynome und es gibt $P(\lambda), Q(\lambda)$ unimodular, so daß

$$P(\lambda)(\lambda I - A)Q(\lambda) = \lambda I - B$$

oder mit $M(\lambda) = [P(\lambda)]^{-1}$

$$M(\lambda)(\lambda I - B) = (\lambda I - A)Q(\lambda)$$

Division von $M(\lambda)$ von links mit $\lambda I - A$ und $Q(\lambda)$ von rechts mit $\lambda I - B$ ergibt

$$\begin{aligned} M(\lambda) &= (\lambda I - A)S(\lambda) + M_0, \\ Q(\lambda) &= R(\lambda)(\lambda I - B) + Q_0, \end{aligned}$$

wobei M_0, Q_0 konstante Matrizen sind wegen grad $(M_0) <$ grad $(\lambda I - A)$ und grad $(Q_0) <$ grad $(\lambda I - B)$.

$$\begin{aligned} \Longrightarrow \quad & ((\lambda I - A)S(\lambda) + M_0)\,(\lambda I - B) &&= (\lambda I - A)\,(R(\lambda)(\lambda I - B) + Q_0) \\ \Longleftrightarrow \quad & (\lambda I - A)\,(S(\lambda) - R(\lambda))\,(\lambda I - B) &&= (\lambda I - A)Q_0 - M_0(\lambda I - B) \end{aligned}$$

Die rechte Seite hat Grad 1, also folgt, daß $\lambda^2\,(S(\lambda) - R(\lambda)) = 0$

$$\begin{aligned} &\Longrightarrow \quad R(\lambda) = S(\lambda) \\ &\Longrightarrow \quad M_0(\lambda I - B) = (\lambda I - A)Q_0 \\ &\Longrightarrow \quad M_0 = Q_0,\ M_0 B = AQ_0 = AM_0. \end{aligned}$$

Es reicht damit zu zeigen, daß M_0 invertierbar ist. Bilde $P(\lambda) = (\lambda I - B)U(\lambda) + P_0$

$$
\begin{aligned}
\Longrightarrow I &= M(\lambda)P(\lambda) \\
&= ((\lambda I - A)S(\lambda) + M_0)\,((\lambda I - B)U(\lambda) + P_0) \\
&= (\lambda I - A)S(\lambda)(\lambda I - B)U(\lambda) + (\lambda I - A)M_0U(\lambda) \\
&\quad +(\lambda I - A)S(\lambda)P_0 + M_0P_0 \\
&= (\lambda I - A)\,(Q(\lambda)U(\lambda) + S(\lambda)P_0) + M_0P_0
\end{aligned}
$$

$$
\begin{aligned}
&\Longrightarrow \quad Q(\lambda)U(\lambda) + S(\lambda)P_0 = 0 \quad \text{und} \\
&\Longrightarrow \quad M_0P_0 = I \implies \det M_0 \neq 0 \implies M_0PM_0^{-1} = A.
\end{aligned}
$$

□

Kapitel 5

Standard-Tripel

Wir hatten bereits die Begleitmatrix eines Matrixpolynoms im Zusammenhang mit Systemen von linearen Differentialgleichungen höherer Ordnung kennengelernt in (3.8), Def. 3.12. Seien also $L(\lambda) = \sum_{j=0}^{l} \lambda^j L_j$ und

$$C_{\hat{L}} = \begin{bmatrix} 0 & I_n & & \\ & & \ddots & \\ & & & I_n \\ -\hat{L}_0 & \cdots & & -\hat{L}_{l-1} \end{bmatrix},$$

wobei $\hat{L}_i = L_l^{-1} L_i$.

Satz 5.1 *Die Matrixpolynome* $\lambda I_{n \cdot l} - C_{\hat{L}}$ *und* $\begin{bmatrix} L(\lambda) & 0 \\ 0 & I_{(l-1)n} \end{bmatrix}$ *sind äquivalent.*

Beweis: Setze

$$F(\lambda) = \begin{bmatrix} I_n & & & & 0 \\ -\lambda I_n & I_n & & & \\ & -\lambda I_n & \ddots & & \\ & & \ddots & \ddots & \\ 0 & & & -\lambda I_n & I_n \end{bmatrix}, \quad E(\lambda) = \begin{bmatrix} B_{l-1}(\lambda) & \cdots & \cdots & B_0(\lambda) \\ -I_n & \ddots & & \\ & \ddots & \ddots & \\ & & -I_n & 0 \end{bmatrix},$$

wobei $B_0(\lambda) = L_l$, $B_{r+1}(\lambda) = \lambda B_r(\lambda) + L_{l-r-1}$ für $r = 0, 1, 2, \ldots, l-2$. Dann gilt $\det(F(\lambda)) \equiv 1$ und $\det E(\lambda) \equiv \pm \det L_l$.

Da $\det(F(\lambda)) \equiv 1$, so ist $F(\lambda)^{-1}$ nach Lemma 4.3 auch ein Matrixpolynom und es gilt sofort

$$E(\lambda)(\lambda I - C_{\hat{L}}) = \begin{bmatrix} L(\lambda) & 0 \\ 0 & I_{(l-1)n} \end{bmatrix} F(\lambda).$$

□

Definition 5.2 *Die Nullstellen von* $\det A(\lambda)$ *für* $A(\lambda) \in K^{n,n}[\lambda]$ *heißen* latente Wurzeln *von* $A(\lambda)$ *und die Gesamtmenge der latenten Wurzeln heißt* Spektrum *von* $A(\lambda)$ *und wird im folgenden mit* $\sigma(A(\lambda))$ *bezeichnet.*
Die matrixwertige Funktion $A(\lambda)^{-1}$*, welche für alle* $\lambda \notin \sigma(A(\lambda))$ *definiert ist, heißt* Resolvente *von* $A(\lambda)$*.*

Satz 5.3 *Sei* $L(\lambda) \in \mathbb{C}^{n,n}[\lambda]$*. So gilt* $L(\lambda)^{-1} = P_1(\lambda I - C_{\hat{L}})^{-1} R_1$*, wobei*

$$P_1 = [\, I_n \, 0 \, \ldots \, 0 \,] \in \mathbb{C}^{n,n\cdot l} \qquad \textit{und} \qquad R_1 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ L_l^{-1} \end{bmatrix} \in \mathbb{C}^{l\cdot n,n}.$$

Beweis: Da $\det L(\lambda) = \det(\lambda I - C_{\hat{L}}) \cdot \det(L_l)$, so ist $L(\lambda)^{-1}$ definiert genau dort, wo $\det(\lambda I - C_{\hat{L}}) \neq 0$ ist.
Aus dem Beweis von Satz 5.1 folgt

$$\begin{bmatrix} L(\lambda)^{-1} & \\ & I_{(l-1)n} \end{bmatrix} = F(\lambda)(\lambda I - C_{\hat{L}})^{-1} E(\lambda)^{-1}. \tag{5.4}$$

Nach Definition von $E(\lambda)$ folgt

$$E(\lambda) \begin{bmatrix} 0 \\ \vdots \\ 0 \\ L_l^{-1} \end{bmatrix} = \begin{bmatrix} I_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

also ist die erste Blockspalte von $E(\lambda)^{-1}$ gerade R_1 und die ersten n Spalten von $F(\lambda)$ sind gerade P_1. □

Definition 5.5 *Ein Matrix-Tripel* (U, V, W) *heißt* zulässig *für ein Matrixpolynom* $L(\lambda)$ *vom Grad* l*, wenn*

$$U \in K^{n,ln},\ V \in K^{ln,ln} \quad \textit{und} \quad W \in K^{ln,n}.$$

Zwei zulässige Tripel (U_1, V_1, W_1) *und* (U_2, V_2, W_2) *heißen* ähnlich*, falls es eine nichtsinguläre Matrix* S *gibt, so daß*

$$U_1 = U_2 S,\ V_1 = S^{-1} V_2 S,\ W_1 = S^{-1} W_2. \tag{5.6}$$

Das Tripel $(P_1, C_{\hat{L}}, R_1)$ *ist zulässig und jedes zu* $(P_1, C_{\hat{L}}, R_1)$ *ähnliche Tripel heißt* Standard-Tripel.

Satz 5.7 *Sei* (U, V, W) *ein Standard-Tripel für* $L(\lambda)$ *und* $\lambda \notin \sigma(L(\lambda))$*, so gilt*

$$L(\lambda)^{-1} = U(\lambda I - V)^{-1} W.$$

Beweis:

$$P_1 = US, \quad C_{\hat{L}} = S^{-1}VS, \quad R_1 = S^{-1}W$$
$$\Longrightarrow \ (\lambda I - C_{\hat{L}})^{-1} = (S^{-1}(\lambda I - V)S)^{-1} = S^{-1}(\lambda I - V)^{-1}S.$$

□

Damit erhalten wir die Resolvente für $L(\lambda)$ mit Hilfe der Resolvente von $C_{\hat{L}}$. Umgekehrt können wir die Matrix $C_{\hat{L}}$ aus jedem Standard-Tripel erhalten.

Lemma 5.8 *Sei (U, V, W) ein Standard-Tripel für $L(\lambda)$ und*

$$Q = \begin{bmatrix} U \\ UV \\ \vdots \\ UV^{l-1} \end{bmatrix}. \tag{5.9}$$

So ist Q invertierbar und $C_{\hat{L}} = QVQ^{-1}$.

Beweis: Da $P_1 = [\, I_n, 0, \ldots, 0\,]$ und $C_{\hat{L}}$ sich aus (5.4) berechnen läßt, erhält man

$$\begin{bmatrix} P_1 \\ P_1C_{\hat{L}} \\ \vdots \\ P_1C_{\hat{L}}^{l-1} \end{bmatrix} = I_{l\cdot n}. \tag{5.10}$$

Mit (5.6) erhalten wir $QS = I_{l\cdot n} \Longrightarrow Q$ invertierbar und $S = Q^{-1} \Longrightarrow C_{\hat{L}} = QVQ^{-1}$. □

Satz 5.11 *Ein zulässiges Tripel (U, V, W) für $L(\lambda)$ ist ein Standard-Tripel genau dann, wenn die folgenden Bedingungen gelten:*

(a) Die Matrix Q aus (5.9) ist invertierbar.

(b) $L_lUV^l + L_{l-1}UV^{l-1} + \ldots + L_1UV + L_0U = 0$.

(c) $W = Q^{-1}R_1$ mit Q wie in (5.9) und R_1 wie in Satz 5.3.

Beweis: Wir zeigen zuerst, daß das Standard-Tripel $(P_1, C_{\hat{L}}, R_1)$ die Bedingungen (a)–(c) erfüllt.

(a) folgt aus Lemma 5.8.
Aus (5.10) folgt

$$P_1C_{\hat{L}}^l \ = \ (P_1C_{\hat{L}}^{l-1})C_{\hat{L}} \ = \ [\,0, \ldots, 0, I_n\,]C_{\hat{L}} \ = \ [-\hat{L}_0, \ldots, -\hat{L}_{l-1}].$$

$$\Longrightarrow\ L_l P_1 C_L^l = [\,-L_0,\,\ldots,\,-L_{l-1}],\quad \text{aber}$$

$$\sum_{j=0}^{l-1} L_j P_1 C_{\hat{L}}^j = [\,L_0,\,\ldots,\,L_{l-1}] \begin{bmatrix} P_1 \\ P_1 C_{\hat{L}} \\ \vdots \\ P_1 C_{\hat{L}}^{l-1} \end{bmatrix} = [\,L_0,\,\ldots,\,L_{l-1}]$$

$$\Longrightarrow\ \sum_{j=0}^{l} L_j P_1 C_{\hat{L}}^j = 0\,,$$

d.h., (b) gilt mit (U,V,W) ersetzt durch $(P_1, C_{\hat{L}}, R_1)$. Dann ist $Q = I$ und (c) gilt auch.

Sei nun (U,V,W) ein Standard-Tripel und sei

$$U = P_1 S,\quad V = S^{-1} C_{\hat{L}} S,\quad W = S^{-1} R_1\,.$$

$\Longrightarrow$

$$Q = \begin{bmatrix} U \\ UV \\ \vdots \\ UV^{l-1} \end{bmatrix} = \begin{bmatrix} P_1 \\ P_1 C_{\hat{L}} \\ \vdots \\ P_1 C_{\hat{L}}^{l-1} \end{bmatrix} S = S$$

und mit (5.10) folgt (a).

$\Longrightarrow \sum_{j=0}^{l} L_j U V^j = \left(\sum_{j=0}^{l} L_j P_1 C_{\hat{L}}^j \right) S = 0 \Longrightarrow$ (b), und mit $QW = S(S^{-1}R_1) = R_1$ folgt (c).

Für die Umkehrung nehmen wir an, daß ein zulässiges Tripel für $L(\lambda)$ die Bedingungen (a), (b), (c) erfüllt. Dann ist zu zeigen, daß (U,V,W) ähnlich zu $(P_1, C_{\hat{L}}, R_1)$ ist.
Aus der Definition für Q folgt, daß $U = P_1 Q$ und aus (a), daß Q invertierbar ist.
Aus (b) folgt dann $QV = C_{\hat{L}} Q \Longrightarrow V = Q^{-1} C_{\hat{L}} Q$.
Mit (c) folgt $U = P_1 Q,\quad V = Q^{-1} C_{\hat{L}} Q,\quad W = Q^{-1} R_1$. □

Damit folgt, daß ein Standard-Tripel durch die ersten zwei Terme festgelegt ist, man nennt die ersten zwei Terme auch *Standard-Paar* für $L(\lambda)$.

Satz 5.12 *Sei (U,V,W) ein zulässiges Tripel für $L(\lambda)$ und sei*

$$L(\lambda)^{-1} = U(\lambda I - V)^{-1} W. \tag{5.13}$$

Dann ist (U,V,W) Standard-Tripel für $L(\lambda)$.

Beweis: Machen wir hier nicht. □

Kapitel 6

Jordan–Tripel, Differentialgleichungen und Differenzengleichungen

Wir haben gesehen, daß alle Standard-Tripel ähnlich sind, wie in (5.6) definiert. Also können wir die Ähnlichkeitstransformation so wählen, daß

$$L(\lambda)^{-1} = X(\lambda I - J)^{-1}Y$$

mit J in Jordan'scher Normalform.

Solche Tripel (X, J, Y) heißen *Jordan–Tripel* und (X, J) heißt *Jordan–Paar* für $L(\lambda)$.
Da jedes Jordan–Paar ein Standard–Paar ist, ist die Matrix

$$Q = \begin{bmatrix} X \\ XJ \\ \vdots \\ XJ^{l-1} \end{bmatrix} \tag{6.1}$$

invertierbar nach Lemma 5.8 und

$$L_l X J^l + \ldots + L_1 X J + L_0 X = 0. \tag{6.2}$$

Sei $X = [x_1, \ldots, x_{ln}]$.

Sei für einen Moment angenommen, daß J diagonalisierbar, d.h. $J = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_{ln} \end{bmatrix}$,

so folgt für alle x_j, daß $x_j \neq 0$, denn sonst hätte Q eine Nullspalte. Weiter gilt mit (6.2), daß

$$L_l x_j \lambda_j^l + \ldots + L_1 x_j \lambda_j + L_0 x_j = 0 \quad \Longrightarrow \quad L(\lambda_j)x_j = 0 \quad \Longrightarrow \quad x_j \in \text{Kern } (L(\lambda_j)).$$

Definition 6.3 *Die Vektoren in* Kern $(L(\lambda_j))$ *heißen rechte* latente Vektoren *von* $L(\lambda)$ *zur* latenten Wurzel λ_j.

Falls J *nicht diagonalisierbar ist, so sei* $L^{(r)}(\lambda) = \dfrac{d^r}{d\lambda^r} L(\lambda)$. *Dann heißt eine Menge von Vektoren* $x_0, \ldots, x_k$ *mit* $x_0 \neq 0$ *eine* latente Jordan–Kette *der Länge* $k+1$ *für* $L(\lambda)$ *zur latenten Wurzel* λ_0, *falls*

$$\left\{ \begin{array}{lcl} L(\lambda_0)x_0 & = & 0, \\ L(\lambda_0)x_1 + \dfrac{1}{1!}L^{(1)}(\lambda_0)x_0 & = & 0, \\ \vdots & & \\ L(\lambda_0)x_k + \dfrac{1}{1!}L^{(1)}(\lambda_0)x_{k-1} + \ldots + \dfrac{1}{k!}L^{(k)}(\lambda_0)x_0 & = & 0\,. \end{array} \right. \tag{6.4}$$

Sei (X, J) ein Jordan–Paar von $L(\lambda)$ mit $J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{bmatrix}$, wobei $J_j \in \mathbb{C}^{n_j, n_j}$ und $X = [X_1, \ldots, X_s]$ mit $X_j \in \mathbb{C}^{n, n_j}$. Dann gilt

$$XJ^r = [X_1 J_1^r, \ldots, X_s J_s^r]$$

und damit aus (6.2)

$$L_l X_j J_j^l + \ldots + L_1 X_j J_j + L_0 X_j = 0\,.$$

Mit $X_j = [x_1^{(j)}, \ldots, x_{n_j}^{(j)}]$ und

$$J_j^r = \begin{bmatrix} \lambda_j^r & \binom{j}{1}\lambda_j^{r-1} & \ldots & \binom{j}{n_j-1}\lambda_j^{r-n_j+1} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \binom{j}{1}\lambda_j^{r-1} \\ 0 & & & \lambda_j^r \end{bmatrix}$$

folgt, daß die Spalten von X_j eine latente Jordan–Kette zum Eigenwert λ_j bilden.

Beispiel 6.5

$$\begin{array}{rcl} L(\lambda) & = & \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \lambda^2 & 0 \\ \lambda+1 & \lambda(\lambda-1) \end{bmatrix}, \\ \Longrightarrow \qquad L^{(1)}(\lambda) & = & \begin{bmatrix} 2\lambda & 0 \\ 1 & 2\lambda-1 \end{bmatrix}, \qquad L^{(2)}(\lambda) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}. \end{array}$$

$\det L(\lambda) = \lambda^3(\lambda - 1)$ hat 2 latente Wurzeln $\lambda_1 = 0, \lambda_2 = 1$.

$$L(0) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, L(1) = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}.$$

Damit kann jeder Vektor $\begin{bmatrix} 0 \\ z \end{bmatrix}$ als latenter Vektor zu λ_1, λ_2 genommen werden.

