Z Zerlegung von Moduln

Im Kapitel über Eigenwerte haben wir gezeigt, daß jeder Endomorphismus f in einem endlich-dimensionalen Vektorraum eine Matrix-Darstellung in Diagonalform besitzt, sofern eine Basis aus Eigenvektoren vorhanden ist, aber auch gesehen, daß solche Basen nicht zu existieren brauchen. Wir wollen nun diese Frage unter allgemeineren Gesichtspunkten neu behandeln. Wir werden insbesondere Normalformen für Matrix-Darstellungen erhalten, aus denen sich dann unter speziellen Voraussetzungen auch die Diagonalform ergeben wird. Darüberhinaus werden wir noch wesentlich allgemeinere Sachverhalte mit einbeziehen können.

Zur Vorbereitung brauchen wir etwas Algebra.

Die Polynomalgebra

Definition Z.1 (K-Algebra) *Ein K-Vektorraum A, in dem noch eine weitere Operation, genannt "Multiplikation" $\cdot : A \times A \to A$ erklärt ist, heißt eine K-Algebra, wenn gelten:*

(i) Die Multiplikation $\cdot$ ist associativ,

(ii) Es gelten die Distributivgesetze $(a, b, c \in A, \alpha, \beta, \gamma \in K)$

$$(\alpha a + \beta b) \cdot c = \alpha(a \cdot c) + \beta(b \cdot c)$$
$$a \cdot (\beta b + \gamma c) = \beta(a \cdot b) + \gamma(a \cdot c).$$

Gibt es ein Element $1 \in A$, sodaß $\forall_{a \in A}\ 1 \cdot a = a \cdot 1 = a$, so heißt A eine K-Algebra mit 1. Für $\alpha \in K$ schreibt man statt $\alpha \cdot 1$ kurz α. Ist die Multiplikation kommutativ, so heißt A eine kommutative K-Algebra. Der Multiplikations-Punkt wird meist nicht geschrieben.

Im weiteren betrachten wir nur K-Algebren mit 1-Element.

Zwei simple Beispiele sind etwa K selbst oder im Falle $K = \mathbb{R}$ die "Polynome". Beide sind kommutativ mit 1. Ein für das Weitere wichtiges Beispiel enthält

Lemma Z.2 *V sei ein K-Vektorraum. Dann ist* $\mathrm{Hom}(V, V)$ *mit der üblichen Addition und dem Hintereinanderausführen von Homomorphismen als Multiplikation eine im allgemeinen nicht kommutative K-Algebra mit* id_V *als 1-Element.*

Ein weiteres wichtiges Beispiel ist folgendes:

Zu einem Körper K bilden wir die sogenannte "direkte Summe" über $\mathbb{N}$ viele Exemplare von K :

$$P := \bigoplus_{i \in \mathbb{N}} K := \{(\alpha_0, \alpha_1, ...) \mid \alpha_i \in K, \neq 0 \text{ nur endlich oft.}\}.$$

In Analogie zu den üblichen Einheitsvektoren in K^n bezeichnen wir die Folgen $(0, ..., 0, 1, 0, ...)$, die genau an der Position i eine 1, sonst nur 0 haben, mit dem Symbol e_i. Offensichtlich ist P ein Vektorraum und $(e_i \mid i = 0, 1, 2, ...)$ eine Basis von P.

Wir definieren nun in P eine Multiplikation $\cdot$ aus der Festsetzung

$$e_i \cdot e_j := e_{i+j},$$

was, um die Distributivgesetze zu erhalten, zwangsläufig zu

$$\left(\sum_{i=0}^{n} \alpha_i e_i\right) \cdot \left(\sum_{j=0}^{m} \beta_j e_j\right) = \sum_{k=0}^{n+m} \gamma_k e_k \text{ mit } \gamma_k = \sum_{i+j=k} \alpha_i \beta_j$$

führt.

Satz und Definition Z.3 *Mit der eben eingeführten Multiplikation ist der Vektorraum* $P := \bigoplus_{i\in\mathbb{N}} K$ *eine kommutative* K*-Algebra mit dem 1-Element* e_0*.* P *heißt die "freie, von einem Element erzeugte* K*-Algebra" oder auch "Algebra der Polynome in einer Veränderlichen über* K*" und wird auch mit* $K[T]$ *- gelesen als* K *adjungiert* T *- bezeichnet. Die Elemente von* $K[T]$ *heißen "Polynome" in einer Veränderlichen über* K*.*

Der **Beweis**, daß es sich hier um eine Algebra handelt, d.h. das Nachrechnen der Axiome sei als Übung gelassen. Die zunächst befremdlich erscheinende Bezeichnung "Polynomalgebra" wir verständlich, wenn wir die Symbole anders wählen.

Schreiben wir in P einmal statt e_0 das Symbol 1, statt e_1 das Symbol T. Dann ist wegen $e_i \cdot e_j := e_{i+j}$ offensichtlich

$$e_2 = e_1 \cdot e_1 = T \cdot T = T^2,$$

$$e_3 = e_2 \cdot e_1 = T^2 \cdot T = T^3,$$

u.s.w. Jedes Element von P läßt sich also statt als

$$p = \alpha_0 e_0 + \alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n$$

auch schreiben als

$$\alpha_0 1 + \alpha_1 T + \alpha_2 t^2 + ... + \alpha_n T^n,$$

und die Multiplikation von solche Elementen verläuft genauso, wie wir es naiv von "Polynomen" gewohnt sind.

Im weiteren werden wir nur noch $K[T]$ statt P schreiben und analog T statt e_1, sowie 1 oder auch T^0 statt e_0. Diese freie, von einem Element erzeugte K-Algebra besitzt eine (sie auch bis auf Isomorphie kennzeichnende) *universelle Eigenschaft* über die Existenz und Eindeutigkeit gewisser Homomorphismen. Dazu zunächst

Definition Z.4 (K-Algebra-Homomorphismus) *Sind* A, A' K*-Algebren mit 1-Elementen* 1_A *bzw.* $1_{A'}$*, so heißt eine Abbildung* $\varphi : A \to A'$ *ein* K*-Algebra-Homomorphismus, wenn gilt:*

(i) φ *ist Vektorraum-Homomorphismus,*

(ii) $\forall_{a,b\in A} \quad \varphi(ab) = \varphi(a)\varphi(b).$

(iii) $\varphi(1_A) = 1_{A'}.$

Die freie Algebra $K[T]$ besitzt nun die folgende *universelle Eigenschaft*:

Satz Z.5 *Ist* A *irgend eine* K*-Algebra mit 1-Element* 1_A*, so gibt es zu jedem Element* $a \in A$ *genau einen* K*-Algebra-Homomorphismus*

$$\varphi_a : K[T] \to A \quad \textit{mit} \quad \varphi_a(T) = a.$$

Beweis: Ist φ ein solcher Homomorphismus, so ist notwendig für

$$p = \sum_0^n \alpha_i T^i \in K[T]$$

dann

$$\varphi(p) = \varphi\left(\sum_0^n \alpha_i T^i\right) = \sum_0^n \alpha_i \varphi(T^i) = \sum_0^n \alpha_i \varphi(T)^i = \sum_0^n \alpha_i a^i,$$

wobei noch als Notation $a^0 := 1_A$ verwendet ist.