Die Jordan–Kette zu $\lambda_2 = 1$ hat die Länge 1, denn mit $x_0 = \begin{bmatrix} 0 \\ z \end{bmatrix}$ und $z \neq 0$ folgt:

$$L(1)x_1 + L^{(1)}(1)x_0 = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} x_1 + \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

hat keine Lösung. Für $\lambda_1 = 0$ ergibt sich aus

$$L(0)x_1 + L^{(1)}(0)x_0 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x_1 + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$x_1 = \begin{bmatrix} z \\ y \end{bmatrix}$ mit beliebigem $y \in \mathbb{C}$.
Aus

$$L(0)x_2 + L^{(1)}(0)x_1 + \tfrac{1}{2}L^{(2)}(0)x_0 =$$

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x_2 + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} z \\ y \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ z \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x_2 + \begin{bmatrix} 0 \\ 2z - y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

ergibt sich $x_2 = \begin{bmatrix} -2z + y \\ v \end{bmatrix}$ mit $v \in \mathbb{C}$ beliebig. Für

$$L(0)x_3 + L^{(1)}(0)x_2 + \tfrac{1}{2}L^{(2)}(0)x_1 + 0x_0 =$$

$$= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x_3 + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} -2z + y \\ v \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} z \\ y \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x_3 + \begin{bmatrix} z \\ -2z + 2y - v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

gibt es keine Lösung. Mit $z = 1$, $y = v = 0$ erhalten wir

$$x_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ x_2 = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

als latente Kette für $\lambda_1 = 0$ und

$$X = \begin{bmatrix} 0 & 0 & 1 & -2 \\ 1 & 1 & 0 & 0 \end{bmatrix}, \quad J = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

ist Jordan–Paar.

Betrachte nun wieder die Differentialgleichung

$$L_l x^{(l)}(t) + \ldots + L_1 x^{(1)}(t) + L_0 x(t) = f(t) \tag{6.6}$$

mit $\det L_l \neq 0$ und die homogene Gleichung

$$L_l x^{(l)}(t) + \ldots + L_1 x^{(1)}(t) + L_0 x(t) = 0. \tag{6.7}$$

Falls λ_0 latente Wurzel von $L(\lambda)$ mit latentem Vektor x_0 ist, so gilt $L(\lambda_0)x_0 = 0$ und $x_0 e^{\lambda_0 t}$ löst (6.7). Falls x_0, x_1 die ersten beiden Vektoren einer latenten Jordan–Kette sind, so sind

$$u_0(t) = x_0 e^{\lambda_0 t}, \quad u_1(t) = (tx_0 + x_1)e^{\lambda_0 t}$$

linear unabhängige Lösungen von (6.7), denn $\left(\frac{d}{dt} - \lambda_0 I\right) u_0(t) = 0$ und

$$\begin{aligned}
&\left(\frac{d}{dt} - \lambda_0 I\right)(tx_0 + x_1)e^{\lambda_0 t} = x_0 e^{\lambda_0 t}, && \text{d.h.}\\
&\left(\frac{d}{dt} - \lambda_0 I\right) u_1 = u_0, && \text{bzw.}\\
&\left(\frac{d}{dt} - \lambda_0 I\right)^2 u_1 = 0.
\end{aligned}$$

Wende Taylor-Entwicklung an für $L(\lambda)$ und $L(\lambda_0)$. Mit

$$\begin{aligned}
L(\lambda) &= L(\lambda_0) + \frac{1}{1!}L^{(1)}(\lambda_0)(\lambda - \lambda_0) + \ldots + \frac{1}{l!}L^{(l)}(\lambda_0)(\lambda - \lambda_0)^l\\
L\left(\frac{d}{dt}\right) &= L(\lambda_0) + \frac{1}{1!}L^{(1)}(\lambda_0)\left(\frac{d}{dt} - \lambda_0 I\right) + \ldots + \frac{1}{l!}L^{(l)}(\lambda_0)\left(\frac{d}{dt} - \lambda_0 I\right)^l\\
\Longrightarrow &\quad \left(\frac{d}{dt} - \lambda_0 I\right)^j u_1(t) = 0, \quad j = 2, 3, \ldots\\
\Longrightarrow\ L\left(\frac{d}{dt}\right) u_1 &= L(\lambda_0)u_1 + \frac{1}{1!}L^{(1)}(\lambda_0)u_0(t)\\
&= (L(\lambda_0)x_0)\, te^{\lambda_0 t} + \left(L(\lambda_0)x_1 + \frac{1}{1!}L^{(1)}(\lambda_0)x_0\right) e^{\lambda_0 t}\\
&= 0.
\end{aligned}$$

$\Longrightarrow$ u_1 ist Lösung von (6.7). Seien α, β so, daß

$$\alpha u_0 + \beta u_1 = 0 \quad \forall t \in \mathbb{R}.$$

$$\begin{aligned}
&\Longrightarrow \quad t(\alpha x_0) + (\alpha x_1 + \beta x_0) = 0 \quad \forall t \in \mathbb{R}\\
&\Longrightarrow \quad \alpha = 0 \quad \Longrightarrow \quad \beta x_0 = 0 \quad \Longrightarrow \quad \beta = 0\\
&\Longrightarrow \quad u_0, u_1 \quad \text{sind linear unabhängig.}
\end{aligned}$$

Analog zeigen wir:

Satz 6.8 *Sei $x_0, \ldots, x_{k-1}$ eine latente Jordan–Kette für $L(\lambda)$ zu λ_0. Dann sind die Funktionen*

$$\begin{aligned} u_0(t) &= x_0 e^{\lambda_0 t} \\ u_1(t) &= (tx_0 + x_1) e^{\lambda_0 t} \\ &\vdots \\ u_{k-1}(t) &= \left(\sum_{j=0}^{k-1} \frac{t^j}{j!} x_{k-1-j} \right) e^{\lambda_0 t} \end{aligned} \tag{6.9}$$

linear unabhängige Lösungen von (6.7).

Umgekehrt seien $x_0, \ldots, x_{k-1} \in \mathbb{C}^n$ mit $x_0 \neq 0$. Falls

$$u_{k-1}(t) = \left(\frac{t^{k-1}}{(k-1)!} x_0 + \ldots + t x_{k-2} + x_{k-1} \right) e^{\lambda_0 t}$$

Lösung von (6.7) ist, so ist λ_0 latente Wurzel von $L(\lambda)$ und $x_0, \ldots, x_{k-1}$ ist latente Jordan–Kette von $L(\lambda)$ zu λ_0.

Beweis: Erster Teil siehe Übung.
Für den zweiten Teil seien $u_0, \ldots, u_{k-2}$ wie in (6.9) und definiere $u_i \equiv 0$ für $i = -1, -2, \ldots$.
Dann gilt

$$\begin{aligned} \left(\frac{d}{dt} - \lambda_0 I \right)^j u_{k-1} &= u_{k-j-1}, \qquad j = 0, 1, 2, \ldots \qquad \text{und} \\ 0 &= L\left(\frac{d}{dt} \right) u_{k-1} \\ &= L(\lambda_0) u_{k-1} + L^{(1)}(\lambda_0) u_{k-2} + \ldots + \frac{1}{l!} L^{(l)}(\lambda_0) u_{k-l-1} \end{aligned}$$

Dann kürzen wir $e^{\lambda_0 t}$ aus den Gleichungen und erhalten (6.4) und da $x_0 \neq 0$, so folgt die Behauptung. □

Um die allgemeine Lösung von (6.6) zu bestimmen, betrachten wir wieder die Umformung als System durch $x_j = x^{(j)}(t)$. Dann folgt, daß (6.6) äquivalent ist zu

$$L_l \dot{x}_{l-1} + L_{l-1} x_{l-1} + \ldots + L_0 x_0 = f, \tag{6.10}$$

oder mit $z = \begin{bmatrix} x_0 \\ \vdots \\ x_{l-1} \end{bmatrix}$, $g = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ L_l^{-1} f \end{bmatrix}$

$$\dot{z} = C_{\hat{L}} z + g\,. \tag{6.11}$$

Die Lösungen von (6.11) haben die Form

$$z = e^{C_{\hat{L}} t} q + \int_0^t e^{C_{\hat{L}}(t-s)} g(s) ds \tag{6.12}$$

für $q \in \mathbb{C}^{nl} \Longrightarrow$ (mit $P_1, C_{\hat{L}}, R_1$ wie in Satz 5.3)

$$x(t) = P_1 e^{C_{\hat{L}} t} q + P_1 \int_0^t e^{C_{\hat{L}}(t-s)} R_1 f(s) ds,$$

da $P_1 z = x$. Damit haben wir gezeigt, daß gilt:

Lemma 6.13 *Jede Lösung von (6.6) hat die Form (6.12) für beliebiges $q \in \mathbb{C}^{ln}$.*

Satz 6.14 *Sei (U, V, W) Standard-Tripel für $L(\lambda)$, so hat jede Lösung von (6.6) die Form*

$$x(t) = U e^{Vt} q + U \int_0^t e^{V(t-s)} W f(s) ds \tag{6.15}$$

für ein beliebiges $q \in \mathbb{C}^{ln}$.

Beweis: Nach Definition des Standard-Tripels gibt es $S \in \mathbb{C}^{ln,ln}$ invertierbar, so daß

$$\begin{aligned} P_1 = US, \quad C_{\hat{L}} = S^{-1} V S, \quad R_1 = S^{-1} W \qquad & \text{und} \\ e^{C_{\hat{L}} \alpha} &= S^{-1} e^{V\alpha} S, \qquad \forall \alpha \in \mathbb{R}. \\ \Longrightarrow P_1 e^{C_{\hat{L}} \alpha} &= U e^{V\alpha} S \\ P_1 e^{C_{\hat{L}} \alpha} R_1 &= U e^{V\alpha} W. \end{aligned}$$

□

Satz 6.16 *Sei (U, V, W) ein Standard-Tripel für $L(\lambda)$. Dann gibt es eine eindeutige Lösung von (6.6), welche die Anfangsbedingungen*

$$x^{(j)}(0) = x_j, \qquad j = 0, \ldots, l-1$$

für beliebige gegebene Vektoren

$$x_0, \ldots, x_{l-1} \in \mathbb{C}^n$$

erfüllt. Die Lösung ist gegeben durch (6.15) mit

$$q = [\, W, VW, \ldots, V^{l-1} W \,] S_L \begin{bmatrix} x_0 \\ \vdots \\ x_{l-1} \end{bmatrix} \tag{6.17}$$

mit

$$S_L = \begin{bmatrix} L_1 & L_2 & L_3 & \cdots & L_l \\ L_2 & L_3 & & & \\ L_3 & & \iddots & & \\ \vdots & & & 0 & \\ L_l & & & & \end{bmatrix} \tag{6.18}$$

Beweis: Übung. □

Wir können auch analog Differenzengleichungen

$$L_l x_{j+l} + \ldots + L_1 x_{j+1} + L_0 x_j = f_j, \qquad j = 0, 1, 2, \ldots \tag{6.19}$$

lösen, wobei $(f_0, f_1, \ldots)$ eine Folge von Vektoren in $\mathbb{C}^n$ ist. Eine Folge $(x_0, x_1, \ldots)$, die (6.19) erfüllt, heißt Lösung von (6.19). Betrachte dazu zuerst die homogene Gleichung

$$L_l x_{j+l} + L_{l-1} x_{j+l-1} + \ldots + L_0 x_j = 0 \tag{6.20}$$

(Beachte im folgenden, daß wir setzen $\binom{j}{r} := 0$ für $j < r$.)

Satz 6.21 *Sei $x_0, \ldots, x_{k-1}$ latente Jordan–Kette für $L(\lambda)$ zu λ_0. So bilden die Folgen $(u_0^{(s)}, u_1^{(s)}, \ldots)$, $s = 0, 1, \ldots, k-1$, mit*

$$\begin{aligned} u_j^{(0)} &= \lambda_0^j x_0 && (6.22) \\ u_j^{(1)} &= \binom{j}{1} \lambda_0^{j-1} x_0 + \lambda_0^j x_1 \\ &\vdots && (6.23) \\ u_j^{(k-1)} &= \binom{j}{k-1} \lambda_0^{j-k+1} x_0 + \ldots + \binom{j}{1} \lambda_0^{j-1} x_{k-2} + \lambda_0^j x_{k-1} \end{aligned}$$

für $j = 0, 1, 2, \ldots$ linear unabhängige Lösungen von (6.20) im Vektorraum der Folgen von Vektoren in $\mathbb{C}^n$.

Beweis: Analog zu Beweis von Satz 6.8 □

Satz 6.24 *Seien $x_0, \ldots, x_{k-1} \in \mathbb{C}^n$ mit $x_0 \neq 0$. Falls die Folge $\left(u_0^{(k-1)}, u_1^{(k-1)}, \ldots\right)$ mit*

$$u_j^{(k-1)} = \sum_{s=0}^{k-1} \binom{j}{s} \lambda_0^{j-s} x_{k-1-s} \tag{6.25}$$

eine Lösung von (6.20) ist, so ist λ_0 latente Wurzel von $L(\lambda)$ und $x_0, \ldots, x_{k-1}$ latente Jordan–Kette von $L(\lambda)$ zu λ_0.

Beispiel 6.26 Löse

$$x_{j+2}^{(1)} = 0$$
$$x_{j+2}^{(2)} + x_{j+1}^{(1)} - x_{j+1}^{(2)} + x_j^{(1)} = 0$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x_{j+2} + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} x_{j+1} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x_j = 0.$$

Aus Beispiel 6.5 folgt, daß mit

$$u_j^{(0)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad j = 0, 1, \ldots$$
$$v_0^{(0)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad v_j^{(0)} = 0, \qquad j = 1, 2, \ldots$$
$$v_0^{(1)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad v_1^{(1)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad v_j^{(1)} = 0, \qquad j = 2, 3, \ldots$$
$$v_0^{(2)} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}, \; v_1^{(2)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; v_2^{(2)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad v_j^{(2)} = 0, \qquad j = 3, 4, \ldots$$

$\alpha_1 u^{(0)} + \alpha_2 v^{(0)} + \alpha_3 v^{(1)} + \alpha_4 v^{(2)} \quad \forall \alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{C}$ die allgemeine Lösung bildet.

Satz 6.27 *Sei (U, V, W) Standard-Tripel für $L(\lambda)$, so hat jede Lösung von (6.20) die Form $x_0 = Uz$ und für $j = 1, 2, \ldots$*

$$x_j = UV^j z + U \sum_{k=0}^{j-1} V^{j-k-1} W f_k \tag{6.28}$$

mit $z \in \mathbb{C}^{nl}$ beliebig.

Satz 6.29 *Sei (U, V, W) Standard-Tripel. Dann gibt es eine eindeutige Lösung von (6.19), welche die Anfangsbedingungen $x_j = a_j, \;\; j = 0, 1, \ldots, l-1$ für gegebene $a_0, a_1, \ldots, a_{l-1} \in \mathbb{C}^n$ erfüllt. Die Lösung ist wie in (6.28) mit $z = [\, W, VW, \ldots, V^{l-1}W \,] S_L \begin{bmatrix} a_0 \\ \vdots \\ a_{l-1} \end{bmatrix}$ und S_L wie in (6.18).*

Beweis: Wie Satz 6.16. □

Kapitel 7

Laplace-Transformation und Rationale Matrix-Funktionen

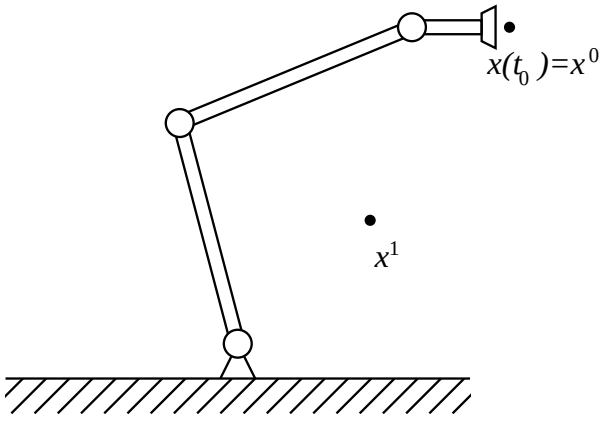
Wir kommen wieder zurück auf lineare Differentialgleichungen, aber diesmal als Steuerungsproblem, z. B. der Roboter aus Kapitel 3.

$$\begin{cases} \dot{x} = Ax + Bu & A \in \mathbb{C}^{n,n},\ B \in \mathbb{C}^{n,m} \\ y = Cx + Du & C \in \mathbb{C}^{p,n},\ D \in \mathbb{C}^{p,m} \\ x(t_0) = x_0 \end{cases} \tag{7.1}$$

Dabei ist x der Zustand des Systems (Ortskoordinaten, Winkel, Geschwindigkeiten),
u ist die Steuerung,
y ist der Ausgang (Meßgrößen).

Diese Formulierung heißt *Zustandsraumbeschreibung.*

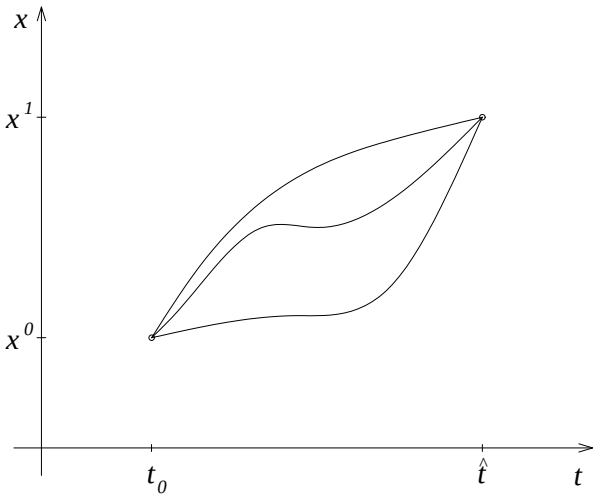
Beispiel 7.2



Bewege den Putzkopf von Position x^0 nach x^1.

Bild 7.3

Für vorgegebene $u(t)$ bestimmt die Lösung $x(t)$ die Bahn, es muß dann $u(t)$ so gewählt werden, daß $x(\hat{t}) = x^1$ für ein $\hat{t} > t_0$. Es gibt allerdings viele mögliche $u(t)$ und es ist auch nicht eindeutig, für welches $\hat{t}$ sich dann der Wert x^1 ergibt. Diese Theorie und die Algorithmen dafür werden dann in der Steuerungstheorie behandelt. Es gibt verschiedene Lösungsansätze. Ein wichtiger und vielfach in der Praxis verwendeter Ansatz ist die Laplace–Transformation.



Definition 7.4 *Sei* $f : [0, \infty) \to \mathbb{R}$ *integrierbar und es gelte* $|f(t)| \leq ke^{ct}$. *Dann heißt* $\mathcal{L}(f) = \hat{f}(s) = \int_0^\infty e^{-st} f(t)dt$ *die* Laplace–Transformierte von f.