Andrerseits prüft man leicht nach, daß die Abbildung

$$\varphi_a : \sum_0^n \alpha_i T^i \mapsto \sum_0^n \alpha_i a^i$$

tatsächlich ein Homomorphismus mit den gewünschten Eigenschaften ist. □

Dieser durch die Vorgabe von $a \in A$ eindeutig bestimmte Homomorphismus φ_a "setzt also in jedes Polynom p statt T das Element a ein". Wir schreiben dafür auch abkürzend

$$\varphi_a(p) =: p(a).$$

Ist speziell $A = K = \mathbb{R}$, so ist dies das übliche Auswerten eines reellen Polynoms an der reellen Stelle a. Wir sind jetzt aber nicht mehr auf diesen Spezialfall beschränkt.

Ist V ein K-Vektorraum, so ist Hom(V,V) eine K-Algebra und wir können Satz Z.5 hierauf anwenden. Ist dann $f \in \text{Hom(V,V)}$, $p = \sum_0^n \alpha_i T^i$, so ist

$$\varphi_f(p) = p(f) = \sum_0^n \alpha_i f^i \in \text{ Hom(V,V)},$$

also wieder ein Endomorphismus in V. Da φ_f die 1-Elemente in sich überführt, ist insbesondere für jedes $f \in \text{Hom(V,V)}$:

$$\varphi_f(1) = \text{id}_V.$$

Die Algebra Hom(V,V) ist im allgemeinen nicht kommutativ. Die durch Einsetzen eines Endomorphismus f in verschiedene Polynome p, q entstehenden Endomorphismen sind jedoch stets vertauschbar, da $K[T]$ kommutativ ist:

$$p(f)q(f) = \varphi_f(p)\varphi_f(q) = \varphi_f(pq) = \varphi_f(qp) = \varphi_f(q)\varphi_f(p) = q(f)p(f).$$

All dies werden wir in Kapitel A noch ausführlicher behandeln.

Wir brauchen noch einige Ergebnisse über $K[T]$.

Ist $p = \sum_0^n \alpha_i T^i$, so heißt $m := \max\{i \mid \alpha_i \neq 0\}$ der *"Grad"* von p : $\deg p$ und α_m der *"höchste"* Koeffizient von p. Ist er 1, so heißt p *"normiert"*. Ist $p = 0$, d.h. das Nullpolynom, in dem alle Koeffizienten $= 0$ sind, so setzen wir $\deg 0 := -\infty$. Für den Grad gelten folgende Regeln:

$$\deg(p+q) \leq \max\{\deg p, \deg q\}$$

$$\deg(p \cdot q) = \deg p + \deg q.$$

Satz Z.6 (Division mit Rest) *Zu je zwei Polynomen p_1, p_2 $(p_2 \neq 0)$ gibt es eindeutig bestimmte Polynome q, r mit $\deg r < \deg p_2$, sodaß*

$$p_1 = qp_2 + r.$$

Beweis: Ist $\deg p_1 < \deg p_2$, so setze $q := 0, r := p_1$.

Sei also $m_1 := \deg p_1 \geq m_2 := \deg p_2$, $p_1 = \sum_0^{m_1} \alpha_i T^i$ und $p_2 = \sum_0^{m_2} \beta_i T^i$, wobei $\alpha_{m_1} \neq 0, \beta_{m_2} \neq 0$. Wir setzen $q_1 := \frac{\alpha_{m_1}}{\beta_{m_2}} T^{m_1 - m_2}$ und $r_1 := p_1 - q_1 p_2$. Dann ist

$$p_1 = q_1 p_2 + r_1$$

und $\deg r_1 \leq m_1 - 1 < \deg p_1$, da wir gerade den höchsten Koeffizienten annulliert haben.

Ist schon $\deg r_1 < \deg p_2$, so sind wir fertig.

Andernfalls ist noch $\deg r_1 \geq \deg p_2$ und wir können mit r_1 statt p_1 dieselbe Konstruktion nochmal machen und erhalten q_2, r_2 mit $\deg r_2 < \deg r_1$, sodaß

$$r_1 = q_2 p_2 + r_2.$$

Dann ist

$$p_1 = q_1 p_2 + r_1 = q_1 p_2 + q_2 p_2 + r_2 = (q_1 + q_2) p_2 + r_2,$$

wobei $\deg r_2 \leq \deg p_1 - 2$.

Dies kann man iterieren, bis schließlich bei einem k-ten Schritt der dann erhaltene Rest r_k die gewünschte Ungleichung $\deg r_k < \deg p_2$ erfüllt. Die Details seien als Übung gelassen.

Die *Eindeutigkeit* der obigen Darstellung folgt so:

Sei $p_1 = q_1 p_2 + r_1$ und $p_1 = q_2 p_2 + r_2$ mit $\deg r_i < \deg p_2$ $(i = 1, 2)$. Dann ist $(q_1 - q_2) p_2 = r_2 - r_1$. Ist $q_1 \neq q_2$, so steht links ein Polynom vom Grad $\geq \deg p_2$, rechts eines vom Grad $< \deg p_2$, was nicht sein kann. Also ist $q_1 = q_2$ und damit auch $r_1 = r_2$. □

Wir sagen: Ein Polynom p_2 *"teilt"* ein Polynom p_1, wenn ein Polynom q existiert, sodaß $p_1 = q p_2$. Es heißt p ein *"gemeinsamer Teiler"* von $p_1, ..., p_n$, wenn p jedes p_i teilt. p ist *"größter gemeinsamer Teiler"* von $(p_1, ..., p_n)$, abgekürzt $\mathrm{ggT}(p_1, ..., p_n)$, wenn

p gemeinsamer Teiler von $(p_1, ..., p_n)$ und
jeder gemeinsame Teiler von $(p_1, ..., p_n)$ auch Teiler von p ist.

Polynome heißen *"teilerfremd"*, wenn bis auf Normierung 1 ggT ist.