Lemma 7.5 *Es gelten die folgenden Regeln:*

(i) $\mathcal{L}\left(a_1 f_1(t) + a_2 f_2(t)\right) = a_1 \mathcal{L}(f_1(t)) + a_2 \mathcal{L}(f_2(t))$;

(ii) $\mathcal{L}\left(\int_0^t f_1(t-\tau) f_2(\tau) d\tau\right) = \mathcal{L}\left(f_1(t)\right) \cdot \mathcal{L}\left(f_2(t)\right)$;

(iii) $\mathcal{L}\left(\int_0^t f(\tau) d\tau\right) = \frac{1}{s}\hat{f}(s)$;

(iv) $\mathcal{L}\left(f^{(n)}(t)\right) = s^n \hat{f}(s) - s^{n-1} g - \ldots - s g^{(n-2)} - g^{(n-1)}$, *wobei* $g^{(\nu)} = \lim_{t \to 0^+} \left(\frac{d^\nu f(t)}{dt^\nu}\right)$;

(v) $\mathcal{L}\left(f(t-b)\right) = e^{-bs} \mathcal{L}\left(f(t)\right)$;

(vi) für $a > 0$ *gilt* $\mathcal{L}\left(f(at)\right) = \frac{1}{a}\hat{f}\left(\frac{s}{a}\right)$;

(vii) $\mathcal{L}\left(e^{-\alpha t} f(t)\right) = \hat{f}(s + \alpha)$;

(viii) $\mathcal{L}\left(t^n f(t)\right) = (-1)^n \hat{f}^{(n)}(s)$;

(ix) falls $\frac{1}{t} f(t)$ *Laplace-transformierbar ist, so gilt* $\mathcal{L}\left(\frac{1}{t} f(t)\right) = \int_s^\infty \hat{f}(r) dr$.

Beweis: Übung □

Nach Laplace–Transformation ergibt sich

$$\begin{aligned} (\lambda I - A)\hat{x} &= B\hat{u} \\ \hat{y} &= C\hat{x} + D\hat{u} \end{aligned} \tag{7.6}$$

oder

$$\begin{aligned} \hat{y} &= [C(\lambda I - A)^{-1}B + D]\,\hat{u}, \\ \hat{x} &= (\lambda I - A)^{-1}B\hat{u} \end{aligned} \tag{7.7}$$

Diese Formulierung in (7.4) heißt *Frequenzraumbeschreibung*, die Matrix

$$C(\lambda I - A)^{-1}B + D = W(\lambda) = [w_{ij}(\lambda)] \tag{7.8}$$

heißt *Transferfunktion* vom Eingang $\hat{u}$ zum Ausgang $\hat{y}$ im Frequenzraum. Voraussetzung ist, daß $\lambda I - A$ regulär ist, d.h. daß $\det(\lambda I - A) \not\equiv 0$. Was hat nun $W(\lambda)$ für eine Form?

$$(\lambda I - A)^{-1} = \frac{1}{\det(\lambda I - A)}\mathrm{adj}(\lambda I - A) \tag{7.9}$$

$\Longrightarrow$

$$W(\lambda) = [w_{ij}] = \left[\frac{p_{ij}(\lambda)}{q_{ij}(\lambda)}\right] \in Q\,(R[\lambda])^{p,m}\,. \tag{7.10}$$

Betrachte nun $Q(R[\lambda])^{p,m} = \left\{\left[\frac{p_{ij}(\lambda)}{q_{ij}(\lambda)}\right]\right\}$, die Menge der rationalen Matrixfunktionen.

Im folgenden setzen wir voraus, daß grad $p_{ij}(\lambda) <$ grad $q_{ij}(\lambda)$, $\forall i = 1, \ldots, p, j = 1, \ldots, m$, wie das natürlich bei $(\lambda I - A)^{-1}$ in (7.9) und damit auch bei $W(\lambda)$ aus (7.8) der Fall ist.

Eine direkte Frage ist nun, ob wir eine gegebene rationale Matrixfunktion $W(\lambda) \in Q(R[\lambda])^{p,m}$ immer in der Form (7.8) schreiben können.

Definition 7.11 *Sei $W(\lambda) \in Q(R[\lambda])^{p,m}$. Falls es $n \in \mathbb{N}$ und Matrizen $A \in R^{n,n}, B \in R^{n,m}, C \in R^{p,n}$ und $D \in R^{p,m}$ gibt, so daß (7.8) gilt, so heißt das Quadrupel*

$$(A, B, C, D)_n$$

Realisierung *von $W(\lambda)$ im Ring R.*

Realisierungen von rationalen Matrixfunktionen spielen eine große Rolle in der Modellbildung dynamischer Systeme.

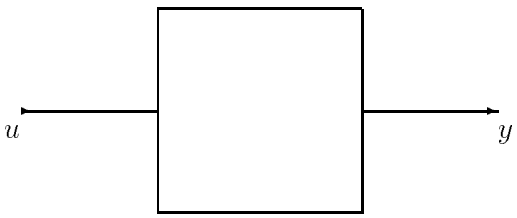


Bild 7.12

Für gegebene Eingangssignale $\hat{u}_0, \hat{u}_1, \hat{u}_2, \hat{u}_3, \ldots$ und Ausgangssignale $\hat{y}_0, \hat{y}_1, \hat{y}_2, \hat{y}_3, \ldots$ im Frequenzraum suche eine Transferfunktion $W(\lambda) = D + C(\lambda I - A)^{-1}B$, so daß

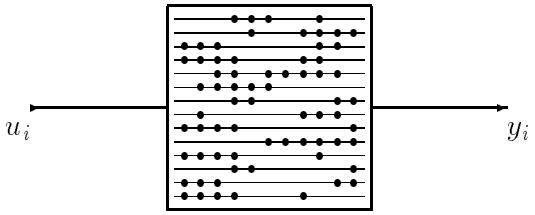
$$\hat{y}_i = W(\lambda)\hat{u}_i \quad \forall_i = 0, 1, 2, 3, \ldots .$$

Dies ist eine rationale Interpolationsaufgabe, soll hier nicht behandelt werden, hat aber immer eine Lösung.

Beispiel 7.13 Chip testen.

$\{\hat{u}_i\}$ ist Folge in $\mathbb{F}_2$,
$\{\hat{y}_i\}$ ist Folge in $\mathbb{F}_2$.
Suche $W(\lambda) \in Q(\mathbb{F}_2[\lambda])^{p,m}$, so daß

$$\hat{y}_i = W(\lambda)\hat{u}_i \qquad \forall i = 0, 1, 2, \ldots .$$



Wenn man so eine rationale Matrixfunktion hat, finde eine Realisierung und dann eine *minimale Realisierung*, d.h. Realisierung mit minimalem n.

Wir beschränken uns im folgenden auf $Q(\mathbb{C}[\lambda])^{p,m}$.

Lemma 7.14 *$W(\lambda) \in Q(\mathbb{C}[\lambda])^{p,m}$ habe eine Realisierung*

$$W(\lambda) = D + C(\lambda I - A)^{-1}B \tag{7.15}$$

mit $A \in \mathbb{C}^{n,n}, B \in \mathbb{C}^{n,m}, C \in \mathbb{C}^{p,m}, D \in \mathbb{C}^{p,m}$. So gilt

$$D = \lim_{\lambda \to \infty} W(\lambda). \tag{7.16}$$

Beweis: Betrachte $\lim_{\lambda \to \infty} (\lambda I - A)^{-1}$. Dazu sei

$$A = T \begin{bmatrix} J_{r_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{r_m}(\lambda_m) \end{bmatrix} T^{-1} = TJT^{-1}$$

die Jordan'sche Normalform von A. So gilt

$$\begin{aligned} (\lambda I - A)^{-1} &= T(\lambda I - J)^{-1}T^{-1} \\ &= T \begin{bmatrix} (\lambda I - J_{r_1}(\lambda_1))^{-1} & & \\ & \ddots & \\ & & (\lambda I - J_{r_m}(\lambda_m))^{-1} \end{bmatrix} T^{-1}. \end{aligned}$$

Aber

$$\left(\lambda I - J_{r_j}(\lambda_j)\right)^{-1} = \begin{bmatrix} (\lambda-\lambda_j)^{-1} & (\lambda-\lambda_j)^{-2} & \cdots & (\lambda-\lambda_j)^{-r_j} \\ & \ddots & \ddots & \vdots \\ & & \ddots & (\lambda-\lambda_j)^{-2} \\ & & & (\lambda-\lambda_j)^{-1} \end{bmatrix}$$

$$\Longrightarrow \lim_{\lambda\to\infty}(\lambda I - A)^{-1} = T\left(\lim_{\lambda\to\infty}(\lambda I - J)\right)^{-1} T^{-1} = 0$$

$$\Longrightarrow \lim_{\lambda\to\infty} W(\lambda) = \lim_{\lambda\to\infty}\left(D + C(\lambda I - A)^{-1}B\right) = D.$$

□

Lemma 7.17 *Seien* $H(\lambda) = \sum_{j=0}^{l-1} \lambda^j H_j \in \mathbb{C}^{n,n}[\lambda]$ *und* $L(\lambda) = \lambda^l I + \sum_{j=0}^{l-1} \lambda_j L_j \in \mathbb{C}^{n,n}[\lambda]$. *Setze*

$$B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I_n \end{bmatrix},\ A = \begin{bmatrix} 0 & I & & \\ & & \ddots & \\ & & & I \\ -L_0 & -L_1 & \ldots & -L_{l-1} \end{bmatrix},\ C = [\, H_0, \ldots, H_{l-1}\,].$$

So gilt

$$H(\lambda)L(\lambda)^{-1} = C(\lambda I - A)^{-1}B. \tag{7.18}$$

Beweis: Seien

$$F(\lambda) = \begin{bmatrix} I & & & \\ -\lambda I & I & & \\ & \ddots & \ddots & \\ & & -\lambda I & I \end{bmatrix} \quad \text{und} \quad E(\lambda) = \begin{bmatrix} B_{l-1}(\lambda) & \cdots & \cdots & B_0(\lambda) \\ -I & & & 0 \\ & \ddots & & \vdots \\ & & -I & 0 \end{bmatrix}$$

wobei $B_0(\lambda) = I,\ B_{r+1}(\lambda) = \lambda B_r(\lambda) + L_{l-r-1}$, für $r = 0, \ldots, l-2$. Dann gilt

$$E(\lambda)(\lambda I - A) = \begin{bmatrix} L(\lambda) & 0 \\ 0 & I \end{bmatrix} F(\lambda),$$

siehe Kapitel über Standard–Tripel.

$$\begin{bmatrix} L(\lambda)^{-1} & 0 \\ 0 & I \end{bmatrix} = F(\lambda)(\lambda I - A)^{-1}E(\lambda)^{-1}. \tag{7.19}$$

Sei $P = [\, I_n, 0, \ldots, 0\,]$, so gilt $PF(\lambda) = P$ und damit

$$L(\lambda)^{-1} = P(\lambda I - A)^{-1}B. \tag{7.20}$$

Führe $C_1(\lambda), \ldots, C_l(\lambda)$ ein durch

$$\begin{bmatrix} C_1(\lambda) \\ \vdots \\ C_l(\lambda) \end{bmatrix} = (\lambda I - A)^{-1} B \quad \Longrightarrow \quad C_1(\lambda) = L(\lambda)^{-1}$$

und da $(\lambda I - A) \begin{bmatrix} C_1(\lambda) \\ \vdots \\ C_l(\lambda) \end{bmatrix} = B$, so folgt $C_i(\lambda) = \lambda^{i-1} C_1(\lambda)$ für $1 \leq i \leq l$,

$$\Longrightarrow C(\lambda I - A)^{-1} B = \sum_{j=0}^{l-1} H_j C_{j+1}(\lambda) \; = \; H(\lambda) L(\lambda)^{-1}.$$

□

Damit folgt nun ein sehr wichtiger Satz.

Satz 7.21 *Für jede Matrixfunktion $W(\lambda) \in Q(\mathbb{C}[\lambda])^{p,m}$, mit $\lim_{\lambda \to \infty} W(\lambda) = D$ endlich, gibt es $n \in \mathbb{N}$ und Matrizen $A \in \mathbb{C}^{n,n}, B \in \mathbb{C}^{n,m}$ und $C \in \mathbb{C}^{p,m}$, so daß $W(\lambda)$ die Realisierung*

$$W(\lambda) = D + C(\lambda I - A)^{-1} B$$

hat.

Beweis: Sei $W(\lambda) \in Q(\mathbb{C}[\lambda])^{p,m}$. So gibt es ein normiertes skalares Polynom $\ell(\lambda) \in \mathbb{C}[\lambda]$ (das kgV aller Nenner), so daß $\ell(\lambda) W(\lambda) \in \mathbb{C}^{p,m}[\lambda]$.

Setze $H(\lambda) = \ell(\lambda)(W(\lambda) - D)$, so ist $H(\lambda) \in \mathbb{C}^{p,m}[\lambda]$. Sei $L(\lambda) = \ell(\lambda) I_n$, ($L(\lambda)$ hat führenden Koeffizienten I_n). Dann ist $W(\lambda) = D + H(\lambda) L(\lambda)^{-1}$ und da

$$\lim_{\lambda \to \infty} H(\lambda) L(\lambda)^{-1} = \lim_{\lambda \to \infty} (W(\lambda) - D) = 0 \text{ ist,}$$

so ist grad $(H(\lambda))$ < grad $(L(\lambda))$ und nach Lemma 7.17 gibt es A, B, C, so daß

$$W(\lambda) = D + C(\lambda I - A)^{-1} B.$$

□

Realisierungen sind im allgemeinen nicht eindeutig. Falls zum Beispiel (A, B, C, D) eine Realisierung ist, so auch $(\hat{A}, \hat{B}, \hat{C}, D)$ mit

$$\hat{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A & A_{23} \\ 0 & 0 & A_{33} \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} B_1 \\ B \\ 0 \end{bmatrix}, \quad \hat{C} = [\, 0 \; C \; C_1 \,]$$

für alle Matrizen A_{ij}, B_1, C_1 von passender Größe. Außerdem gilt, falls (A, B, C, D) eine Realisierung von $W(\lambda)$ ist, so auch $(SAS^{-1}, SB, CS^{-1}, D)$.

Wenn wir also viele Realisierungen finden können, so ist es in der Praxis meistens wünschenswert, die Systemdimension n, d.h. die Größe von A möglichst klein zu machen.

Anstelle von $\lim_{\lambda \to \infty} W(\lambda)$ verwenden wir auch die Schreibweise $W(\infty)$.

Definition 7.22 *Sei $W(\lambda) \in Q(\mathbb{C}[\lambda])^{p,m}$ mit $W(\infty) = D$ endlich. Eine Realisierung (A, B, C, D) von $W(\lambda)$ mit $A \in \mathbb{C}^{n,n}, B \in \mathbb{C}^{n,m}, C \in \mathbb{C}^{p,n}, D \in \mathbb{C}^{p,m}$ heißt* minimal, *falls es keine Realisierung $(\tilde{A}, \tilde{B}, \tilde{C}, D)$ gibt mit $\tilde{A} \in \mathbb{C}^{r,r}, \tilde{B} \in \mathbb{C}^{r,m}, \tilde{C} \in \mathbb{C}^{p,r}$ und $r < n$.*

Wir werden nun zeigen, daß die Minimalität einer Realisierung direkt vom Rang der Matrizen

$$\begin{aligned} \mathcal{C}(A,B) &= [B, AB, A^2B, \ldots, A^{n-1}B] \text{ und} \\ \mathcal{O}(A,C) &= \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \end{aligned} \tag{7.23}$$

abhängt.

Definition 7.24 *Das Matrizenpaar (A, B) mit $A \in \mathbb{C}^{n,n}, B \in \mathbb{C}^{n,m}$ heißt* steuerbar, *wenn* $\operatorname{Rang} \mathcal{C}(A,B) = n$. *Das Paar (A, C) mit $A \in \mathbb{C}^{n,n}, C \in \mathbb{C}^{p,n}$ heißt* beobachtbar, *wenn* $\operatorname{Rang} \mathcal{O}(A,C) = n$.

Diese Begriffe sind aus der Steuerungstheorie motiviert. Das Steuerungsproblem

$$\dot{x} = Ax + Bu, \quad y = Cx, \quad x(0) = x^0 \tag{7.25}$$

mit $A \in \mathbb{C}^{n,n}, B \in \mathbb{C}^{n,m}$ ist steuerbar (d.h., zu gegebenem x^1 gibt es $t_1 \in \mathbb{R}$ und $u(t)$ stückweise stetig, so daß für die Lösung von (7.25) gilt, daß $x(t_1) = x^1$) genau dann, wenn (A, B) steuerbar.

Analog ist das System (7.25) beobachtbar (d.h., für jedes t_0 und Lösungen $x, \tilde{x}$ mit gleichem u und $Cx = C\tilde{x}$ gilt $x(t) = \tilde{x}(t)$, $\forall\, t \le t_0$) genau dann, wenn (A, C) beobachtbar ist.

Satz 7.26 *Gegeben sei ein System (A, B, C), so gibt es S invertierbar, so daß*

$$S^{-1}AS = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{bmatrix}, \quad S^{-1}B = \begin{bmatrix} B_1 \\ B_2 \\ 0 \end{bmatrix}, \quad CS = [0\ C_2\ C_3] \quad \textit{mit}$$

$$\left(\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \right) \qquad \textit{steuerbar},$$

und das System (A_{22}, B_2, C_2) ist beobachtbar und steuerbar.

Beweis: erfolgt hier nicht. □

Im folgenden bezeichnen wir für eine Matrix $Z \in \mathbb{C}^{m,n}$ mit Z^{-L} eine Linksinverse von Z und mit Z^{-R} eine Rechtsinverse von Z, d.h.

$$\begin{aligned} ZZ^{-R} &= I_m, \\ Z^{-L}Z &= I_n. \end{aligned}$$

Diese Inversen müssen nicht exisiteren und sind eindeutig nur für $m = n$.

Satz 7.27 *Seien $(A_1, B_1, C_1), (A_2, B_2, C_2)$ Realisierungen einer rationalen Matrixfunktion $W(\lambda)$ mit*

(A_1, C_1), (A_2, C_2) *beobachtbar und*
(A_1, B_1), (A_2, B_2) *steuerbar und* $A_1 \in \mathbb{C}^{n_1, n_1}$, $A_2 \in \mathbb{C}^{n_2, n_2}$.