Den ggT kann man algorithmisch berechnen:

Satz Z.7 (Euklids Algorithmus) *Zu zwei nichttrivialen Polynomen p_1, p_2 definieren wir weitere Polynome durch fortlaufende Division mit Rest:*

Ist $p_{n+1} \neq 0$, so sei p_{n+2} der bei Division von p_n durch p_{n+1} auftretende Rest mit $\deg p_{n+2} < \deg p_{n+1}$. Wir haben also für $n = 1, 2, ...$ solange $p_{n+1} \neq 0$ die Darstellung

$$p_n = q_n \cdot p_{n+1} + p_{n+2}.$$

Ist dann m die größte Nummer mit $p_m \neq 0$, so ist p_m größter gemeinsamer Teiler von p_1 und p_2.

Beweis: Œ sei $\deg p_1 \geq \deg p_2$. Dann haben wir nach Satz Z.6 für die Grade jeweils $\deg p_1 \geq \deg p_2 > \deg p_3 > .....$. Der Algorithmus bricht also ab. Ist nun p_m der letzte nicht verschwindende Rest bei dieser Kette von Divisionen, so haben wir

$$\begin{array}{rcccc}
p_1 & = & q_1 p_2 & + & p_3 \\
p_2 & = & q_2 p_3 & + & p_4 \\
 & & \vdots & & \\
p_{m-3} & = & q_{m-3} p_{m-2} & + & p_{m-1} \\
p_{m-2} & = & q_{m-2} p_{m-1} & + & p_m \\
p_{m-1} & = & q_{m-1} p_m & &
\end{array}$$

Nach der letzten Zeile ist p_m Teiler von p_{m-1}. Also ist nach der vorletzten Zeile p_m auch Teiler von p_{m-2} und indem man das Divisions-Schema rückwärts durchgeht, folgt, daß p_m alle p_i teilt, also auch p_1 und p_2.

Hat man in d einen gemeinsamen Teiler von p_1 und p_2, so gehe man das Schema von oben durch und erhält sukzessive d als Teiler von $p_2, p_3, ...$ und schließlich von p_m. Somit ist p_m der ggT. □

Euklids Algorithmus liefert uns auch eine geschlossene Darstellung für den ggT:

Satz Z.8 *Sind p_1, p_2 Polynome, d ein* ggT(p_1, p_2)*, so existieren Polynome s_1, s_2 sodaß*

$$d = s_1 p_1 + s_2 p_2.$$

Sind insbesondere p_1, p_2 teilerfremd, so gibt es eine Darstellung

$$1 = s_1 p_1 + s_2 p_2.$$

Beweis: Wir verwenden die Divisionskette von Euklids Algorithmus. Sei $d := p_m$ der letzte nicht verschwindende Rest. Dann ist (vorletzte Zeile!)

$$d = p_m = p_{m-2} - q_{m-2} p_{m-1}.$$

Aus der davor liegenden Zeile ergibt sich die Darstellung

$$p_{m-1} = p_{m-3} - q_{m-3} p_{m-2},$$

womit wir aus der eben gewonnen Formel p_{m-1} eliminieren können. Es entsteht

$$d = p_m = p_{m-2} - q_{m-2}(p_{m-3} - q_{m-3} p_{m-2}) = (-q_{m-2}) p_{m-3} + (1 + q_{m-2} q_{m-3}) p_{m-2},$$

sodaß d durch p_{m-2} und p_{m-3} dargestellt ist. Dies kann man in naheliegender Weise iterieren, was die behauptete Darstellung liefert. □

Wir hatten bemerkt, daß jeder Körper K insbesondere eine K-Algebra ist. Dann können wir also in ein Polynom $p \in K[T]$ ein Körperelement λ einsetzen und erhalten wieder ein Körperelement $p(\lambda)$. Wir nennen λ eine *"Nullstelle"* von p, wenn $p(\lambda) = 0$.

Lemma Z.9 *λ ist Nullstelle von P, genau wenn $(T - \lambda)$ Teiler von p.*

Beweis:
Sei $(T - \lambda)$ Teiler von p: Dann haben wir eine Darstellung $p(T) = q(T) \cdot (T - \lambda)$, sodaß (setze λ ein!) $p(\lambda) = q(\lambda) \cdot (\lambda - \lambda) = 0$.
Sei λ Nullstelle von p: Per Division mit Rest gibt es eine Darstellung

$$p(T) = q(T) \cdot (T - \lambda) + r(T)$$

mit $\deg r < \deg(T - \lambda) = 1$. Demnach besitzt das Polynom r nur einen einzigen Koeffizienten, den zu T^0 und der muß wegen $p(\lambda) = 0$ selbst $= 0$ sein. Also ist $r = 0$ und damit $(T - \lambda)$ Teiler von p. □

Da in einem algebraisch abgeschlossenen Körper — $\mathbb{C}$ ist ein solcher — jedes Polynom vom Grad ≥ 1 eine Nullstelle hat, erhalten wir

Satz Z.10 *Ist K ein algebraisch abgeschlossener Körper, $p \in K[T]$, $\deg p = n$ mit $n \geq 1$, so existieren Körperelemente $\lambda_1, ..., \lambda_n$ sodaß, wenn α_n den höchsten Koeffizienten von p bezeichnet*

$$p(T) = \alpha_n (T - \lambda_1) \cdot (T - \lambda_2) \cdot ... \cdot (T - \lambda_n).$$

Genau die λ_i sind die Nullstellen von p.

Beweis: Da K algebraisch abgeschlossen ist und $n \geq 1$, hat p eine Nullstelle λ_1. Dann ist nach dem Lemma also

$$p(T) = p_1(T)(T - \lambda_1) \text{ wobei } \deg p_1 = \deg p - 1 = n - 1.$$

Ist $n = 1$, so ist als $p_1 \in K$ und damit $= \alpha_n$.

Ist $n > 1$, so ist noch $\deg p_1 \geq 1$, p_1 besitzt somit einen Nullstelle λ_2 ..... . Es ist wohl klar, wie man zu argumentieren hat. □

Diese Nullstellen brauchen nicht paarweise verschieden zu sein. Dies ergibt

Korollar Z.11 *Über einem algebraisch abgeschlossenen Körper K besitzt jedes normierte Polynom p vom Grad $n \geq 1$ eine Darstellung*

$$p(T) = \prod_1^m (T - \lambda_i)^{\nu_i}$$

mit paarweise verschiedenen $\lambda_i \in K$ und Exponenten $\nu_i \in \mathbb{N}$. Die ν_i heißen die Ordnungen der λ_i als Nullstelle von p.

Aus dieser Darstellung folgt sofort

Satz Z.12 *Über einem algebraisch abgeschlossenen Körper sind zwei Polynome genau dann teilerfremd, wenn sie keine gemeinsame Nullstelle haben.*

Beweis als Übung.

Da viele Körper, etwa $\mathbb{Q}, \mathbb{R}$ und alle endlichen Körper nicht algebraisch abgeschlossen sind, wollen wir noch untersuchen, in welcher Form die Aussage von Korollar Z.11 dort gilt.

Betrachten wir das Polynom $p(T) = T^4 - 2$.

Als Polynom in $\mathbb{C}[T]$ können wir es zerlegen in

$$p(T) = (T - \sqrt[4]{2})(T + \sqrt[4]{2})(T - i\sqrt[4]{2})(T + i\sqrt[4]{2}).$$

In $\mathbb{R}[T]$ geht das so nicht mehr, aber wir finden noch die Darstellung

$$p(T) = (T - \sqrt[4]{2})(T + \sqrt[4]{2})(T^2 + \sqrt{2}),$$

während über $\mathbb{Q}$ abgesehen von Trivialitäten wie

$$p(T) = \frac{1}{3} \cdot \left(3 \cdot (T^4 - 2)\right)$$

überhaupt keine Zerlegung in Faktoren mehr möglich ist.

Die Zerlegbarkeit von Polynomen hängt also vom Körper ab.

Definition Z.13 (Irreduzibel, prim)

(i) Ein Polynom $p \in K[T]$ heißt irreduzibel (über K), wenn p nur triviale Teiler besitzt, d.h. wenn gilt:
$p \notin K$ und ist $p = q_1 \cdot q_2$ so ist $q_1 \in K$ oder $q_2 \in K$, d.h. einer der Teiler ist ein Polynom vom Grad 0.

(ii) Ein Polynom $p \in K[T]$ heißt prim (über K), wenn gilt:
Sind $p_1, p_2 \in K[T]$ und teilt p das Produkt $p_1 \cdot p_2$, so teilt p schon einen der Faktoren also p_1 oder p_2.

Satz Z.14 *In $K[T]$ gilt: Jedes irreduzible Polynom ist prim.*

Beweis: Nehmen wir an, das irreduzible Polynom p (Œvom Grad ≥ 1) teile $p_1 p_2$, aber nicht p_1. Dann sind p und p_1 teilerfremd. Nach Satz Z.8 gibt es also Polynome s, s_1, sodaß $1 = sp + s_1 p_1$ ist. Dann ist auch $p_2 = p_2 sp + s_1 p_1 p_2$. Hier teilt p die rechte Seite und damit auch die linke, d.h. $p \mid p_2$. □

Satz Z.15 *Jedes normierte Polynom $p \in K[T]$ vom Grad $n \geq 1$ besitzt eine - bis auf die Nummerierung der Faktoren - eindeutig bestimmte Zerlegung*

$$p(T) = \prod_{i=1}^{m} q_i(T)^{\nu_i},$$

wobei die q_i normierte, paarweise verschiedene irreduzible Polynome vom Grad ≥ 1 sind, die Exponenten $\nu_i \in \mathbb{N}$, ≥ 1 und $\sum_1^m \nu_i \deg q_i = n$.

Beweis: *Existenz:* Induktion über $n := \deg p$:

Ist $n = 1$, so hat p die Form $p(T) = (T - \lambda)$ und dies ist mit $m = 1, \nu_1 = 1$, $q_1(T) = (T - \lambda)$ die behauptete Darstellung.

Ist $n > 1$ und (Induktionsannahme) für Polynome vom Grad < 1 die Aussage schon gezeigt, so sind zwei Fälle zu unterscheiden:

– Ist p irreduzibel, dann liegt wieder wie eben ein triviale Form der behaupteten Darstellung vor.

– Zerfällt dagegen p in zwei nichttriviale Faktoren $p = p_1 p_2$, mit $\deg p_1 < n$, $\deg p_2 < n$, so können wir nach Induktionsannahme jeden der Faktoren auf die genannte Weise als Produkt von irreduziblen darstellen, die Darstellungen multiplizieren und evtl. doppelt auftretende Faktoren zusammenfassen.

Eindeutigkeit: Wir nehmen an, das Polynom besitze zwei Zerlegungen in irreduzible Faktoren:

$$p(T) = \prod_{i=1}^{m} q_i(T)^{\nu_i} = \prod_{j=1}^{k} r_j(T)^{\mu_j}.$$

Ist für ein i und ein j gerade $q_i = r_j$, so bleibt die rechte Gleichung erhalten, wenn wir diesen gemeinsamen Faktor herauskürzen und dies können wir solange iterieren, bis beide Seiten keine gemeinsamen Faktor mehr enthalten. Bleibt auf beiden Seiten nur das leere Produkt übrig, so waren die Zerlegungen — bis auf die Reihenfolge — identisch. Waren sie dagegen echt verschieden, so tritt etwa links ein Faktor q_1 auf, der von allen noch rechts vorhandenen Faktoren r_j verschieden ist. Da er irreduzibel ist, und das noch rechts stehende Produkt teilt, muß er nach dem vorigen Satz einen Faktor teilen und so mit einem der r_j übereinstimmen, was aber doch gar nicht mehr geht. Somit könne die Darstellungen nicht echt verschieden sein. □

1

Zerlegung eines Raumes nach einem Endomorphismus I

Im Weiteren sei stets V ein endlich-dimensionaler Vektorraum $\neq (0)$ über einem beliebigen Körper K.

Definition Z.16 *Zu $f \in \mathrm{Hom}(V, V)$ heißt ein Unterraum $U \subset V$ f-invariant, wenn $f(U) \subset U$.*

Beispiele sind etwa Eigenräume von f, der ganze Raum V, aber auch der triviale Raum (0). Ein f-invarianter Unterraum U enthält mit einem Element u auch alle Elemente der Form $f(u), f^2(u), f^3(u), \ldots$ und deren Linearkombinationen, somit für jedes Polynom $p \in K[T]$ auch das Bild $(p(f))(u)$ von u unter dem Homomorphismus $p(f)$. Damit haben wir

Lemma Z.17 *Ist U f-invariant, so ist U für jedes $p \in K[T]$ auch $p(f)$-invariant.*

Die Betrachtung solcher Polynome in f führt uns überhaupt hier weiter:

Lemma Z.18 *Zu $p \in K[T]$ ist $U := \ker p(f)$ ein f-invarianter Unterraum.*

Beweis: Für $u \in \ker p(f)$ ist

$$p(f)(f(u)) = (p(f)f)(u) = (fp(f))(u) = f(p(f)(u)) = f(0) = 0,$$

somit auch $f(u) \in \ker p(f)$. □

Es sind zwar i.a. nicht alle f-invarianten Unterräume von V solche Kerne, aber sie liegen alle in solchen:

Satz und Definition Z.19 *Wir setzen $\mathcal{J}(f) := \{p \in K[T] \mid p(f) = 0\}$. Dann ist $\mathcal{J}(f) \neq \{0\}$ und das normierte Polynom ψ_f kleinsten Grades in $\mathcal{J}(f)$ heißt das "Minimalpolynom von f auf V". Es ist eindeutig bestimmt und teilt jedes Polynom aus $\mathcal{J}(f)$. Man nennt $\mathcal{J}(f)$ das von f erzeugte "Ideal".*

Beweis: Da V endlich-dimensional ist, ist auch $\mathrm{Hom}(V,V)$ endlich-dimensional. Somit gibt es zu der Folge $f^0 = \mathrm{id}_V, f^1 = f, f^2, f^3, ...$ in $\mathrm{Hom}(V,V)$ eine Nummer $m > 0$, sodaß $f^m \in \mathrm{span}(f^0, f^1, f^2, ..., f^{m-1})$, d.h. mit gewissen $\alpha_i \in K$

$$f^m = \sum_{i=0}^{m-1} \alpha_i f^i$$

oder $p(f) = 0$ für das nichttriviale Polynom

$$p(T) := T^m - \sum_{i=0}^{m-1} \alpha_i T^i \in K[T].$$

Also ist $p \in \mathcal{J}(f)$.