Dann gilt $n_1 = n_2$ und es gibt S nichtsingulär, so daß

$$A_1 = S^{-1} A_2 S, \; B_1 = S^{-1} B_2, \; C_1 = C_2 S. \tag{7.28}$$

Weiterhin ist S eindeutig und gegeben durch

$$\begin{aligned} S &= \begin{bmatrix} C_2 \\ C_2 A_2 \\ \vdots \\ C_2 A_2^{p-1} \end{bmatrix}^{-L} \begin{bmatrix} C_1 \\ C_1 A_1 \\ \vdots \\ C_1 A_1^{p-1} \end{bmatrix} \\ &= [B_2, A_2 B_2, \ldots, A_2^{m-1} B_2][B_1, A_1 B_1, \ldots, A_1^{p-1} B_1]^{-R} \end{aligned} \tag{7.29}$$

Dabei ist $p \in \mathbb{N}$ so, daß

$$\text{Kern} \begin{bmatrix} C_i \\ C_i A_i \\ \vdots \\ C_i A_i^{p-1} \end{bmatrix} = \{0\}, \quad \textit{und} \quad \text{Rang}\,[B_i, A_i B_i, \ldots, A_i^{p-1} B_i] = n, \quad i = 1, 2.$$

Beweis: ersparen wir uns hier. □

Korollar 7.30 *Sei (A, B, C) eine Realisierung von $W(\lambda)$. Dann sind (A, C) beobachtbar und (A, B) steuerbar genau dann, wenn die Realisierung minimal ist.*

Beweis: Sei (A, B, C) eine minimale Realisierung. Nach Satz 7.26 gibt es eine transformierte Realisierung (A', B', C') mit (A', B') steuerbar und (A', C') beobachtbar.
Da aber nach Satz 7.27 alle minimalen Realisierungen ähnlich sind, folgt, daß die Dimension von A und A' gleich ist.
Umgekehrt sei (A, B, C) eine nicht minimale Realisierung mit (A, B) steuerbar und (A, C) beobachtbar. Dann gibt es eine minimale Realisierung (A', B', C') und damit Untermatrizen (A'', B'', C'') von (A', B', C'), die steuerbar und beobachtbar sind.
Da aber dann A'' und A ähnlich sind, folgt die Behauptung. □

Kapitel 8

Matrixgruppen

Wir hatten schon in Teil 1 und Teil 2 gesehen, daß die invertierbaren Matrizen in $\mathbb{C}^{n,n}(\mathbb{R}^{n,n})$ eine Gruppe bilden und analog auch die unitären (orthogonalen) Matrizen $\mathcal{O}_n(\mathbb{C})$ bzw. $\mathcal{O}_n(\mathbb{R})$. Wir werden nun weitere solcher Matrixgruppen kennenlernen.

Definition 8.1 *Sei* $J = \begin{bmatrix} I_p & \\ & -I_q \end{bmatrix} \in \mathbb{R}^{n,n}$ *und sei* $\langle x, y\rangle_J = y^\top Jx$ *eine Abbildung von* $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, *so heißt eine Matrix* $Q \in \mathbb{R}^{n,n}$ J–orthogonal[1], *falls*

$$\langle Qx, Qy\rangle_J = \langle x, y\rangle_J, \tag{8.2}$$

und $S \in \mathbb{R}^{n,n}$ *heißt* J–symmetrisch, *falls*

$$\langle Sx, y\rangle_J = \langle x, Sy\rangle_J. \tag{8.3}$$

Die Abbildung $\langle \cdot, \cdot\rangle_J : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ definiert ein indefinites Skalarprodukt, bei dem die gleichen Bedingungen wie bei einem normalen Skalarprodukt gelten, bis auf die letzte Bedingung, denn

$$\langle x, x\rangle_J = 0 \not\Rightarrow x = 0.$$

Beispiel 8.4 $J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad x = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \neq 0.$

$$x^\top Jx = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0.$$

Satz 8.5 *Sei* $\mathcal{S}_J = \{Q \mid Q \text{ ist } J\text{-orthogonal}\}$, *so gilt für alle* $Q \in \mathcal{S}_J$, *daß*

$$Q^\top JQ = J$$

und $\mathcal{S}_J$ *ist eine Gruppe bezüglich der normalen Matrixmultiplikation, und damit Untergruppe von* $GL_n(\mathbb{R})$.

[1] manchmal werden auch solche Matrizen *symplektisch* genannt (hier aber siehe Def. 8.27)

Beweis: Der Beweis der Gruppeneigenschaft ist Übungsaufgabe.
Sei $\{x_1, \ldots, x_n\}$ eine Basis von $\mathbb{R}^n$, so gilt

$$\begin{aligned}
&\langle Qx_i, Qx_j\rangle = x_j^\top Q^\top JQx_i = x_j^\top Jx_i \quad \forall_{i,j} = 1, \ldots, n \\
&\Longrightarrow \text{ mit } X = [x_1, \ldots, x_n] \\
&X^\top Q^\top JQX = X^\top JX
\end{aligned}$$

und da die x_i eine Basis bilden, ist X nichtsingulär $\Longrightarrow$ $Q^\top JQ = J$. □

Wie sehen nun die Elemente von $\mathcal{S}_J$ aus, und was für Eigenschaften hat $\mathcal{S}_J$?
Falls $J = \begin{bmatrix} I_p & \\ & -I_q \end{bmatrix}$ und $p = 0$ oder $q = 0$ ist, so ist $J = -I$ oder $J = I$ und damit $\mathcal{S}_J = \mathcal{O}_n(\mathbb{R})$.

Falls jedoch $p, q \neq 0$, so ist die Situation komplizierter.

Lemma 8.6 *Sei* $J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, *so haben alle* J*–orthogonalen Matrizen die Form*

$$\begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix} \begin{bmatrix} \cosh(\varphi) & \sinh(\varphi) \\ \sinh(\varphi) & \cosh(\varphi) \end{bmatrix}, \quad \textit{mit } \varepsilon_i^2 = 1,\ i = 1, 2$$

wobei $\cosh(\varphi) = \dfrac{e^\varphi + e^{-\varphi}}{2}, \quad \sinh(\varphi) = \dfrac{e^\varphi - e^{-\varphi}}{2}$

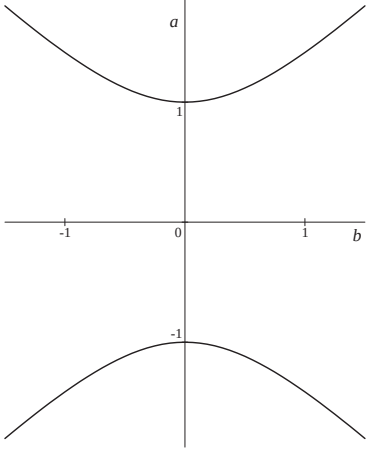
Beweis: Sei $Q = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ und $Q^\top \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} Q = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

$$\begin{aligned}
\Longleftrightarrow \quad & \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} a & b \\ -c & -d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\
\Longleftrightarrow \quad & a^2 - c^2 = 1, \quad ab - cd = 0, \quad b^2 - d^2 = -1 \\
& a^2 = \quad 1 + c^2 \implies a = \pm\sqrt{1 + c^2} \\
& b^2 = -1 + d^2 \implies b = \pm\sqrt{d^2 - 1} \\
\Longleftrightarrow \quad & \pm\sqrt{1 + c^2}\sqrt{d^2 - 1} - cd = 0 \\
\Longrightarrow \quad & (1 + c^2)(d^2 - 1) = c^2 d^2 \\
& d^2 + c^2 d^2 - c^2 - 1 = c^2 d^2 \\
\Longrightarrow \quad & d^2 - c^2 = 1
\end{aligned}$$

Mit $a^2 - c^2 = 1 \implies d^2 - a^2 = 0 \implies d = \pm a$ und mit $ab = cd \implies c = \pm b$.

Also gilt $Q = \begin{bmatrix} a & b \\ \varepsilon b & \varepsilon a \end{bmatrix}$ mit $\varepsilon \in \{\pm 1\}$ und $a^2 = 1 + b^2$.

Die Gleichung $a = \pm\sqrt{1 + b^2}$ definiert eine Hyperbel.



$$\Longrightarrow Q = \begin{bmatrix} 1 & \\ & \varepsilon \end{bmatrix} \begin{bmatrix} a & b \\ b & a \end{bmatrix} \text{ mit } \varepsilon \in \{\pm 1\} \text{ und } a^2 - b^2 = 1.$$

Diese Beziehung gilt für $a = \pm\cosh(\varphi)$, $b = \sinh(\varphi)$ mit $\varphi = \pm\text{arcosh}(a) = \ln(a \pm \sqrt{a^2 - 1})$ und wir haben mit $\varepsilon_1 = \pm 1$ und $\varepsilon_2 = \varepsilon_1\varepsilon$:

$$\begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix} \begin{bmatrix} \cosh(\varphi) & \sinh(\varphi) \\ \sinh(\varphi) & \cosh(\varphi) \end{bmatrix}.$$

□

Da wir schon wissen, daß alle Matrizen $Q \in \mathbb{R}^{2,2}$, für die $Q^\top I Q = I$ ist, die Form

$$\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} 1 & \\ & \varepsilon \end{bmatrix} \quad \text{mit} \quad \varepsilon \in \{\pm 1\}$$

haben, können wir jetzt alle Matrizen in $\mathcal{S}_J$ beschreiben.

Satz 8.7 *Sei* $J = \begin{bmatrix} I_p & \\ & -I_q \end{bmatrix} \in \mathbb{R}^{n,n}$ *und* $G \in \mathcal{S}_J = \{G \in \mathbb{R}^{n,n} \mid G^\top J G = J\}$, *so läßt sich* G *als Produkt von Matrizen der Form*

$$H_{ij}(\varphi) = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & \cosh\varphi & & \sinh\varphi & \leftarrow i \\ & & & \ddots & & \\ & & \sinh\varphi & & \cosh\varphi & \leftarrow j \\ & & & & & \ddots \\ & & & & & 1 \end{bmatrix}$$

$$\begin{matrix} & & \uparrow & & \uparrow & \\ & & i & & j & \end{matrix}$$

$$R_{ij}(\varphi) \;=\; \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & \cos\varphi & & \sin\varphi & & \leftarrow i \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & -\sin\varphi & & \cos\varphi & & \leftarrow j \\ & & & & & 1 & \\ & & & & & & \ddots \\ & & & & & & 1 \end{bmatrix}$$

$$\begin{matrix} & & \uparrow & & \uparrow & \\ & & i & & j & \end{matrix}$$

und Vorzeichenmatrizen $D = \operatorname{diag}(d_1, \ldots, d_n)$ *mit* $d_i \in \{\pm 1\}$ *darstellen.*

Beweis: Sei $G^\top J G = J$ und $G = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ mit $A \in \mathbb{R}^{p,p}, D \in \mathbb{R}^{q,q}$.

Nach Teil I, Kap. 10 (QR–Zerlegung) gibt es orthogonale Matrizen $P_1 \in \mathbb{R}^{p,p}, P_2 \in \mathbb{R}^{q,q}$, so daß

$$\tilde{A} = P_1^\top A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & a_{p2} & \cdots & a_{pp} \end{bmatrix}, \quad \tilde{C} = P_2^\top C = \begin{bmatrix} c_{11} & c_{12} & \ldots & c_{1p} \\ 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & c_{q2} & \cdots & c_{qp} \end{bmatrix}$$

und $P^\top = \begin{bmatrix} P_1^\top & 0 \\ 0 & P_2^\top \end{bmatrix} \in \mathcal{S}_J$ und damit auch $P^\top \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{bmatrix} \in \mathcal{S}_J$.

$\Longrightarrow \tilde{A}^\top \tilde{A} - \tilde{C}^\top \tilde{C} = I_p \quad \Longrightarrow \quad a_{11}^2 - c_{11}^2 = 1.$

Multiplikation mit $H = \begin{bmatrix} a_{11} & & c_{11} & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ c_{11} & & a_{11} & \\ & & & 1 \\ & & & \ddots \\ & & & 1 \end{bmatrix}^{-1} \in \mathcal{S}_J$ von links ergibt

$$\begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix} = H P^\top \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \left[\begin{array}{cccc|c} 1 & * & \cdots & * & \\ 0 & \vdots & & \vdots & \\ \vdots & \vdots & & \vdots & * \\ 0 & * & \cdots & * & \\ \hline 0 & * & \cdots & * & \\ \vdots & \vdots & & \vdots & * \\ 0 & * & \cdots & * & \end{array}\right] \in \mathcal{S}_J.$$

$\Longrightarrow \quad \hat{A}^\top \hat{A} - \hat{C}^\top \hat{C} = I, \quad \hat{A}^\top \hat{B} - \hat{C}^\top \hat{D} = 0.$

Für die erste Zeile ergibt sich daraus

$$\hat{A} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & * & \\ 0 & & & \end{array}\right] \quad \text{und} \quad \hat{B} = \left[\begin{array}{ccc} 0 & \cdots & 0 \\ \hline & & \\ & * & \\ & & \end{array}\right].$$

Per Induktion folgt dann die Behauptung. Es ist dabei klar, daß wir jederzeit noch mit Diagonalmatrizen mit $\{\pm 1\}$ in der Diagonale multiplizieren können. □

Beachte: Hyperbolische Rotationen sind keine orthogonalen Matrizen, denn i.a. gilt $H_{ij}^\top I H_{ij} \neq I$. Sie sind nur orthogonal im Sinne des Skalarprodukts $\langle \cdot, \cdot \rangle_J$.

Die Zerlegung von $G \in \mathcal{S}_J$ als Produkt von Rotationen und hyperbolischen Rotationen ist natürlich nicht eindeutig. Allerdings lassen sich die Elemente von $\mathcal{S}_J$ über die Winkel φ_{ij} der Rotationen und die Vorzeichen parametrisieren. Man kann zeigen, daß $\frac{n(n-1)}{2}$ Winkel in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ für Rotationen bzw. Parameter φ_{ij} für hyperbolische Rotationen und n Vorzeichen (sind eigentlich auch Winkel) jedes Element charakterisieren.

Beispiel 8.8 $J = \left[\begin{array}{cc|c} 1 & & \\ & 1 & \\ \hline & & -1 \end{array}\right]$

$$\begin{aligned} G &= \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{3\sqrt{2}}{2} & \frac{3\sqrt{2}}{2} & 2\sqrt{2} \\ -2 & 2 & 3 \end{bmatrix}, \quad H_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 2\sqrt{2} \\ 0 & 2\sqrt{2} & 3 \end{bmatrix}, \\ H_{23}^{-1} &= \frac{1}{9-(2\sqrt{2})^2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -2\sqrt{2} \\ 0 & -2\sqrt{2} & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -2\sqrt{2} \\ 0 & -2\sqrt{2} & 3 \end{bmatrix}, \\ \varphi_{23} &= \operatorname{arcosh}(3), \\ H_{23}^{-1} G &= \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \varphi_{12} = \frac{\pi}{4},\ \varphi_{13} = 0. \end{aligned}$$

Korollar 8.9 *Sei* $J = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$, *so gilt* $\det G = \pm 1$ *für alle* $G \in \mathcal{S}_J$.

Beweis: $\det(G^\top J G) = \det(J) \implies \det(G^2) \cdot \det(J) = \det(J) \implies (\det G)^2 = 1$
$\implies \det(G) = \pm 1$. □

Eine weitere wichtige Matrixgruppe sind die Matrizen in $\mathbb{R}^{n,n}$ mit Determinante 1. Diese Gruppe wird mit $SL_n(\mathbb{R})$ bezeichnet.

Satz 8.10 *Sei* $J = \begin{bmatrix} I_p & \\ & -I_q \end{bmatrix}$, $p+q=n$. *Die folgenden Klassen von Matrizen bilden Untergruppen von* $GL_n(\mathbb{R})$.

(a) $\mathcal{S}_J = \{G \in \mathbb{R}^{n,n} \mid G^\top J G = J\}$

(b) $SL_n(\mathbb{R}) = \{A \in \mathbb{R}^{n,n} \mid \det A = 1\}$

(c) $SL_n(\mathbb{R}) \cap \mathcal{S}_J$

(d) $SL_n(\mathbb{R}) \cap \mathcal{O}_n(\mathbb{R})$

(e) $\mathcal{S}_J \cap \mathcal{O}_n(\mathbb{R})$

Beweis: Übung □

Analoge Betrachtungen lassen sich auch im Komplexen durchführen.

Wir hatten in Teil 2 gesehen, daß orthogonale (unitäre Matrizen) alle Eigenwerte auf dem Einheitskreis haben. Kann man für J-orthogonale Matrizen etwas analoges zeigen?

Betrachte $Q \in \mathcal{S}_J$ mit $J = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$, $p+q=n$.

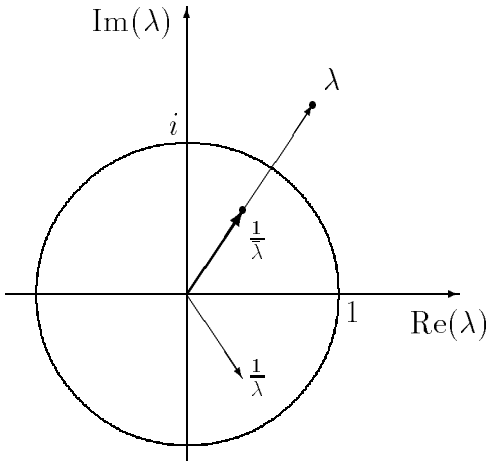
Sei $Qx = \lambda x$ mit $x \neq 0 \implies \underline{Jx} = Q^\top J Q x = \underline{\lambda Q^\top J x}$, also ist $y = (x^\top J^\top)$ Linkseigenvektor von Q und es gilt

$$y^\top = \lambda y^\top Q. \tag{8.11}$$

Da $Q \in \mathcal{S}_J$ nichtsingulär ist, gilt $\lambda \neq 0$ und es folgt

$$y^\top Q = \frac{1}{\lambda} y^\top \quad \text{oder} \quad Q^\top y = \frac{1}{\lambda} y.$$

Da aber Q und $Q^\top$ die gleichen Eigenwerte haben, ergibt sich, daß mit λ auch immer $\frac{1}{\lambda}$ ein Eigenwert ist (im Komplexen $\frac{1}{\overline{\lambda}}$).



$$\lambda = a + ib$$

$$\frac{1}{\overline{\lambda}} = \frac{a+ib}{a^2+b^2}$$

$$\frac{1}{\lambda} = \frac{a-ib}{a^2+b^2}$$

Es folgt, daß für $p+q$ ungerade mindestens ein Eigenwert auf dem Einheitskreis liegen muß, und zwar im reellen Fall bei $+1$ oder -1, denn sonst kann nicht $\lambda = \frac{1}{\lambda}$ sein.

Man kann für Matrizen aus $\mathcal{S}_J$ auch eine Jordanform bestimmen unter Transformation mit Matrizen aus $\mathcal{S}_J$, d.h., man sucht für $Q \in \mathcal{S}_J$ ein $S \in \mathcal{S}_J$, so daß $S^{-1}QS$ in einer Normalform ist, die so etwas wie die Jordan'sche Normalform ist. Diese Form wurde erst kürzlich das erste Mal bestimmt und ist ziemlich schrecklich (siehe meine Webseite, bzw. `http://www.tu-chemnitz.de/sfb393/preprints.html`, Nr. 98-07 und 98-29)

Was kann man nun im J-symmetrischen oder J-schiefsymmetrischen Fall aussagen?
Dazu führen wir zwei neue Matrixmultiplikationen ein.

Definition 8.12 *Sei* $J = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$, $p+q=n$ *und*

$$\mathcal{A}_J = \{A \in \mathbb{R}^{n,n} \mid (AJ)^\top = AJ\}$$

die Menge der J-symmetrischen Matrizen in $\mathbb{R}^{n,n}$ und

$$\mathcal{C}_J = \{A \in \mathbb{R}^{n,n} \mid (AJ)^\top = -AJ\}$$

die Menge der J-schiefsymmetrischen Matrizen in $\mathbb{R}^{n,n}$.
Definiere für $A, B \in \mathcal{A}_J$ die Multiplikation

$$(A,B) = AB + BA \tag{8.13}$$

und für $A, B \in \mathcal{C}_J$ die Multiplikation

$$[A,B] = AB - BA. \tag{8.14}$$

Lemma 8.15

(i) Mit $A, B \in \mathcal{A}_J$ ist auch $(A,B) \in \mathcal{A}_J$.

(ii) Mit $A, B \in \mathcal{C}_J$ ist auch $[A,B] \in \mathcal{C}_J$.

(iii) Für $A \in \mathcal{C}_J$ ist $A^2 \in \mathcal{A}_J$.

Beweis: Beachte $J^\top = J$.

$$\begin{aligned}
(i)\quad A,B \in \mathcal{A}_J &\implies (AJ)^\top = AJ, \quad (BJ)^\top = BJ \\
&\implies (AB+BA)J = ABJ + BAJ \\
&= AJ^\top B^\top + BJ^\top A^\top = AJB^\top + BJA^\top \\
&= J^\top A^\top B^\top + J^\top B^\top A^\top = J(AB+BA)^\top.
\end{aligned}$$

$$\begin{aligned}
(ii)\quad A,B \in \mathcal{C}_J &\implies (AJ)^\top = -AJ, \quad (BJ)^\top = -BJ \\
&\implies (AB-BA)J = ABJ - BAJ \\
&= -AJ^\top B^\top + BJ^\top A^\top = JA^\top B^\top - JB^\top A^\top \\
&= J(AB-BA)^\top
\end{aligned}$$

$$(iii)\quad A^2 J = -AJA^\top = J(A^\top)^2.$$

□

Definition 8.16 *Ein Vektorraum V mit einer bilinearen Operation „Multiplikation“*

$$\begin{aligned} * \;:\; & V \times V \to V \\ & (x,y) \to x * y = xy \end{aligned}$$

heißt nichtassoziative Algebra. *Falls die Operation $*$ assoziativ ist, so heißt die Algebra* assoziativ, *d.h., es gilt $(xy)z = x(yz)$.*

Beispiel 8.17 Die Mengen $\mathbb{R}^{n,n}, \mathbb{C}^{n,n}, \mathbb{F}^{n,n}$ sind assoziative Algebren mit der normalen Matrixmultiplikation. Mit der Multiplikation $[A,B] = AB - BA$ sind sie nichtassoziative Algebren.

Definition 8.18 *Ein Vektorraum L mit einem Produkt $[\cdot,\cdot]$, für das gilt*

$$[x,y] = -[y,x], \qquad \forall x,y \in L, \tag{8.19}$$

$$[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0, \quad \forall x,y,z \in L \quad \textit{(Jacobi-Identität)} \tag{8.20}$$

heißt Lie-Algebra.

Beispiel 8.21

(i) $\mathbb{R}^3$ mit dem $\times$-Produkt (Vektorprodukt) ist eine Lie-Algebra, denn es gilt

$$a \times (b \times c) = (a^\top c)b - (a^\top b)c\,.$$

(ii) Für $J = \begin{bmatrix} I_p & \\ & -I_q \end{bmatrix}$ ist $\mathcal{C}_J = \{A \in \mathbb{R}^{p+q,p+q} \mid AJ + J^\top A^\top = 0\}$ eine Lie-Algebra mit dem *Lie-Produkt* $[A,B] = AB - BA$.

Definition 8.22 *Eine* Derivation *einer nichtassoziativen Algebra $\mathcal{A}$ ist ein linearer Operator d auf $\mathcal{A}$, der die formale Leibniz–Regel für Ableitungen erfüllt.*

$$d(x,y) = (dx)y + x(dy) \quad \forall x,y \in \mathcal{A} \tag{8.23}$$

Beispiel 8.24 Für die Algebra der Polynome in x und ein festes Polynom $p(x)$ ist $d = p(x)\dfrac{d}{dx}$ eine Derivation.

Die Menge der Derivationen einer nichtassoziativen Algebra $\mathcal{A}$ ist ein Untervektorraum von der Menge aller linearen Abbildungen von $\mathcal{A}$ in $\mathcal{A}$.

Das Produkt von Derivationen ist im allgemeinen keine Derivation jedoch das Lie-Produkt $[\cdot,\cdot]$. Denn für $[d_1,d_2]$ gilt:

$$\begin{aligned} (d_1d_2 - d_2d_1)(xy) &= d_1(d_2(xy)) - d_2(d_1(xy)) \\ &= d_1((d_2x)y + x(d_2y)) - d_2((d_1x)y + x(d_1y)) \\ &= (d_1d_2x)y + (d_2x)(d_1y) + (d_1x)(d_2y) + x(d_1d_2y) \\ &\quad -(d_2d_1x)y - (d_1x)(d_2y) - (d_2x)(d_1y) - x(d_2d_1y) \\ &= ((d_1d_2 - d_2d_1)x)\,y + x\,((d_1d_2 - d_2d_1)y)\,. \end{aligned}$$

Korollar 8.25 *Sei $\mathcal{A}$ eine nichtassoziative Algebra und D die Menge der Derivationen von $\mathcal{A}$, so ist D eine Lie–Algebra. Man nennt D die Derivationsalgebra.*

Bevor wir den Zusammenhang zwischen den Lie–Algebren und den Matrixgruppen studieren, wollen wir noch eine weitere klassische Matrixgruppe einführen.

Definition 8.26 *Sei $J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$. Die Menge der J–orthogonalen Matrizen in $\mathbb{R}^{2n,2n}$ heißt* symplektische Gruppe.

$$Sp_{2n}(\mathbb{R}) = \{Q \in \mathbb{R}^{2n,2n} \mid Q^\top JQ = J\}.$$

Lemma 8.27

(a) Sei $K \in \mathbb{R}^{2n,2n}$ schiefsymmetrisch und nichtsingulär, so ist K kongruent zu J aus Definition 8.26.

(b) Sei $K \in \mathbb{R}^{n,n}$ symmetrisch und nichtsingulär, so ist K kongruent zu $\begin{bmatrix} I_p & \\ & -I_q \end{bmatrix}$.

Beweis:

(a) Da K reell schiefsymmetrisch ist, gibt es Q orthogonal, so daß

$$Q^\top KQ = \begin{bmatrix} K_{11} & & \\ & \ddots & \\ & & K_{ss} \end{bmatrix}$$

mit K_{ii} reell-schiefsymmetrisch 1×1 oder 2×2 (je nachdem, ob die Eigenwerte reell oder komplex konjugierte Paare sind). Da aber alle Eigenwerte von K auf der imaginären Achse liegen und K nichtsingulär ist, so folgt, daß alle K_{ii} die Form $\begin{bmatrix} 0 & a_i \\ -a_i & 0 \end{bmatrix}$ haben und Paare komplex konjugierter Eigenwerte haben, also $s = n$

Sei $D = \operatorname{diag}(D_{11}, \ldots, D_{nn})$ mit $D_{ii} = \begin{bmatrix} \frac{1}{\sqrt{a_i}} & 0 \\ 0 & \frac{\operatorname{sgn}(a_i)}{\sqrt{a_i}} \end{bmatrix}$, so folgt

$$D^\top Q^\top KQD = \begin{bmatrix} 0 & 1 & & & \\ -1 & 0 & & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & -1 & 0 \end{bmatrix}.$$

Mit $P = [e_1, e_3, \ldots, e_{2n-1}, e_2, e_4, \ldots, e_{2n}]$ ist dann $P^\top D^\top Q^\top KQDP = J$.

(b) klar.

□

Damit können wir uns auf die Fälle

$$J = \begin{bmatrix} I_p & \\ & -I_q \end{bmatrix} \text{ oder } J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$$

beschränken.

Lemma 8.28 *Sei* $J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$. *Die Menge der* J*–symmetrischen Matrizen*

$$\mathcal{A}_J(\mathbb{R}) = \{A \in \mathbb{R}^{n,n} \mid (AJ)^\top = AJ\}$$

ist eine Lie–Algebra mit dem Produkt $[\cdot,\cdot]$. *Alle Matrizen in* $\mathcal{A}_J(\mathbb{R})$ *haben die Form*

$$A = \begin{bmatrix} F & G \\ H & -F^\top \end{bmatrix} \text{ mit } H = H^\top, G = G^\top.$$

Diese Matrizen heißen Hamiltonische Matrizen.

Beweis:

$$A = \begin{bmatrix} F & G \\ H & L \end{bmatrix} \in \mathcal{A}_J(\mathbb{R})$$

$$\Longrightarrow \quad \begin{bmatrix} -G & F \\ -L & H \end{bmatrix} = (AJ)^\top = AJ = \begin{bmatrix} -G^\top & -L^\top \\ F^\top & H^\top \end{bmatrix}$$

$$\Longrightarrow \quad G = G^\top,\ H = H^\top,\ L = -F^\top.$$

$A, B \in \mathcal{A}_J(\mathbb{R})$

$$\begin{aligned} \Longrightarrow (AB - BA)J &= ABJ - BAJ \\ &= AJ^\top B^\top - BJ^\top A^\top \\ &\overset{(J^\top = -J)}{=} -AJB^\top + BJA^\top \\ &= -J^\top A^\top B^\top + J^\top B^\top A^\top \\ &= J^\top (B^\top A^\top - A^\top B^\top) \end{aligned}$$

□

Lemma 8.29 *Sei* $X \in GL_n(\mathbb{R})$. *Falls* AX *symmetrisch (schiefsymmetrisch) für alle symmetrischen (schiefsymmetrischen) Matrizen* $A \in \mathbb{R}^{n,n}$, *so gilt*

$$X = \alpha I_n \quad \textit{mit} \quad \alpha \in \mathbb{R} \setminus \{0\}.$$

Beweis: Es folgt mit $A = I_n$, daß X symmetrisch ist. Mit $A = e_k e_j^\top + e_j e_k^\top$ folgt

$$X(e_k e_j^\top + e_j e_k^\top) = (e_k e_j^\top + e_j e_k^\top)X, \quad \forall j,k \in \{1,\dots,n\}.$$

$$\begin{aligned}
\Longrightarrow\ & x_{rk} e_j^\top e_l + x_{rj} e_k^\top e_l = e_r^\top e_k x_{jl} + e_r^\top e_j x_{kl} \qquad \forall j,k,r,l \in \{1,\dots,n\} \\
\Longrightarrow\ & \text{für } k \neq r,\ l = j = k \text{ folgt } 2x_{rk} = 0 \text{ und} \\
& \text{für } j = r, l = k \text{ folgt } x_{kk} = x_{jj} \implies x = \alpha I_n
\end{aligned}$$

Im schiefsymmetrischen Fall analoger Beweis (Übung). □

Satz 8.30 *Sei* $J = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$ *oder* $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$ *und sei* $K = \mathcal{A}_J(\mathbb{R})$ *oder* $K = \mathcal{C}_J(\mathbb{R})$. *Dann gilt* $T^{-1}MT \in K$ *für alle* $M \in K$ *genau dann, wenn*

$$T^\top J T = \alpha J \quad \textit{für} \quad \alpha \in \mathbb{R}\backslash\{0\}$$

(wobei natürlich das passende J *zu wählen ist).*

Beweis: Falls $T^\top J T = \alpha J$ (oder $T^\top J = \alpha J T^{-1}$) ist, so gilt für $M \in K$, daß

$$\begin{aligned}
J(T^{-1}MT) &= \frac{1}{\alpha} T^\top J M T = \pm \frac{1}{\alpha} T^\top M^\top J^\top T = \pm \frac{1}{\alpha} T^\top M^\top \alpha T^{-\top} J^\top \\
&= \pm (T^{-1}MT)^\top J^\top.
\end{aligned}$$

Umgekehrt, falls für $M \in K$ auch $T^{-1}MT \in K$, so folgt

$$\begin{aligned}
J(T^{-1}MT) &= \pm (T^{-1}MT)^\top J^\top = \pm T^\top M^\top T^{-\top} J^\top = \pm T^\top M^\top J^\top J T^{-\top} J^\top \\
&= T^\top J M J T^{-\top} J^\top, \qquad \Longrightarrow \\
JM(TJT^\top J^\top) &= (JTJ^\top T^\top) JM
\end{aligned}$$

und damit ist $JM(TJT^\top J^\top)$ symmetrisch.
Wenn also $T^{-1}MT \in K\ \forall M \in K$, so ist $TJT^\top J^\top$ symmetrisch für alle symmetrischen (schiefsymmetrischen) $A = JM$, also folgt nach Lemma 8.29

$$JTJ^\top T^\top = \alpha I \quad \Longrightarrow \quad TJ^\top T^\top = \alpha J^\top \text{ oder } TJT^\top = \alpha J \implies JT^\top = \alpha T^{-1} J$$

$$\begin{aligned}
\Longrightarrow\ & JT^\top J^\top T &&= \alpha T^{-1} J J^\top T = \alpha I \\
\Longrightarrow\ & T^\top J^\top T &&= \alpha J^\top \\
\Longrightarrow\ & T^\top J T = && \alpha J.
\end{aligned}$$

□

Wir erhalten also, daß Ähnlichkeitstransformation mit Elementen der Gruppe die entsprechende Algebra erhalten und eigentlich nur diese (bis auf eine Konstante). Man kann nun für Matrizen in $\mathcal{A}_J, C_J$ mit $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$ eine Schurform mit Matrizen in $\mathcal{O}_{2n}(\mathbb{R}) \cap Sp_{2n}(\mathbb{R})$ herleiten.

Satz 8.31 *Sei* $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$.

(i) Sei $H \in \mathcal{A}_J = \{A \in \mathbb{R}^{2n,2n} \mid (AJ)^\top = AJ\}$ *und* H *habe keine Eigenwerte mit Realteil* 0*, so gibt es* $S \in \mathcal{O}_{2n}(\mathbb{R}) \cap Sp_{2n}(\mathbb{R})$*, so daß*

$$S^{-1}HS = S^\top HS = \begin{bmatrix} F & G \\ 0 & -F^\top \end{bmatrix}$$

mit F *in reeller Schurform, d.h.* $F = \begin{bmatrix} F_{11} & \cdots & F_{1k} \\ & \ddots & \vdots \\ & & F_{kk} \end{bmatrix}$ *mit* $F_{ii} \quad 1 \times 1$ *oder* 2×2.

(ii) Sei $H \in \mathcal{C}_J = \{A \in \mathbb{R}^{2n,2n} \mid (AJ)^\top = -AJ\}$*, so gibt es* $S \in \mathcal{O}_{2n}(\mathbb{R}) \cap Sp_{2n}(\mathbb{R})$*, so daß*

$$S^{-1}HS = S^\top HS = \begin{bmatrix} F & G \\ 0 & F^\top \end{bmatrix}$$

mit F *in reeller Schurform.*

Beweis: Wir haben bereits die Schurform für allgemeine Matrizen betrachtet. Der Beweis geht analog und soll hier nicht gemacht werden. □

Bemerkung: Man kann diesen Satz als notwendige und hinreichende Bedingung formulieren und dabei Eigenwerte auf der imaginären Achse zulassen.
Siehe Webseite.

Kapitel 9

Grundlagen der Gruppentheorie

Wir wollen nun noch einige wichtige Grundlagen der Gruppentheorie behandeln.

Definition 9.1 *Eine nichtleere Menge H mit einer assoziativen inneren Verknüpfung $\circ$ heißt* Halbgruppe.

Beispiel 9.2 $(\mathbb{Z},+),(\mathbb{Z},\cdot)$ sind Halbgruppen.

Satz 9.3 *Sei $(H,\circ)$ eine Halbgruppe $a \in H, m,n \in \mathbb{N}$, so gilt*

$$a^n \circ a^m = a^{n+m},\ (a^n)^m = a^{n\cdot m}.$$

(Beachte Potenzen sind wie üblich definiert durch $a^1 = a, a^{n+1} = a \circ a^n$.)

Beweis: Per Induktion. □

Definition 9.4 *Eine Halbgruppe $(H,\circ)$ heißt* kommutativ *(*abelsch*), falls $a \circ b = b \circ a$ für alle $a,b \in H$. Ein Element e heißt linksneutrales (rechtsneutrales) Element, falls $e \circ a = a$ $(a \circ e = a)$. Ein Element, welches sowohl rechts- als auch linksneutral ist, heißt* neutral.

Beispiel 9.5 Sei $H = \left\{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}, a,b \in \mathbb{R} \right\}$, so ist H eine Halbgruppe bezüglich der Matrizenmultiplikation. $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ ist linksneutral aber nicht rechtsneutral.

Definition 9.6 *Es sei $(H,\circ)$ eine Halbgruppe mit neutralem Element e. Ein Element $a \in H$ heißt* linksinvertierbar *(*rechtsinvertierbar*), wenn es $b \in H$ gibt mit $b \circ a = e (a \circ b = e)$. b heißt dann* Linksinverses *(*Rechtsinverses*) zu a.*

Satz 9.7 *Ist $(H,\circ)$ eine Halbgruppe mit neutralem Element e und H^* die Menge der (links- und rechts-) invertierbaren Elemente aus H, so ist $(H^*,\circ)$ eine Gruppe.*

Beweis: $e \circ e = e \implies e \in H^*$ und $e^{-1} = e$. $a^{-1} \circ a = a \circ a^{-1} = e \implies a^{-1}$ ist invertierbar mit Inversem a. Da das Inverse eindeutig ist, folgt $a = (a^{-1})^{-1}$. □

Beispiel 9.8

(a) $\mathbb{R}^{n,n}$ ist eine Halbgruppe mit neutralem Element $e = I_n$. $(\mathbb{R}^{n,n})^* = GL_n(\mathbb{R})$.

(b) Die Menge $E(X)$ aller Abbildungen einer Menge X in sich mit Komposition $\circ$ als Verknüpfung ist eine Halbgruppe mit $e = Id_X$.
Die bijektiven Abbildungen von $X \to X$ sind die Permutationen und bilden gerade $E(X)^*$.
Die Gruppe der Permutationen von X heißt *symmetrische Gruppe* $S(X)$.