Sei nun $\psi \in \mathcal{J}(f)$, normiert, von minimalem Grad, ferner $p \in \mathcal{J}(f)$ beliebig, $\neq 0$. Dann ist $\deg p \geq \deg \psi$ und per Division mit Rest ist

$$p(T) = q(T)\psi(T) + r(T),$$

wobei $\deg r < \deg \psi$. Da $p, \psi \in \mathcal{J}(f)$ sind, ist $p(f) = \psi(f) = 0$, also auch $r(f) = 0$, somit $r \in \mathcal{J}(f)$. Da es kleineren Grad als ψ hat, bleibt dann nur $r = 0 \in K[T]$, sodaß ψ ein Teiler von p ist. Damit folgt auch die Eindeutigkeit. □

Bemerkung Z.20 *Abgesehen von dem trivialen und deshalb im allgemeinen ausgeschlossenen Fall $V = (0)$ ist stets $\deg \psi_f \geq 1$.*

Bezeichnung Z.21 *Ist $U \subset V$ ein f-invarianter Unterraum, so ist $f' := f_{|U} \in \mathrm{Hom}(U,U)$. Dann nennen wir das Minimalpolynom ψ' von f' auch kurz das "Minimalpolynom von f auf U" oder, wenn f aus dem Kontext klar ist, nur "Minimalpolynom zu U".*

Über invariante Unterräume und ihre Minimalpolynome gilt der folgende

Satz Z.22 *Es sei $f \in \mathrm{Hom}(V,V)$ mit Minimalpolynom ψ.*

(i) Ist $U \subset V$ f-invariant, ψ' das Minimalpolynom von f auf U, so ist ψ' ein Teiler von ψ.

(ii) Für jedes $p \in K[T]$ gilt: Das Minimalpolynom ψ' zu f auf dem Unterraum $\ker p(f)$ teilt p.

(iii) Ist $\psi = p_1 p_2$ mit normierten teilerfremden (nichttrivialen) Polynomen p_1, p_2 und ist $U_i := \ker p_i(f)$, $(i = 1, 2)$, so ist $V = U_1 \oplus U_2$ und p_i das Minimalpolynom von f auf U_i:

Beweis:

(i) Es ist $\psi(f) = 0 \in \mathrm{Hom}(V, V)$, d.h. $\psi(f)(v) = 0$ für alle $v \in V$ und damit auch für alle $v \in U$. Damit ist für $f' = f_{|U}, \psi \in \mathcal{J}(f')$ also Vielfaches von ψ'.

(ii) $U := \ker p(f)$ ist nach Lemma Z.18 f-invariant und per definitionem verschwindet $p(f)$ auf U, sodaß $p \in \mathcal{J}(f_{|U})$.

(iii) Da p_1 und p_2 teilerfremd sind, gibt es nach Satz Z.8 Polynome s_1, s_2, sodaß

$$1 = s_2(T)p_2(T) + s_1(T)p_1(T)$$

und damit auch

$$\mathrm{id}_V = s_2(f)p_2(f) + s_1(f)p_1(f).$$

Setze für $v \in V$: $u_1 := s_2(f)p_2(f)(v), \quad u_2 := s_1(f)p_1(f)(v)$. Dann ist

$$p_1(f)(u_1) = p_1(f)s_2(f)p_2(f)(v) = s_2(f)(p_1(f)p_2(f))(v) = s_2(f)\psi(f)(v) = 0,$$

und analog

$$p_2(f)(u_2) = 0, \text{ d.h. } u_i \in U_i \ (i = 1, 2).$$

Ferner ist

$$v = \mathrm{id}_V(v) = s_2(f)p_2(f)(v) + s_1(f)p_1(f)(v) = u_1 + u_2,$$

also $V = U_1 + U_2$.

Schließlich ist diese Summe direkt, da die Räume komplementär sind; denn für $u \in U_1 \cap U_2$ ist $p_1(f)(u) = p_2(f)(u) = 0$, somit

$$u = s_2(f)p_2(f)(u) + s_1(f)p_1(f)(u) = 0,$$

d.h. $u = 0$.

Bleibt noch die Aussage über die Minimalpolynome zu zeigen:

Als Kerne sind die U_i f-invariant und die Minimalpolynome ψ_i von f auf U_i nach 2. Teiler von p_i. Mit der direkten Summenzerlegung $u = u_1 + u_2$ folgt

$$\psi_1(f)\psi_2(f)(u) = \psi_2(f)\psi_1(f)(u_1) + \psi_1(f)\psi_2(f)(u_2) = 0,$$

sodaß $\psi_1\psi_2 \in \mathcal{J}(f)$.

Damit haben wir also: $\psi = p_1p_2$ teilt $\psi_1\psi_2$, ψ_1 teilt p_1, ψ_2 teilt p_2 und alle sind normiert. Dann ist aber notwendig $p_i = \psi_i (i = 1, 2)$. □

Damit erhalten wir einen ersten Zerlegungssatz:

Satz Z.23 *Es sei $f \in \mathrm{Hom}(V, V)$ mit Minimalpolynom*

$$\psi(T) = \prod_{i=1}^{k} \psi_i(T), \text{ wobei } \psi_i(T) = \varphi_i^{\nu_i}(T)$$

mit $\nu_i \geq 1$, und die φ_i paarweise verschiedene, (nichttriviale,) normierte, irreduzible Polynome sind. Dann ist mit $U_i := \ker \psi_i(f)$

$$V = U_1 \oplus U_2 \oplus ... \oplus U_k$$

eine direkte Summenzerlegung in (nichttriviale) f-invariante Unterräume. Die zugehörigen Minimalpolynome von f sind genau die $\psi_i(T)$.

Wählt man zu jedem U_i eine Basis B_i so ergeben diese zusammen eine Basis B von V. Die Matrix-Darstellung F von f bezüglich dieser Basis B hat dann die Block-Diagonal-Gestalt

$$F = \mathrm{diag}(F_1, F_2, ..., F_k),$$

wobei die F_i die Matrixdarstellungen von $f_{|U_i}$ bezüglich der Basen B_i sind.