Für endliche Gruppen können wir die Multiplikation (Verknüpfung) mittels einer Tafel angeben

$\circ$	x_1	$\cdots$	x_n
x_1	$x_1 \circ x_1$	$\cdots$	$x_1 \circ x_n$
$\vdots$			
x_n	$x_n \circ x_1$	$\cdots$	$x_n \circ x_n$

Satz 9.9 *Die Verknüpfungstafel einer endlichen Halbgruppe G ist die Tafel einer Gruppe genau dann, wenn in jeder Zeile und Spalte jedes Element von G höchstens einmal vorkommt.*

Beweis: Sei $G = \{x_1, \ldots, x_n\}$. In der i-ten Zeile sind die Elemente $x_i \circ x_j$. Die Behauptung ist äquivalent zu

$$x_i \circ x_j = x_i \circ x_k \iff j = k.$$

Kürzungsregel in Gruppen $\implies x_j = x_k$.
Analoges gilt für Spalten. □

Beispiel 9.10

(i)

$\bullet$	e	a	b
e	e	a	b
a	a		
b	b		

In der $(2,2)$ Position kann nur b oder e stehen. Wäre es e, so müßte darunter b stehen, dann wären aber zwei b's in einer Zeile, also gibt es nur eine Möglichkeit bei 3 Elementen.

$\bullet$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

(ii) Klein'sche Vierergruppe $G = \{e, a, b, c\}$, $a^2 = b^2 = e$.

	e	a	b	c
e	e	a	b	c
a	a	e	$*$	
b	b		e	
c	c			

$*$ muß c sein, also folgt

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Es gibt also nur eine Gruppe mit 4 Elementen und $a^2 = b^2 = e$. Man muß allerdings noch die Assoziativität nachweisen.

Wir hatten schon Homomorphismen zwischen Vektorräumen und Ringen kennengelernt, für Gruppen definieren wir analog:

Definition 9.11 *Es seien $(G, \circ), (H, *)$ Gruppen.*
Eine Abbildung $\varphi : G \to H$ heißt Gruppenhomomorphismus, *wenn für alle $a, b \in G$ gilt*

$$\varphi(a \circ b) = \varphi(a) * \varphi(b).$$

Beispiel 9.12 $\varphi(x) = e^x$ ist ein Gruppenhomomorphismus von $(\mathbb{R}, +)$ in $(R\backslash\{0\}, \cdot)$, da $e^{x+y} = e^x \cdot e^y$.

Wiederholung: Sei $\varphi : G \to H$ Gruppenhomomorphismus.

Monomorphismus	φ injektiv
Epimorphismus	φ surjektiv
Isomorphismus	φ bijektiv
Endomorphismus	$\varphi : G \to G$
Automorphismus	$\varphi : G \to G$ bijektiv .

Satz 9.13 *Seien G, H Gruppen, $\varphi : G \to H$ Homomorphismus, e neutrales Element von G. Dann folgt*

(a) $\varphi(e)$ *ist neutrales Element von* H,

(b) $\varphi(a^{-1}) = [\varphi(a)]^{-1} \quad \forall a \in G$,

(c) $\varphi(a^n) = (\varphi(a))^n \qquad \forall a \in G,\ n \in \mathbb{Z}$.

Beweis: Übung □

Wiederholung: Seien G, H Gruppen mit neutralen Elementen e_g, e_h

$$\begin{aligned} \text{Bild}\,\varphi &= \{\varphi(x) \in H \mid x \in G\}, \\ \text{Kern}\,\varphi &= \{x \in G \mid \varphi(x) = e_h\} \end{aligned}$$

Satz 9.14

(a) Ein Gruppenhomomorphismus $\varphi : G \to H$ ist injektiv genau dann, wenn

$$\text{Kern}\,\varphi = \{e_g\}.$$

(b) Sind $\varphi : G \to H$, $\psi : H \to K$ Gruppenhomomorphismen, so ist auch $\psi \circ \varphi : G \to K$ ein Gruppenhomomorphismus.

(c) Mit φ ist auch $\varphi^{-1} : H \to G$ ein Isomorphismus.

(d) Sind $\varphi : G \to H$, $\psi : H \to K$ Isomorphismen, so auch $\psi \circ \varphi$.

(e) Id_G ist ein Isomorphismus.

Beweis: Übung □

Satz 9.14 zeigt, daß Isomorphie auf einer Menge von Gruppen eine Äquivalenzrelation definiert.

Definition 9.15 *Die Menge* $\text{Aut}\,G = \{\varphi : G \to G \mid \varphi \text{ ist Automorphismus}\}$ *(die natürlich wieder eine Gruppe ist) heißt* Automorphismengruppe *von G.*

Beispiel 9.16 Sei G eine Gruppe und $x \in G$. Sei

$$\begin{aligned} \varphi_x : \; G &\to G \\ y &\mapsto xyx^{-1}. \end{aligned}$$

Da $\varphi_x(yz) = xyzx^{-1} = xyx^{-1}xzx^{-1} = \varphi_x(y)\varphi_x(z)$, so ist φ_x ein Homomorphismus und da $\varphi_x \circ \varphi_{x^{-1}} = \varphi_{x^{-1}} \circ \varphi_x = Id_G$, so ist φ_x sogar ein Automorphismus. D.h., Ähnlichkeitstransformationen in Matrixgruppen sind Automorphismen.

Definition 9.17 *Sei G eine Gruppe. Ein Element $\varphi \in \text{Aut}G$ heißt* innerer Automorphismus, *wenn es $x \in G$ gibt mit $\varphi = \varphi_x$.*
Elemente $a, b \in G$ heißen konjugiert, *wenn es $x \in G$ gibt mit $\varphi_x(b) = xbx^{-1} = a$.*

In kommutativen Gruppen ist nur die Identität ein innerer Automorphismus. Die Abbildung

$$\begin{aligned} \Phi : \; G &\to \text{Aut}\,G \\ x &\mapsto \varphi_x \end{aligned} \tag{9.18}$$

ist ein Homomorphismus, denn

$$\begin{aligned} \varphi_{xy}(z) &= xyz(xy)^{-1} = xyzy^{-1}x^{-1} \\ &= (\varphi_x \circ \varphi_y)(z) \end{aligned}$$

$$\Longrightarrow \quad \Phi(xy) = \Phi(x) \circ \Phi(y).$$

Definition 9.19 *Für die Abbildung* Φ *aus (9.18) heißt* KernΦ *das* Zentrum von G, *bezeichnet mit* $Z(G)$.

Da

$$\begin{aligned} \text{Kern } \Phi &= \{x \in G \mid \varphi_x = Id_G\} \\ &= \{x \in G \mid xyx^{-1} = y \ \ \forall y \in G\} \\ &= \{x \in G \mid xy = yx \ \ \forall y \in G\} \end{aligned}$$

mißt $Z(G)$ den Grad der Kommutativität von G, denn G ist kommutativ genau dann, wenn $Z(G) = G$.

Beispiel 9.20 Betrachte die Gruppe

$$G = GL_n(\mathbb{R}),$$

so ist $Z(G) = \{\alpha I_n \mid \alpha \in \mathbb{R} \setminus \{0\}\}$. Verwende die Matrizen $I + E_{ij} \in G$ für $i \neq j$, dann folgt aus $X(I + E_{ij}) = (I + E_{ij})X \quad \Longrightarrow \quad XE_{ij} = E_{ij}X$. So folgt $x_{ij} = 0 \ \forall i \neq j$ und $x_{ii} = x_{jj}$.

Satz 9.21 *Sind* X, Y *nichtleere Mengen von gleicher Mächtigkeit, so sind die symmetrischen Gruppen* $S(X)$ *und* $S(Y)$ *isomorph.*

Beweis: Da X, Y gleich viele Elemente haben, so gibt es eine bijektive Abbildung $f : X \to Y$. Zu diesem $f : X \to Y$ betrachte

$$\begin{aligned} \varphi : \ & S(X) \to S(Y) \\ & g \mapsto f \circ g \circ f^{-1} \end{aligned}$$

φ ist auch bijektiv mit $\varphi^{-1} = f^{-1} \circ h \circ f$ und man sieht sofort, daß φ Homomorphismus ist. □

Satz 9.22 (Cayley)
Jede Gruppe G *ist isomorph zu einer Gruppe von Permutationen von* G.

Beweis: Für $g \in G$ sei $L(g) \in S(G)$ definiert durch

$$L(g) : x \mapsto gx.$$

Da G assoziativ ist, so folgt

$$L(gh) = L(g)L(h) \ \ \forall g, h \in G.$$

Sei $\mathcal{L}(G) = \{L(g) \mid g \in G\}$, so ist $\mathcal{L}(G)$ eine Halbgruppe und da $Id = L(e) \in \mathcal{L}(G)$, so gibt es ein neutrales Element und jede Permutation $L(g)$ hat als Inverses $L(g^{-1})$.
Damit ist $\mathcal{L}(G)$ eine Gruppe von Permutationen und die Abbildung

$$\begin{aligned} \Lambda : \ & G \to \mathcal{L}(G) \\ & g \mapsto L(g) \end{aligned}$$

ist ein Epimorphismus (surjektiv). Sei $g \in \operatorname{Kern} \Lambda$ mit $\Lambda(g) = Id$, so folgt

$$\Lambda(g)e = \Lambda(g) = e \Longrightarrow g = e.$$

Also ist Λ ein Isomorphismus. □

Korollar 9.23 *Jede Gruppe der Ordnung n ist isomorph zur Gruppe der Permutationen vom Grad n.*

Für Teilmengen $A, B \subset G$ einer Gruppe G definieren wir

$$\begin{aligned} A \cdot B &:= \{a \cdot b \mid a \in A,\ b \in B\} \\ A^{-1} &:= \{a^{-1} \mid a \in A\} \end{aligned}$$

Satz 9.24 *Sei G eine Gruppe und $U \subseteq G$ eine nichtleere Teilmenge.*

(a) Dann sind die folgenden Aussagen äquivalent:

(i) U ist Untergruppe von G,
(ii) $x, y \in U \Longrightarrow x \cdot y \in U, x^{-1} \in U$,
(iii) $UU \subset U$ und $U^{-1} \subset U$,
(iv) $UU = U$ und $U^{-1} = U$
(v) $x, y \in U \Longrightarrow xy^{-1} \in U$
(vi) $UU^{-1} \subset U$
(vii) $UU^{-1} = U$.

(b) Falls U eine endliche Menge ist, so ist U Untergruppe genau dann, wenn $UU \subset U$.

Beweis: Übung. □

Beispiel 9.25

(a) Sei $G = \{e, a, b, c\}$ die Klein'sche Vierergruppe, dann sind

$$U_1 = \{e, a\}, U_2 = \{e, b\}, U_3 = \{e, c\}$$

Untergruppen von G.

(b) Sei in $S(\mathbb{N})$ für $n \in \mathbb{N}$

$$S'_n = \{f \in S(\mathbb{N}) \mid f(i) = i\ \forall i \in \mathbb{N} \setminus \{1, 2, \ldots, n\}\}$$

so ist S'_n Untergruppe von $S(\mathbb{N})$ und für $n < m$ ist S'_n Untergruppe von S'_m. Dies sieht man sofort, denn $S'_n = S_n$ ist die symmetrische Gruppe von $\{1, \ldots, n\}$.

(c) Eine Realisierung der Klein'schen Vierergruppe ist

$$V_4 = \left\{ e = id \approx I_4,\ a = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \approx \begin{bmatrix} 0 & 1 & & \\ 1 & 0 & & \\ & & 0 & 1 \\ & & 1 & 0 \end{bmatrix}, \right.$$

$$b = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} \approx \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

$$\left. c = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} \approx \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \right\}.$$

Satz 9.26 *Seien G, H Gruppen, $U \subset G$, $V \subset H$ Untergruppen und $\varphi : G \to H$ ein Homomorphismus. Dann gilt*

(i) Bild $(\varphi(U))$ *ist Untergruppe von H.*

(ii) $\varphi^{-1}(V)$ *ist Untergruppe von G.*

Beweis:

(i) $\varphi(U)\varphi(U)^{-1} = \varphi(U)\varphi(U^{-1}) = \varphi(UU^{-1}) = \varphi(U)$, denn $UU^{-1} = U$.

(ii) $\varphi^{-1}(V) \neq \emptyset$, da $\varphi(e) \in V$.
$x, y \in \varphi^{-1}(V) \Rightarrow \varphi(x), \varphi(y) \in V \Rightarrow \varphi(xy^{-1}) = \varphi(x)\varphi(y)^{-1} \Rightarrow xy^{-1} \in \varphi^{-1}(V)$. □

Korollar 9.27 *Für jede Untergruppe U einer Gruppe G und für alle $g \in G$ ist gUg^{-1} auch Untergruppe.*

Beweis: $g\,x\,g^{-1}$ ist innerer Homomorphismus. □

Definition 9.28 *Sei G Gruppe und $U \subset G$ Untergruppe. Die Untergruppen gUg^{-1}, für $g \in G$ heißen die zu U* konjugierten Untergruppen. *Zwei Untergruppen U, V von G heißen* konjugiert, *falls es $g \in G$ gibt mit*

$$U = gVg^{-1}.$$

Konjugation ist eine Äquivalenzrelation auf der Menge der Untergruppen von G.

Satz 9.29 *Es seien U, V Untergruppen von G. UV ist Untergruppe genau dann, wenn $UV = VU$.*

Beweis: Sei UV Untergruppe,

$$\Longrightarrow UV = (UV)^{-1} = V^{-1}U^{-1} = VU.$$

Gilt $UV = VU$, so haben wir $(UV)(UV) = U(VU)V = U(UV)V = (UU)(VV) = UV$ und $(UV)^{-1} = V^{-1}U^{-1} = VU = UV$. □

Definition 9.30 *Sei G eine Gruppe und $S \subset G$. Die Gruppe*

$$\langle S\rangle = \bigcap\{U \mid U \text{ ist Untergruppe von } G \text{ mit } S \subset U\}$$

heißt die von S erzeugte Untergruppe von G.
Ist $G = \langle S\rangle$, so heißt G von S erzeugt und S heißt Erzeugendensystem von G. G heißt endlich erzeugt, falls $G = \langle a_1, \ldots, a_n\rangle$ für $\{a_1, \ldots, a_n\} \subset G$.

Satz 9.31 *Sei G eine Gruppe und $S \subset G$ nicht leer. Dann ist $\langle S\rangle$ die Menge aller endlichen Produkte von $S \cup S^{-1}$.*

Beweis: Falls U eine Untergruppe von G ist, und $S \subset U \implies SS \subset U$ und $S^{-1} \subset U$,

$$\Longrightarrow \quad \Sigma = \{x_1 \cdots x_n \mid x_i \in S \cup S^{-1}, n \in \mathbb{N}\} \subset U$$

und $S \subset \Sigma \subset \langle S\rangle$. Falls $x = x_1 \ldots x_n, y = y_1 \ldots y_n \in \Sigma$, so ist auch $xy^{-1} \in \Sigma \Longrightarrow \Sigma$ ist Untergruppe und $S \subset \Sigma \Longrightarrow \langle S\rangle \subset \Sigma \implies \langle S\rangle = \Sigma$. □

Definition 9.32 *Eine Gruppe G heißt zyklisch, wenn es $g \in G$ gibt, so daß $G = \langle g\rangle$.*

Es ist klar, daß $\langle g\rangle = \{g^n \mid n \in \mathbb{Z}\}$, und da $g^n g^m = g^{n+m} = g^m g^n$, so ist jede zyklische Gruppe kommutativ.

Beispiel 9.33

(i) $(\mathbb{Z}, +) = \{m1 \mid m \in \mathbb{Z}\}$ ist zyklisch und von 1 erzeugt.

(ii) $(\mathbb{Z}_n, +)$, die Gruppe der Restklassen modulo n ist zyklisch und wird von der Restklasse 1 erzeugt.

Beachte $(\mathbb{Z}, +)$ ist unendlich und $(\mathbb{Z}_n, +)$ ist endlich.

Satz 9.34 *Jede Untergruppe einer zyklischen Gruppe ist zyklisch.*

Beweis: $G = \langle g \rangle$ und $U \subset G$ Untergruppe und sei $U \neq \{e\}$. ($\{e\}$ ist natürlich zyklisch).

Sei $x \in U, x \neq e \implies \exists\, m \in \mathbb{Z}, m \neq 0$ und $x = g^m$ und mit x ist auch $x^{-1} = g^{-m} \in U$.

Entweder m oder $-m$ ist positiv,

$$\implies \quad K = \{k \in \mathbb{N} \mid g^k \in U\} \neq \emptyset.$$

Sei $t = \min(K)$, so gilt falls $g^l \in U$, daß $l = qt + r$ mit $0 \leq r < t$, $q, r \in \mathbb{Z}$. Es gilt

$$g^r = g^{l-qt} = g^l (g^t)^{-q} \in U.$$

Falls $r \neq 0$, so folgt $r \in K$ und dies ist ein Widerspruch, da $t = \min\{K\}$ und $r < t$. $\implies r = 0, l = qt$ und $g^l = (g^t)^q \implies U \subset \langle g^t \rangle$. Da $\langle g^t \rangle \subset U$, folgt $U = \langle g^t \rangle$. □

Definition 9.35 *Die Ordnung der von $a \in G$ erzeugten Untergruppe $\langle a \rangle$ heißt die* Ordnung von a. *Bezeichnung* ord $(a) = |\langle a \rangle|$.

Satz 9.36

(a) Die Ordnung von $a \in G$ ist unendlich oder gleich der kleinsten positiven Zahl s mit $a^s = e$.

(b) Hat $a \in G$ die endliche Ordnung t, dann ist $\langle a \rangle = \{e, a, \ldots, a^{t-1}\}$.

Beweis: Betrachte $\langle a \rangle = \{a^n \mid n \in \mathbb{Z}\}$, so sind entweder alle Potenzen verschieden und damit $|\langle a \rangle| = \infty$ oder $a^i = a^j$ für ein $i > j \implies a^l = e$ für $l = i - j$. Sei t die kleinste Zahl für die $a^t = e$, dann ist $\langle a \rangle = \{e, a, \ldots, a^{t-1}\}$. □

Beispiel 9.37 Sei $Z = \{z \in \mathbb{C} \mid z^n = 1\} = \{e^{i\frac{2\pi j}{n}} \mid j = 0, \ldots, n-1\}$. So ist $Z = \langle e^{i\frac{2\pi}{n}} \rangle$.

Korollar 9.38 *Sei $a \in G$ und* ord $(a) = t$, *so gilt $a^k = e \iff k \in t\mathbb{Z}$.*

Beweis: $k = t \cdot r \implies a^k = (a^t)^r = e$.

Für die Umkehrung sei $W = \{k \in \mathbb{Z}, a^k = e\}$. W ist Untergruppe von $\mathbb{Z}$ und nach Satz 9.34 folgt, daß jede Untergruppe von $\mathbb{Z}$ die Form $t\mathbb{Z}$ hat, wobei $t = 0$ oder t ist die kleinste in der Untergruppe vorkommende Zahl aus $\mathbb{N}$. $\implies$ ord $(a) = t$, also $W = t\mathbb{Z}$. □

Satz 9.39 *Sei $a \in G$ mit* ord $(a) = n$. *Dann gilt* ord $(a^m) = \frac{n}{(m,n)}$ $\forall m \in \mathbb{Z}$, $(m, n) =$ ggT *von m und n.*

Beweis: Sei $d = (m, n)$, und sei $t =$ord(a^m), $m = d\tilde{m}, n = d\tilde{n}$. Dann sind $\tilde{m}, \tilde{n}$ teilerfremd. $e = (a^m)^t = a^{mt} \overset{\text{Korollar 9.38}}{\implies} mt = nr$ mit $r \in \mathbb{Z}$. Kürzen von d ergibt $\tilde{m}t = \tilde{n}r$, also $\tilde{n} | \tilde{m}t \implies \tilde{n} | t \implies \tilde{n} \leq t$. Aber $(a^m)^{\tilde{n}} = a^{\tilde{m}d\tilde{n}} = (e^n)^{\tilde{m}} = e^{\tilde{m}} = e \implies t \leq \tilde{n} \Rightarrow t = \tilde{n} = \frac{n}{(m,n)}$. □

Korollar 9.40 *Sei $\langle a \rangle$ zyklisch mit* ord $(a) = n$. *Dann gilt*

$$\langle a \rangle = \langle a^m \rangle \quad \Longleftrightarrow \quad m, n \text{ teilerfremd.}$$

Beweis: $\langle a^m \rangle = \langle a \rangle \Longleftrightarrow$ ord (a) = ord $(a^m) \overset{\text{Satz 9.39}}{\Longleftrightarrow} (m, n) = 1$. □

Satz 9.41 *Ist G zyklisch von der Ordnung n, so gibt es zu jedem Teiler d von n genau eine Untergruppe der Ordnung d.*

Beweis: Sei $G = \langle a \rangle$, ord$(a) = n$ und $n = dm$. Nach Satz 9.39 gilt ord$(a^m) = \dfrac{dm}{(m,n)} = d$.

Also ist $U = \langle a^m \rangle$ Untergruppe der Ordnung d.

Sei V eine weitere Untergruppe der gleichen Ordnung. $\Longrightarrow$ $V = \langle a^k \rangle$ und nach Satz 9.39

$$d = \text{ord } (a^k) = \frac{dm}{(n,k)} \Longrightarrow (n,k) = m \Longrightarrow m|k.$$

Das gleiche Argument umgedreht ergibt $k|m$, also

$$V = \langle a^k \rangle \subset \langle a^m \rangle \subset \langle a^k \rangle.$$

□

Beispiel 9.42 Sei $E = \mathbb{R}^2$ und $n \in \mathbb{N}$.
Sei $\delta \in S(E)$ die Drehung im Nullpunkt um den Winkel $\dfrac{2\pi}{n}$ und σ die Spiegelung an der y–Achse. Die Gruppe $\langle \sigma, \delta \rangle$ heißt die *Diedergruppe* D_n.

Da $\delta^n = Id, \sigma^2 = Id$ und $\delta s \delta = s$, so folgt mit Satz 9.31, daß jedes Element von D_n die Form $\sigma^i \delta^j$ hat,

$$\Longrightarrow \quad D_n = \{Id, \delta, \delta^2, \ldots, \delta^{n-1}, \sigma, \sigma\delta, \ldots, \sigma\delta^{n-1}\}.$$

D_n hat damit die Ordnung $2n$.

Definition 9.43 *Sei G eine Gruppe und $U \subset G$ Untergruppe. Die Relation $R_U \subset G \times G$ wird definiert durch $(x, y) \in R_U \Longleftrightarrow xy^{-1} \in U$.*

Beispiel 9.44

(i) Falls $U = \{e\}$, so ist $x, y \in R_U \Longleftrightarrow x = y$.

(ii) Falls $G = \mathbb{Z}$ und $U = n\mathbb{Z}$, so gilt

$$(x, y) \in R_U \Longleftrightarrow x - y \in n\mathbb{Z} \Longleftrightarrow x = y \text{ (modulo } n).$$

Satz 9.45 *R_U ist eine mit der Multiplikation auf G rechtsverträgliche Äquivalenzrelation, d.h. R_U ist Äquivalenzrelation und mit*

$$(x,y) \in R_U \text{ gilt } (xa, ya) \in R_U \quad \forall a \in G.$$

Beweis: U ist Untergruppe, d.h., $e \in U$ und $xx^{-1} = e \ \forall x \in G \Longrightarrow (x,x) \in R_U \ \forall x \in G$.

Ist $xy^{-1} \in U$, so $(xy^{-1})^{-1} = yx^{-1} \in U$. Aus $xy^{-1} \in U$ und $yz^{-1} \in U$ folgt $(xy^{-1})(yz^{-1}) = xz^{-1} \in U$.
$\Longrightarrow R_U$ ist Äquivalenzrelation.
Ist $xy^{-1} \in U$ und $a \in G$

$$\Longrightarrow x(aa^{-1})y^{-1} = (xa)(ya)^{-1}$$
$$\Longrightarrow (xa, ya) \in R_u.$$

□

Wir betrachten nun die Äquivalenzklassen von R_U, $[x]_R = \{y \mid (x,y) \in R_U\}$ und da $y \in [x]_R \Longleftrightarrow yx^{-1} \in U \Longleftrightarrow y \in Ux$, so folgt $[x]_R = Ux$.

Definition 9.46 *Sei G eine Gruppe und U eine Untergruppe von G, so heißen die Mengen Ux, für $x \in G$ die Rechtsnebenklassen von U.*

Satz 9.47 *Sei U Untergruppe einer Gruppe G, dann gilt*

(i) $G = \bigcup_{x \in G} Ux$

(ii) $Ux \cap Uy \neq \emptyset \iff xy^{-1} \in U \Longleftrightarrow Ux = Uy.$

(iii) $x \in U \Longleftrightarrow Ux = U.$

(iv) Die Abbildung von $Ux \to Uy$, gegeben durch $ux \mapsto uy$, ist bijektiv, insbesondere $|Ux| = |U|$ für alle $x \in G$.

Beweis: Übung □

Definition 9.48 *Die Anzahl der verschiedenen Rechtsnebenklassen einer Untergruppe U einer Gruppe G heißt der Index von U in G. Bezeichnung $|G : U|$.*

Beispiel 9.49

(i) $|G : \{e\}| = |G|$.

(ii) $|G : G| = 1$,

(iii) $|\mathbb{Z} : n\mathbb{Z}| = n$.

Satz 9.50 (Satz von Lagrange) *Sei G eine Gruppe und $U \subset G$ Untergruppe. Sind je zwei der Größen $|G|$, $|U|$ und $|G:U|$ endlich, so auch die dritte, und es gilt*

$$|G| = |G:U|\,|U|.$$

Beweis: $G = \bigcup_{x \in G} Ux$ nach Satz 9.47(i), und $|Ux| = |U| \,\forall x \in G$.
Sei $a \in G, a \notin U \Longrightarrow U \cup Ua \subset G$, U und U_a sind disjunkt und $|U| = |Ua|$.
Sei nun $b \in G, b \notin U \cup U_a \implies |U| = |Ua| = |Ub|$ und $U \cup Ua \cup Ub \subset G$. Ist G endlich, so folgt nach endlich vielen Schritten die Behauptung. □

Korollar 9.51 *Sei G eine endliche Gruppe.*

(i) Die Ordnung jeder Untergruppe ist Teiler der Gruppenordnung.

(ii) Die Ordnung jedes Elements der Gruppe teilt die Gruppenordnung.

(iii) Ist p (eine Primzahl) die Ordnung von G, so ist G zyklisch.

(iv) Sind U, V endliche Untergruppen mit teilerfremden Ordnungen so ist $U \cap V = \{e\}$.

Beweis:

(i), (ii) folgen sofort aus Satz 9.50.

(iii) Für $a \in G, a \neq e$ ist ord $a \neq 1$. Da aber ord (a) nach (ii) Teiler von p ist, so folgt ord$(a) = |G|$, d.h., $|\langle a \rangle| = |G|$ für die Untergruppe $\langle a \rangle$ von G.
Das gilt (bei endlichen Gruppen) nur für $\langle a \rangle = G$.

(iv) $U \cap V$ ist Untergruppe von U und V.

$$|U \cap V| \,|\, |U|, |V| \implies |U \cap V| = 1 \implies U \cap V = \{e\}.$$

□

Anstatt Rechtsnebenklassen können wir natürlich auch <u>*Linksnebenklassen*</u> betrachten mit Hilfe der Relation $\tilde{R}_U$, gegeben durch

$$(x, y) \in \tilde{R}_U \iff x^{-1}y \in U.$$

$\tilde{R}_U$ ist linksverträgliche Äquivalenzrelation auf G mit Linksnebenklassen xU, $x \in G$.

Satz 9.52 *Sei U eine Untergruppe der Gruppe G, so ist die Abbildung $Ux \mapsto x^{-1}U$ eine bijektive Abbildung der Rechtsnebenklassen auf die Linksnebenklassen.*

Beweis: $Ug = Uh \iff gh^{-1} \in U \iff (g^{-1})^{-1}h^{-1} \in U \iff g^{-1}U = h^{-1}U$. $\Longrightarrow$ Injektivität.
Da $G = G^{-1}$, so ist die Abbildung auch surjektiv. □

Relationen, die links– und rechtsverträglich sind, heißen verträglich.

Kapitel 10

Normalteiler und Faktorgruppen

Satz 10.1 *Sei U eine Untergruppe einer Gruppe G. Die durch*

$$(x,y) \in R_U \iff xy^{-1} \in U$$

definierte Äquivalenzrelation auf G ist genau dann verträglich, wenn

$$aUa^{-1} \subset U \quad \forall a \in G.$$

Beweis: Sei R_U verträglich, damit ist R_U insbesondere linksverträglich, also gilt für $a \in G$ und $(x,y) \in R_U$ auch

$$(ax,ay) \in R_U \iff axy^{-1}a^{-1} \in U.$$

Mit $y = e$ folgt $axa^{-1} \in U \ \forall x \in U$ und $\forall a \in G$.

Ist umgekehrt $aUa^{-1} \subset U$ und $xy^{-1} \in U \Longrightarrow (ax)(ay)^{-1} \in U$, damit ist R_U linksverträglich und da R_U sowieso rechtsverträglich ist (Satz 9.45), folgt R_U verträglich. □

Definition 10.2 *Eine Untergruppe U einer Gruppe G heißt* Normalteiler, *wenn für alle $a \in G$ gilt, daß*

$$aUa^{-1} \subset U.$$

Satz 10.3 *Sei G eine Gruppe und U eine Untergruppe von G. Dann sind äquivalent:*

(a) U *ist Normalteiler von* G

(b) $aUa^{-1} = U, \ \forall a \in G$

(c) $aU = Ua, \ \forall a \in G$

(d) $aU \subset Ua, \ \forall a \in G$

(e) $xy^{-1} \in U \iff x^{-1}y \in U.$

Beweis: a) $\Longrightarrow$ b) U Normalteiler und $a \in G \Longrightarrow aUa^{-1} \subset U$ und $a^{-1}U(a^{-1})^{-1} \subset U$
$\Longrightarrow U = a(a^{-1}Ua)a^{-1} \subset a^{-1}Ua \subset U \implies aUa^{-1} = U.$

Der Rest des Beweises ist Übungsaufgabe. □

Beispiel 10.4

(i) $\{e\}$ und G sind natürlich Normalteiler, und in kommutativen Gruppen ist jede Untergruppe Normalteiler.

(ii) $SL_n(\mathbb{R})$ ist Normalteiler in $GL_n(\mathbb{R})$.

(iii) $\langle\delta\rangle$ ist Normalteiler in D_n (Bsp. 9.42).

Satz 10.5

(i) Sei I eine nichtleere Indexmenge und $\{U_\alpha\}_{\alpha\in I}$ Familie von Normalteilern von einer Gruppe G, so ist $U = \bigcap\limits_{\alpha\in I} U_\alpha$ auch Normalteiler von G.

(ii) Sei $\varphi : G \to H$ Gruppenhomomorphismus, U Normalteiler von G, V Normalteiler in H. Dann ist $\varphi^{-1}(V)$ Normalteiler in G, und wenn φ surjektiv ist, so ist auch $\varphi(U)$ Normalteiler von H.

Beweis:

(i) Wir haben bereits gezeigt, daß U als Durchschnitt von Untergruppen eine Untergruppe ist. Sei $a \in G, u \in U$, d.h. $u \in U_\alpha \ \ \forall\alpha \in I \Longrightarrow aua^{-1} \in U_\alpha \ \forall\alpha \in I \Longrightarrow aua^{-1} \in U$.

(ii) Die Untergruppeneigenschaft ist bereits gezeigt. Sei $x \in \varphi^{-1}(V), a \in G$. Dann gilt $\varphi(axa^{-1}) = \varphi(a)\varphi(x)\varphi(a^{-1}) \in V$ und damit $axa^{-1} \in \varphi^{-1}(V)$.
Ist U Normalteiler, so gilt $aUa^{-1} \subset U \ \forall a \in G \Longrightarrow \varphi(a)\varphi(U)\varphi(a)^{-1} \subset \varphi(U)$.
Ist φ surjektiv, so kann jedes Element z von H als $z = \varphi(a)$ mit $a \in G$ geschrieben werden $\Longrightarrow z\varphi(U)z^{-1} \subset \varphi(U) \ \forall z \in H$. □

Definition 10.6 *Sei G eine Gruppe und $X \subset G$. Die Menge $N(X) = \{g \in G \mid gX = Xg\}$ heißt <u>Normalisator von X in G.</u>*

Satz 10.7 *Sei G eine Gruppe.*

(i) $X \subset G \Longrightarrow N(X)$ ist Untergruppe von G.

(ii) Eine Untergruppe U von G ist Normalteiler $\Longleftrightarrow N(U) = G$.

(iii) Sei U Untergruppe von G. U ist Normalteiler in $N(U)$, und ist U Normalteiler in einer Untergruppe V von G, so gilt

$$V \subset N(U).$$

Beweis: Übung □

Satz 10.8

(a) Ist U Normalteiler einer Gruppe G, dann ist $G/U = \{aU \mid a \in G\}$ mit einer Verknüpfung $(aU)(bU) = (ab)U$ eine Gruppe, (die Faktorgruppe von G nach U).

(b) $|G/U| = |G : U|$.

(c) Die Abbildung $\Pi : G \to G/U$ mit $\Pi(g) = gU$ ist ein Gruppenhomomorphismus mit $\operatorname{Kern}\Pi = U$. ($\Pi$ ist der kanonische Epimorphismus).

Beweis:

(a) Da $G/U = G/R_U$ und $(aU, bU) \mapsto (ab)U$ innere Verknüpfung, so ist G/U Halbgruppe. $eU = U$ ist neutrales Element und $a^{-1}U$ das Inverse von aU. Damit folgt die Behauptung.

(b) folgt per Definition.

(c) $\Pi(ab) = (ab)U = (aU)(bU) = \Pi(a) \cdot \Pi(b) \Longrightarrow \Pi$ Homomorphismus, U ist neutrales Element. Es folgt: $a \in \operatorname{Kern}\Pi \iff \Pi(a) = U \iff aU = U \iff a \in U$. □

Satz 10.8 b) besagt für endliche Gruppen, daß $|G/U| = \dfrac{|G|}{|U|}$.

Korollar 10.9 *$U \subset G$ (Gruppe) ist Normalteiler genau dann, wenn U Kern eines Homomorphismus $\varphi : G \to H$ ist.*

Beweis: $\varphi : G \to H$ Homomorphismus $\Longrightarrow$ Kern φ ist Normalteiler, denn es ist natürlich eine Untergruppe von G, und mit $a \in G$ und $x \in \operatorname{Kern}\varphi$ folgt

$$\varphi(axa^{-1}) = \varphi(a)\varphi(x)\varphi(a^{-1}) = \varphi(a)\varphi(a^{-1}) = \varphi(aa^{-1}) = e_H \implies axa^{-1} \in \operatorname{Kern}\varphi.$$

Die andere Richtung folgt aus Satz 10.8 c). □

Definition 10.10 *Eine Gruppe G, die außer e und G keine Normalteiler besitzt, heißt* einfache *Gruppe.*

Ein Homomorphismus von einer einfachen Gruppe G in eine andere Gruppe H ist damit immer injektiv oder es gilt $g \to e_H \ \forall g \in G$.

Beispiel 10.11

(i) $\mathbb{Z}$ ist kommutativ $\Longrightarrow n\mathbb{Z}$ ist Normalteiler $\forall n$ und $\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}_n$

(ii) $(\mathbb{Z}, +)$ ist Untergruppe von $(\mathbb{R}, +)$.
$\mathbb{R}/\mathbb{Z}$ ist isomorph zur multiplikativen Untergruppe $\{e^{2\pi\alpha i} \mid \alpha \in \mathbb{R}\}$ von $\mathbb{C} \setminus \{0\}$.

(iii) $A \to \det A$ ist Homomorphismus von $GL_n(\mathbb{R}) \to \mathbb{R}\setminus\{0\} = \mathbb{R}^*$. Der Kern dieser Abbildung ist

$$SL_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid \det A = 1\}$$

$GL_n(\mathbb{R})/SL_n(\mathbb{R})$ ist isomorph zu $\mathbb{R}\setminus\{0\}$.

Satz 10.12 (Homomorphiesatz für Gruppen)
Ist $\varphi : G \to H$ ein Gruppenhomomorphismus, so ist $G/\mathrm{Kern}\,\varphi$ isomorph zu $\varphi(G)$.