Beweis: Die erste Aussage folgt durch rekursive Anwendung von Satz Z.22, der Rest ist trivial. □

Betrachten wir einen Spezialfall:
Das Minimalpolynom ψ habe die Gestalt

$$\psi(T) = \prod_{i=1}^{k} (T - \lambda_i)$$

mit paarweise verschiedenen $\lambda_i \in K$. Es ist also, bezogen auf den vorigen Satz, für alle i: $\psi_i(T) = \varphi_i(T) = T - \lambda_i$, $\nu_i = 1$. Dann ist $u \in U_i := \ker(f - \lambda_i \,\mathrm{id})$ genau wenn $fu = \lambda_i u$, die U_i sind also genau die Eigenräume von f zu den verschiedenen Eigenwerten λ_i und jede Matrixdarstellung von $f_{|U_i}$ hat die Gestalt

$$F_i = \mathrm{diag}(\lambda_i, \lambda_i, ..., \lambda_i).$$

Nach Satz Z.23 ist also f diagonalisierbar.

Ist umgekehrt f diagonalisierbar und sind dabei $\lambda_1, ..., \lambda_k$ die *verschiedenen* Eigenwerte von f, so gibt es eine Basis aus Eigenvektoren und für jeden solchen Basisvektor v ist dann $f(v) = \lambda_i v$ für eines der λ_i. Somit gilt mit dem Polynom

$$\psi(T) := (T - \lambda_1) \cdots (T - \lambda_k) :$$

$\psi(f)(v) = 0$ für jeden Basisvektor, also auf ganz V. Damit ist $\psi(T) \in \mathcal{J}(f)$, woraus man leicht erhält, daß es sogar genau das Minimalpolynom von f ist.

Wir haben also gezeigt:

Satz Z.24 *$f \in \mathrm{Hom}(V, V)$ ist genau dann diagonalisierbar, wenn das Minimalpolynom ψ die Form $\psi(T) = \prod_{i=1}^{k}(T - \lambda_i)$ hat, wobei die $\lambda_i \in K$ paarweise verschieden sind.*

Zerlegung f-zyklischer Räume

Unter den f-invarianten Unterräumen gibt es besonders einfache, die aus einem Element entstehen, die "f-zyklischen". Nach Lemma Z.17 enthält ein f-invarianter Unterraum U mit jedem Element u auch alle Elemente der Form $p(f)(u)$ für $p \in K[T]$. Diese bilden aber selbst schon einen f-invarianten Unterraum.

Satz und Definition Z.25 *Für jeden Vektor $v \in V$ ist*

$$U_v := \{p(f)(v) \mid p \in K[T]\}$$

ein f-invarianter Unterraum. Wir nennen ihn (vorläufig) den "f-zyklischen Unterraum zu v".

Beweis: Sind $u_i := p_i(f)(v) \in U_v$, $\alpha_i \in K$ $(i = 1, 2)$, so ist

$$\alpha_1 u_1 + \alpha_2 u_2 = (\alpha_1 p_1(f) + \alpha_2 p_2(f))(v) = ((\alpha_1 p_1 + \alpha_2 p_2)(f))(v) \in U_v,$$

da $\alpha_1 p_1 + \alpha_2 p_2 \in K[T]$. Also ist U_v ein Unterraum. Da mit $p(f)$ auch $fp(f) = p(f)f$ ein "Polynom im f" ist, ist er f-invariant. □

Daß ein solcher Raum schon durch f und ein einziges Element v bestimmt ist, hat eine Reihe von Konsequenzen.

Satz Z.26 *Es sei U_v der f-zyklische Unterraum zu $v \neq 0$. Dann gilt für Polynome $q \in K[T]$: Genau wenn $q(f)(v) = 0$ ist, gilt $q(f)(u) = 0$ für alle $u \in U_v$. Insbesondere ist damit das Minimalpolynom ψ_v von f auf U_v das normierte Polynom kleinsten Grades, für das $p(f)(v) = 0$. Wir nennen ψ_v auch kurz "Minimalpolynom zu v".*

Ferner ist der Grad $m := \deg \psi_v$ des Minimalpolynoms gleich der Dimension von U_v und $(v, f(v), f^2(v), ..., f^{m-1}(v))$ eine Basis.

Beweis: Jedes $u \in U_v$ hat die Gestalt $u = p(f)(v)$ mit einem $p \in K[T]$. Damit folgt: Ist $q(f)(v) = 0$, so ist $q(f)(u) = q(f)p(f)(v) = p(f)q(f)(v) = p(f)(0) = 0$.

Wegen $v \in U_v$ ist die Umkehrung trivial. Hieraus ergibt sich auch die Charakterisierung des Minimalpolynoms:

Das Minimalpolynom ψ_v habe die Form

$$\psi_v(T) = T^m - \sum_{i=0}^{m-1} \alpha_i T^i.$$

Dann ist

$$0 = \psi_v(f)(v) = (f^m - \sum_{i=0}^{m-1} \alpha_i f^i)(v)$$

oder

$$f^m(v) = \sum_{i=0}^{m-1} \alpha_i f^i(v) \in \text{span}(v, f(v), f^2(v), ..., f^{m-1}(v)).$$

Damit ist

$$\begin{aligned} f^{m+1}(v) = f(f^m)(v) &\in \text{span}(f(v), f^2(v), ..., f^{m-1}(v), f^m(v)) \\ &\subset \text{span}(v, f(v), f^2(v), ..., f^{m-1}(v)), \text{ da} \\ f^m(v) &\in \text{span}(v, f(v), f^2(v), ..., f^{m-1}(v)). \end{aligned}$$

Dies läßt sich iterieren, sodaß für alle $n = 0, 1, 2, ...$

$$f^n(v) \in \text{span}(v, f(v), f^2(v), ..., f^{m-1}(v)).$$

Nach Definition von U_v ist damit

$$U_v = \text{span}(v, f(v), f^2(v), ..., f^{m-1}(v)).$$

Diese Elemente sind auch linear unabhängig ($v \neq 0$!). Denn wäre $0 = \sum_{i=0}^{m-1} \beta_i f^i(v)$ wobei nicht alle $\beta_i = 0$, und dazu $k := \max\{i \mid \beta_i \neq 0\}$, so wäre wegen $v \neq 0$ jedenfalls $1 \leq k < m$ und Œ könnten wir $\beta_k = -1$ annehmen. Dann hätten wir aber eine Gleichung

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

sodaß mit

$$p(T) := T^k - \sum_{i=0}^{k-1} \beta_i T^i$$

auch $p(f)(v) = 0$ gelten würde, im Widerspruch dazu, daß das kleinste nichttriviale Polynom φ mit $\varphi(f)(v) = 0$ den Grad m hat. □

Für solch f-zyklischen Räume gilt natürlich auch die Zerlegung nach Satz Z.23. Schauen wir dies genauer an!