Beweis: Sei $K = \mathrm{Kern}\,\varphi$, dann ist K Normalteiler und durch $(aK)(bK) = (ab)K$ ist die Gruppenstruktur auf G/K definiert.
Definiere $\Phi : G/K \to \varphi(G)$ durch $\Phi(gK) = \varphi(g)$.
Es ist zu zeigen, daß Φ eine wohldefinierte Abbildung ist. Dies gilt, da

$$\begin{aligned} gK = hK &\iff gh^{-1} \in K \iff \varphi(gh^{-1}) = e_H \iff \varphi(g) = \varphi(h) \\ &\iff \Phi(gK) = \Phi(hK). \end{aligned}$$

Φ ist surjektiv und da $\Phi((aK)(bK)) = \Phi(abK) = \varphi(ab) = \varphi(a)\varphi(b) = \Phi(aK)\Phi(bK)$, ist Φ ein Isomorphismus. □

Korollar 10.13 *Ist $\varphi : G \to H$ ein injektiver Gruppenhomomorphismus, so ist G isomorph zu $\varphi(G)$.*

Satz 10.14 *Zu jedem $n \in \mathbb{N}$ gibt es bis auf Isomorphie genau eine zyklische Gruppe der Ordnung n, nämlich $(\mathbb{Z}_n, +)$ und jede unendliche zyklische Gruppe ist isomorph zu $\mathbb{Z}$.*

Beweis: Sei $G = \langle a \rangle$ zyklisch $\Longrightarrow G = \{a^m \mid m \in \mathbb{Z}\}$. Dann ist $\gamma : m \mapsto a^m$ ein Epimorphismus von $\mathbb{Z} \to G$, $\mathrm{Kern}\,\varphi = t\mathbb{Z}$ mit $t = 0$ oder t ist die kleinste natürliche Zahl mit $a^t = e$, also $t = \mathrm{ord}\, a$. Aus dem Homomorphiesatz folgt:

G isomorph $\mathbb{Z}/\{0\}$ isomorph $\mathbb{Z}$, wenn G nicht endlich,
oder
G isomorph $\mathbb{Z}_t = \mathbb{Z}/t\mathbb{Z}$ mit $t = |G|$, sonst. □

Satz 10.15 (Erster Isomorphiesatz)
Sei U Untergruppe von G und V Normalteiler von G. Dann ist UV Untergruppe von G, $U \cap V$ Normalteiler von U und UV/V isomorph $U/(U \cap V)$.

Beweis: Sei V Normalteiler $\Longrightarrow aV = Va\ \forall a \in G \Longrightarrow UV = \bigcup_{u \in U} uV = \bigcup_{u \in U} Vu = VU$.
Nach Satz 9.29 ist UV Untergruppe, und natürlich ist $V \subset uV$ Normalteiler in UV.

Betrachte die Restriktion Π_0 des kanonischen Epimorphismus $\Pi : G \to G/V$ auf U, d.h.

$$\begin{aligned} \Pi_0 :\ & U \to G/V, \\ & u \mapsto uV. \end{aligned}$$

Da $V = vV\ \forall v \in V$, so gilt

$$\Pi_0(U) = \{uV \mid u \in U\} = \{uvV \mid uv \in UV\} = UV/V.$$

V ist das neutrale Element in G/V und $aV = V \iff a \in V$.

$$\Longrightarrow \mathrm{Kern}\,\Pi_0 = \{u \in U \mid uV = V\} = \{u \in U \mid u \in U\} = U \cap V.$$

Damit folgt, daß $U \cap V$ als Kern von Π_0 Normalteiler in U ist, und nach dem Homomorphiesatz folgt die Behauptung. □

Beispiel 10.16 Sei V_4 die Kleinsche Vierergruppe als Teilmenge von S_4.
V_4 ist Normalteiler und es gilt $S_3V_4 = S_4 \Longrightarrow S_4/V_4 = S_3V_4/V_4$ isomorph zu $S_3/(S_3 \cap V_4)$ isomorph zu S_3.

Satz 10.17 (Zweiter Isomorphiesatz) *Seien U, V Normalteiler von G mit $U \subset V$. Dann ist V/U Normalteiler von G/U und $(G/U)/(V/U)$ isomorph G/V.*

Beweis: Sei $\varphi : G/U \to G/V$ gegeben durch $\varphi(gU) = gV$.
Sei $gU = hU \implies gh^{-1} \in U \subset V \implies gV = hV$. Damit ist φ wohldefinierter Homomorphismus mit Bild $(\varphi) = G/V$ und Kern $(\varphi) = \{gU \mid gV = V\} = \{gU \mid g \in V\} = V/U$.
Der Homomorphiesatz liefert die Behauptung. □

Satz 10.18 *Sei N Normalteiler einer Gruppe G und $\Pi : G \to G/N$ der kanonische Epimorphismus.*

(i) Ist U Untergruppe von G, so ist $\Pi(U) = UN/N$ Untergruppe von G/N.

(ii) Ist V Untergruppe von G/N, dann ist $\Pi^{-1}(V)$ Untergruppe von G, die N enthält und $V = \Pi^{-1}(V)/N$.

(iii) $U \mapsto \Pi(U)$ ist bijektiv und $\Pi(U)$ ist Normalteiler in G/N genau dann, wenn U Normalteiler in G ist. Dann gilt

$$|G : U| = |G/N : \Pi(U)|.$$

Beweis:

(i) und (ii) folgen aus der Definition von Π und der Tatsache, daß Bild $(\Pi(U))$ Untergruppe ist.

(ii) Für Untergruppe U mit $N \subset U$ folgt aus (i) und (ii), daß $\Pi^{-1}(\Pi(U)) = U$ und für Untergruppe V von G/N gilt $\Pi(\Pi^{-1}(V)) = V$. Das zeigt, daß die Abbildungen invers zueinander sind, Satz 10.5(ii). Da $N \subset U$, so ist der zweite Isomorphiesatz anwendbar und Satz 10.8 b) ergibt

$$\begin{aligned} |G : U| &= |G/U| = |(G/N)/(U/N)| \\ &= |G/N : \Pi(U)|. \end{aligned}$$

□

Kapitel 11

Die Sylow'schen Sätze

Definition 11.1 *Es sei G eine Gruppe mit neutralem Element e und X eine nichtleere Menge. Wir sagen, daß G auf X operiert, wenn es eine Abbildung*

$$\begin{aligned} \cdot : \ & G \times X \to X \\ & (g, x) \mapsto g \cdot x \end{aligned}$$

gibt, so daß $\forall\, g, h \in G,\ x \in X$

$$(gh) \cdot x = g \cdot (h \cdot x), \tag{11.2}$$

$$ex = x. \tag{11.3}$$

Wir haben schon gesehen, daß jeder Homomorphismus von $G \to S(X)$ eine Operation von G auf X definiert.
Bilde nun für gegebene Operation $(g, x) \mapsto g \cdot x$ zu $g \in G$ eine Abbildung

$$\Psi_g : x \mapsto g \cdot x$$

von X in sich, so folgt aus (11.2) und (11.3), daß

$$\Psi_g \Psi_{g^{-1}}(x) = g(g^{-1}x) = (gg^{-1})x = ex = x, \quad \text{also } \Psi_g \Psi_{g^{-1}} = id.$$

Also ist Ψ_g in $S(X)$, und die Abbildung $\Psi : g \mapsto \Psi_g$ von $G \to S(X)$ ist ein Homomorphismus. Damit operiert G auf X genau dann, wenn es einen Homomorphismus von $G \to S(X)$ gibt.

Satz 11.4 *G operiere auf $X \neq \emptyset$. Dann ist*

$$R(G) = \{(x, y) \in X \times X \mid \exists\, g \in G \text{ mit } g \cdot x = y\}$$

eine Äquivalenzrelation.

Beweis: Da $ex = x$, so ist $(x, x) \in R(G)\ \forall x \in X$.
Sei $(x, y) \in R(G)$, d.h. $g \cdot x = y$ für ein $g \in G$,

$$\begin{aligned} &\Longrightarrow \quad g^{-1}y = g^{-1}(gx) = (g^{-1}g)x = ex = x \\ &\Longrightarrow \quad (y, x) \in R(G). \end{aligned}$$

Für $(x, y) \in R(G), (y, z) \in R(G)$, d.h., $(hg)x = h(gx) = hy = z$, folgt $(x, z) \in R(G)$. □

Definition 11.5 *Die Äquivalenzklassen von $R(G)$ heißen* Orbits *(Bahnen) von G in X.*

$$\begin{aligned} [x] &= \{y \in X \mid (x, y) \in R(G)\} \\ &= \{y \in X \mid \exists\, g \in G \text{ mit } g \cdot x = y\} \\ &= \{g \cdot x \mid g \in G\} = Gx \end{aligned}$$

Definition 11.6 *Eine Gruppe G operiere auf $X \neq \emptyset$. Dann heißt*

$$G_x = \{g \in G \mid g \cdot x = x\}$$

der Stabilisator *von $x \in X$ in G.*

Satz 11.7 *Die Gruppe G operiere auf $X \neq \emptyset$. Dann ist für jedes $x \in X$ der Stabilisator eine Untergruppe von G und es gilt*

$$|Gx| = |G : G_x|.$$

Beweis: $h \in G_x \Longrightarrow h^{-1}x = h^{-1}(hx) = (hh^{-1})x = ex = x \Longrightarrow h^{-1} \in G_x$.
Sind $h, g \in G_x$, so gilt

$$(gh)x = g(hx) = gx = x \Longrightarrow gh \in G_x.$$

Für die zweite Behauptung betrachte die Abbildung $\gamma : g \cdot x \mapsto gG_x$.

$gx = hx \iff h^{-1}g \in G_x \iff hG_x = gG_x$. Also ist γ wohldefiniert und bijektive Abbildung von Gx auf die Menge der Linksnebenklassen von G nach G_x. $\Longrightarrow$ Behauptung. □

Aus dem Satz von Lagrange folgt dann sofort:

Korollar 11.8 *Ist G eine endliche Gruppe, so ist für jedes $x \in X$ die Anzahl der Elemente von $G \cdot x$ ein Teiler der Gruppenordnung $|G|$.*

Beweis: Folgt aus Satz 11.8 und dann Satz von Lagrange. □

Eine Untergruppe U von G operiert auf G durch Linksmultiplikation $(u, g) \mapsto ug$.

Der Orbit von $g \in G$ ist die Rechtsnebenklasse Ug.
Aufgrund der Kürzungsregeln ist jeder Stabilisator trivial, d.h. $U_g = \{e\},\ \forall\, g \in G$.

Satz 11.9 *Die Anzahl der verschiedenen konjugierten Untergruppen einer Untergruppe U von G ist gleich dem Index des Normalisators U von G.*

Beweis: Sei $X \subset G$ und $\mathcal{K} = \{gXg^{-1} \mid g \in G\}$.
Die Elemente aus $\mathcal{K}$ heißen die zu X konjugierten Teilmengen.
G operiert auf $\mathcal{K}$ durch $h \cdot y = hyh^{-1}$, und es gibt offensichtlich nur einen Orbit.
$\Longrightarrow$ Stabilisator von X ist gleich Normalisator von X:

$$G_x = N(x) = \{g \in G \mid gX = Xg\} \implies |\mathcal{K}| = |G : N(x)|.$$

Falls X Untergruppe von G ist, so folgt die Behauptung. □

Beispiel 11.10 Die Wirkung des inneren Automorphismus einer Gruppe G kann man als Operation von G auf sich auffassen. Die Operation $g \cdot h = ghg^{-1}$ heißt *Konjugation* mit den Orbits $G \cdot h = \{ghg^{-1} \mid g \in G\}$. Der Stabilisator G_x stimmt hier mit dem Normalisator $N(x) = \{g \in G \mid gx = xg\}$ überein.

Wir können $X = \bigcup G \cdot x$ in eine Menge disjunkter Orbits zerlegen. So reicht es, jeden Orbit als Vereinigung über eine Teilmenge darzustellen, die aus jedem Orbit genau ein Element enthält. So eine Menge heißt Vertretersystem für Orbits $V \subset X$ ist ein Vertretersystem genau dann, wenn

$$\forall x \in X \ \ \exists v \in V \ \text{ mit } \ G \cdot x = G \cdot v \tag{11.11}$$

$$\forall a, b \in V \ \text{ mit } \ a \neq b \ \text{ ist auch } G \cdot a \neq G \cdot b. \tag{11.12}$$

Lemma 11.13 *Sei $n \in \mathbb{N}$, p Primzahl und $n = p^k m$ mit $(p, m) = 1$, dann ist für jedes r mit $1 \leq r \leq k$ p^{k-r+1} kein Teiler von $\binom{n}{p^r}$.*

Beweis: Übung. □

Satz 11.14 (Erster Sylow'scher Satz) *Ist G endliche Gruppe der Ordnung $n = p^k m$ mit p Primzahl und $(p, m) = 1$, so gibt es zu jedem r mit $1 \leq r \leq k$ eine Untergruppe von G der Ordnung p^r.*

Beweis: Sei Z die Menge der Teilmengen von G, die genau p^r Elemente enthalten,
$Z = \{A \subset G \mid |A| = p^r\}$.
Dann gilt $|Z| = \binom{n}{p^r}$ und da für $A \subset G$ und $g \in G$ stets $|gA| = |A|$ gilt, operiert G vermöge $(g, A) \mapsto gA$ auf Z.
Also können wir Z in disjunkte Orbits zerlegen, $Z = \bigcup\limits_{A \in Z} GA$, und wir erhalten

$$\binom{n}{p^r} = |Z| = \sum_{A \in V_Z} |GA| = \sum_{A \in V_Z} |G : G_A|$$

wobei V_Z ein Vertretersystem für die Orbits ist.

Da $p^{k-r+1} \nmid \binom{n}{p^r}$, so gibt es mindestens einen Summanden, der nicht durch p^{k-r+1} teilbar ist. $\Longrightarrow \exists\, B \in Z$ mit $p^{k-r+1} \nmid |G : G_B|$.
Es ist demnach p^{k-r} die höchste p-Potenz, die als Teiler von $|G : G_B|$ möglich ist. Da

$$p^k m = |G| = |G : G_B| \cdot |G_B|,$$

kommt p als Faktor k-mal in $|G : G_B| \cdot |G_B|$ vor, jedoch höchsten $(k-r)$-mal in $|G : G_B|$
$\Longrightarrow$ p ist als Faktor mindestens r mal in $|G_B|$
$\Longrightarrow$ $|G_B| \geq p^r$.
Ist $b \in B$, so folgt aus der Definition des Stabilisators, daß $G_B B = B$ oder $G_B b \subset B$.
Also folgt für die Anzahl der Elemente des Stabilisators G_B, daß

$$|G_B| = |G_B b| \leq |B| = p^r.$$

Zusammen mit $|G_B| \geq p^r$ folgt $|G_B| = p^r$.
Der Stabilisator ist also die gesuchte Untergruppe der Ordnung p^r. □

Satz 11.15 (von Cauchy) *Ist G eine endliche Gruppe und die Primzahl p ein Teiler von $|G|$, dann enthält G ein Element der Ordnung p.*

Beweis: Nach Satz 11.14 enthält G eine Untergruppe U der Ordnung p. Nach Korollar 9.51 ist eine Untergruppe zyklisch, also $U = \langle a \rangle$ und $\mathrm{ord}(a) = p$. □

Definition 11.16 *Sei p eine Primzahl und G eine Gruppe. G heißt* p–Gruppe*, wenn jedes Element von G eine p–Potenz als Ordnung hat, d.h., zu $g \in G$ gibt es $k \geq 0$ mit $g^{p^k} = e$.*

Korollar 11.17 *Eine endliche Gruppe ist genau dann p–Gruppe, wenn ihre Ordnung eine Potenz von p ist.*

Beweis: Machen wir hier nicht. □

Definition 11.18 *Sei G Gruppe und $U \subset G$ Untergruppe. U heißt* p–Sylow–Gruppe *von G, falls gilt:*

$$U \text{ ist } p\text{-Untergruppe von } G. \tag{11.19}$$
$$\text{Ist } H \text{ eine } p\text{-Untergruppe von } G \text{ mit } U \subset H, \text{ dann folgt } U = H. \tag{11.20}$$

Das bedeutet, daß die p–Sylow–Gruppen die maximalen unter den p–Untergruppen bilden.

Satz 11.21 *Sei G endliche Gruppe der Ordnung $n = p^k m$ mit p Primzahl und $(p, m) = 1$, dann ist jede Untergruppe der Ordnung p^k eine p–Sylow–Gruppe von G.*

Beweis: $U \subset G$ Untergruppe mit $|U| = p^k$. Nach Korollar 11.17 ist U eine p–Gruppe. Sei H eine p–Gruppe mit $U \subset H \subset G$. Da wieder nach Korollar 11.17 $|H| = p^l$ und nach dem Satz von Lagrange $p^l | p^k m$ folgt $l \leq k \implies |H| \leq |U|$ und da $U \subset H$, folgt $U = H$. □

Korollar 11.22 *Hat G die Ordnung $p^k m$, p Primzahl, $(p, m) = 1$, so hat G mindestens eine p–Sylow–Gruppe der Ordnung p^k.*

Beweis: Nach Satz 11.14 enthält G eine Untergruppe der Ordnung p^k, und diese ist nach Satz 11.21 p–Sylow–Gruppe. □

Lemma 11.23 *Ist U eine p–Untergruppe (p–Sylow–Gruppe) von G, dann sind auch alle konjugierten Untergruppen gUg^{-1}, $g \in G$, p–Untergruppen (p–Sylow–Gruppen).*

Beweis: Da x und gxg^{-1} gleiche Ordnung haben, ist mit U auch gUg^{-1} eine p–Gruppe. Ist U eine p–Sylow–Gruppe und wäre die p–Gruppe gUg^{-1} echt in einer p–Gruppe H enthalten, dann wäre U echt in der p–Gruppe $g^{-1}Hg$ enthalten, im Widerspruch zu (11.20). □

Satz 11.24 (Zweiter Sylow'scher Satz)

(i) Sei G eine endliche Gruppe der Ordnung $p^k m (k \geq 1)$ mit $(p, m) = 1$ und p Primzahl, und sei P eine p–Sylow–Gruppe von G. Dann enthält P von jeder p–Untergruppe U von G eine konjugierte, d.h. $\exists\, a \in G$ mit $aUa^{-1} \subset P$.

(ii) Je zwei p–Sylow–Gruppen von G sind konjugiert, damit isomorph und von der Ordnung p^k.

(iii) Eine p–Sylow–Gruppe P von G ist genau dann ein Normalteiler, wenn P die einzige p–Sylow–Gruppe in G ist.

Satz 11.25 (Dritter Sylow'scher Satz) *In einer endlichen Gruppe G, deren Ordnung durch p (Primzahl) teilbar ist, ist die Anzahl der p–Sylow–Gruppen ein Teiler der Gruppenordnung und hat die Form*

$$1 + kp, \quad k \geq 0.$$

Beweise siehe: Meyberg, Algebra.

Man nutzt diese Sätze, um die endlichen Gruppen zu klassifizieren, indem man zuerst mal die p–Gruppen analysiert, um damit Aussagen für beliebige Gruppen zu machen. Mehr dazu in Vorlesungen zur Algebra.

Inhaltsverzeichnis

Index

Numbers written in italic refer to the page where the corresponding entry is described, the ones underlined to the definition, the rest to the pages where the entry is used.