Satz Z.27 *Es sei U_v ein f-zyklischer Raum, zerlegt in f-invariante Unterräume U_i:*

$$U_v = U_1 \oplus U_2.$$

Dann gelten:

(i) Die U_i sind selbst f-zyklisch und ist $v = u_1 + u_2$ mit $u_i \in U_i$, so ist U_i erzeugt von u_i, d.h. $U_i = U_{u_i}$.

(ii) Für die entsprechenden Minimalpolynome ψ, ψ_1, ψ_2 gilt

$$\psi = \psi_1 \psi_2.$$

Beweis:

1. Wegen $v = u_1 + u_2$ ist für jedes Polynom $q \in K[T]$

$$q(f)(v) = q(f)(u_1) + q(f)(u_2) \ \in U_{u_1} + U_{u_2},$$

sodaß $U_v \subset U_{u_1} + U_{u_2}$. Nach Lemma Z.17 ist $U_{u_i} \subset U_i$, da die U_i f-invariant sind, ferner ist $U_1 \oplus U_2 = U_v$. Damit ist notwendig $U_i = U_{u_i}$ $(i = 1, 2)$.

2. Betrachte das Polynom $p := \psi_1 \psi_2$. Es ist

$$p(f)(v) = \psi_2(f)\psi_1(f)(u_1) + \psi_1(f)\psi_2(f)(u_2) = 0 + 0 = 0,$$

sodaß $p(f)$ ganz U_v anulliert. Damit ist ψ ein Teiler von $\psi_1\psi_2$. Sind andrerseits m, m_1, m_2 die Grade der Minimalpolynome, so ist nach Satz Z.26 dann $m = \dim U_v$, $m_i = \dim U_i$ und wegen $U_1 \oplus U_2 = U_v$ auch $m_1 + m_2 = m$. Somit haben ψ und $\psi_1\psi_2$ den selben Grad, sind beide normiert und stimmen wegen der gezeigten Teilbarkeit also überein. □

Definition Z.28 *Wir nennen einen f-invarianten Unterraum U "irreduzibel", wenn jede Zerlegung $U = U_1 \oplus U_2$ als direkte Summe f-invarianter Räume trivial ist, d.h. einer der Räume der Nullraum (0) ist.*

Satz Z.29 *(i) Ein f-zyklischer Raum U_v ist genau dann irreduzibel, wenn sein Minimalpolynom eine Potenz eines irreduziblen Polynoms ist.*

(ii) Besitzt der f-zyklische Raum U eine Zerlegung $U = U_1 \oplus U_2$ in f-invariante Räume, wobei U_1 irreduzibel $(\neq (0))$ ist, so hat sein Minimalpolynom ψ die Form

$$\psi = \varphi^\nu \chi,$$

wobei φ irreduzibel und teilerfremd zu χ. Dabei sind φ^ν und χ die Minimalpolynome zu U_1 bzw. U_2.

Beweis:

(i) Hat das Minimalpolynom (nichttriviale) teilerfremde Faktoren, so kann man nach Satz Z.22 den Raum nichttrivial in f-invariante Räume zerlegen.

Ist umgekehrt das Minimalpolynom von U_v Potenz eines irreduziblen und $U_v = U_1 \oplus U_2$ Zerlegung in f-invariante Räume, dann gilt nach Satz Z.27, daß mit $v = u_1 + u_2$ $(u_i \in U_i)$ dann $U_i = U_{u_i}$. Ferner gilt für die Minimalpolynome ψ, ψ_1, ψ_2 daß $\psi = \psi_1\psi_2$.

Ist also $\psi = \varphi^\nu$ mit irreduziblem φ, so ist $\psi_i = \varphi^{\nu_i}$ mit $0 \le \nu_i \le \nu,\ \nu_1 + \nu_2 = \nu$. Sei Œ $\nu_1 \le \nu_2$. Dann ist

$$\psi_2(f)(v) = \psi_2(f)(u_1) + \psi_2(f)(u_2) = \varphi^{\nu_2 - \nu_1}\psi_1(f)(u_1) + \psi_2(f)(u_2) = 0 + 0 = 0.$$

Nach Satz Z.26 ist dann das Minimalpolynom ψ von U_v ein Teiler von ψ_2, somit $\nu \le \nu_2$, d.h. $\nu_2 = \nu,\ \nu_1 = 0$.

U_v und U_2 haben also das selbe Minimalpolynom, als zyklische Räume nach Satz Z.26 damit die selbe Dimension, sodaß die Zerlegung $U_v = U_1 \oplus U_2$ trivial ist, also U_v irreduzibel.

(ii) Sind ψ_1, ψ_2 die Minimalpolynome von U_1, U_2, so ist nach Satz Z.27 $\psi = \psi_1\psi_2$ und nach Teil 1 hat ψ_1 die Gestalt $\psi_1 = \varphi^\nu$ mit einem irreduziblen Polynom φ. Somit bleibt nur noch zu zeigen, daß φ kein Teiler von ψ_2 ist.

Nehmen wir an, es sei $\psi_2 = \varphi\alpha$. Sei wieder $v = u_1 + u_2$ entsprechend den Unterräumen zerlegt. Dann ist

$$\begin{aligned}\varphi^\nu(f)\alpha(f)(v) &= \alpha(f)\varphi^\nu(f)(u_1) + \varphi^{\nu-1}(f)(\varphi(f)\alpha(f))(u_2)\\ &= \alpha(f)\psi_1(f)(u_1) + \varphi^{\nu-1}(f)\psi_2(f)(u_2) = 0.\end{aligned}$$

Somit ist das U_v-Minimalpolynom ψ ein Teiler von $\varphi^\nu\alpha$, d.h. $\psi = \varphi^\nu(\varphi\alpha) = \varphi^{\nu+1}\alpha$ teilt $\varphi^\nu\alpha$, was unmöglich ist. □

Damit bekommen wir den

Satz Z.30 (Zerlegungssatz für f-zyklische Räume) *Jeder f-zyklische Raum U besitzt eine Zerlegung als direkte Summe irreduzibler f-invarianter Unterräume:*

$$U = U_1 \oplus ... \oplus U_k.$$

Diese Zerlegung ist eindeutig abgesehen von der Reihenfolge und von trivialen Summanden. Die U_i sind selbst f-zyklisch. Diese Zerlegung entspricht der Zerlegung des Minimalpolynoms ψ in irreduzible Faktoren:

$$\psi(T) = \prod_{i=1}^{k} \psi_i(T),$$

wobei die ψ_i von der Form $\psi_i = \varphi_i^{\nu_i}$, mit φ_i irreduzibel, normiert und paarweise verschieden. Bei richtiger Nummerierung ist dann $\varphi_i^{\nu_i}$ das Minimalpolynom von U_i.

Beweis: Satz Z.23 liefert eine direkte Summenzerlegung gemäß der Darstellung des Minimalpolynoms in

$$U = U_1 \oplus ... \oplus U_k,$$

wobei das Minimalpolynom zu U_i gerade $\psi_i = \varphi_i^{\nu_i}$.

Da U f-zyklisch ist, sind nach Satz Z.27 die U_i ebenfalls zyklisch und, da die Minimalpolynome Potenzen von irreduziblen sind, sind diese Räume nach Satz Z.29 irreduzibel. Damit ist die Existenz gezeigt.

Die Eindeutigkeit folgt aus dem zweiten Teil von Satz Z.29, wonach jeder irreduzible (f-zyklische) Summand als Minimalpolynom genau einen solchen Faktor ψ_i aus der Darstellung des Minimalpolynoms haben muß, und die restlichen Summanden zusammen als Minimalpolynom das Produkt der restlichen Faktoren, was rekursiv die Behauptung ergibt. □

Matrix-Darstellungen auf f-zyklischen Räumen

Nach Satz Z.23 liefert eine Zerlegung des Raumes in eine direkte Summe f-invarianter, irreduzibler Räume eine Matrixdarstellung in Block-Diagonalform, wobei jeder Block die Einschränkung von f auf einen solchen Summanden-Raum beschreibt.

In der Situation von Satz Z.30 sind dies alles irreduzible f-zyklische Räume mit einem Minimalpolynom der Gestalt $\psi = \varphi^\nu$, wobei φ normiert, irreduzibel ist. Hat für einen solchen irreduziblen Raum U_v das Minimalpolynom die Form

$$\psi(T) = T^m - \sum_{i=0}^{m-1} \alpha_i T^i,$$

so wissen wir nach Satz Z.26, daß

$$(v, f(v), ..., f^{m-1}(v))$$

Basis von U_v ist, wobei wegen $\psi(f)(v) = 0$ noch

$$f(f^{m-1}(v)) = f^m(v) = \sum_{i=0}^{m-1} \alpha_i f^i(v).$$

Numeriert man diese Basis von links nach rechts, d.h. setzt

$$(v_1, v_2, ..., v_m) := (v, f(v), ..., f^{m-1}(v)),$$

so erhält man als Matrixdarstellung von f auf U_v offenbar

$$f \mathrel{\hat{=}} F = \begin{pmatrix} 0\,0\,0\,... & 0 & \alpha_0 \\ 1\,0\,0\,... & 0 & \alpha_1 \\ 0\,1\,0\,... & 0 & \alpha_2 \\ .\ .\ .\ ... & . & . \\ .\ .\ .\ ... & . & . \\ 0\,0\,0\,... & 1 & \alpha_{m-1} \end{pmatrix}_{(m,m)}$$

Bezeichnung Z.31 *Man nennt die eben erhaltene Matrix F die "Begleitmatrix" zum Polynom $\psi(T) = T^m - \sum_{i=0}^{m-1} \alpha_i T^i$.*

Dieser Name wird noch deutlicher durch den leicht zu verifizierenden

Satz Z.32 *Ist F die Begleitmatrix zu dem (normierten) Polynom $p(T)$, so ist $p(T) = \pm \det(F - T \cdot I)$, d.h. jedes normierte Polynom ist das charakteristische Polynom seiner Begleitmatrix.*

Bemerkung Z.33 *Es ist auch gebräuchlich, die oben verwendete Basis in der anderen Reihenfolge zu numerieren, sodaß man bezeichnet*

$$(v'_1, v'_2, ..., v'_m) := (f^{m-1}(v), f^{m-2}(v), ..., f(v), v).$$

Dann erhält man als Matrixdarstellung von f die andere Form der Begleitmatrix :

$$f \hat{=} F' = \begin{pmatrix} \alpha_{m-1} & 1 & 0 & ... & 0 & 0 \\ \alpha_{m-2} & 0 & 1 & ... & 0 & 0 \\ . & . & . & ... & . & . \\ \alpha_2 & 0 & 0 & ... & 1 & 0 \\ \alpha_1 & 0 & 0 & ... & 0 & 1 \\ \alpha_0 & 0 & 0 & ... & 0 & 0 \end{pmatrix}_{(m,m)}$$

Bei dieser Darstellung haben wir noch nicht benutzt, daß das Minimalpolynom ψ die spezielle Form $\psi = \varphi^\nu$hat. Dem kann man mit der Wahl einer speziellen Basis Rechnung tragen:

Sei $k := \deg \varphi$; damit ist dann $m := \deg \psi = k\nu = \dim U_v$. Wir setzen $g := \varphi(f)$ und betrachten die m Elemente

$$(v, f(v), ..., f^{k-1}(v), g(v), fg(v), ..., f^{k-1}g(v), g^2(v), fg^2(v), ..., f^{k-1}g^{\nu-1}(v)).$$

Diese sind linear unabhängig. Um dies zu sehen, seien $(\alpha_0, ..., \alpha_{m-1})$ die Koeffizienten einer sie annullierenden Linearkombination. Dazu betrachte das Polynom

$$p(T) := \alpha_0 + \alpha_1 T + ...\alpha_{k-1}T^{k-1} + \alpha_k\varphi(T) + \alpha_{k+1}T\varphi(T) + ... + \alpha_{m-1}T^{k-1}\varphi^{\nu-1}(T).$$

Nach Konstruktion ist $p(f)(v) = 0$, andrerseits ist sein Grad

$$\deg p \leq k - 1 + (\nu - 1)k = \nu k - 1 = m - 1 < m,$$

also kleiner als der des Minimalpolynoms, sodaß $p(T) = 0 \in K[T]$, somit $\alpha_0 = ... = \alpha_{m-1} = 0$.

Stellen wir f bezüglich dieser Basis (von links nach rechts numeriert) dar! Dazu ist für jedes Basis-Element v_i das Bild $f(v_i)$ wieder durch die Basis auszudrücken.

Sofern v_i von der Form

$$v_i \;=\; f^\kappa g^\mu(v) \text{ mit } 0 \leq \kappa < k - 1, 0 \leq \mu < \nu$$

ist $f(v_i) \;=\; f^{\kappa+1}g^\mu(v) = v_{i+1}$, d.h. einfach das nächste Basis-Element. Ihm entspricht eine Spalte mit genau einer 1, die gerade unter der Diagonalen steht, und sonst alles $= 0$.

Dann bleiben noch zu betrachten die Basiselemente der Form

$$f^{k-1}g^\mu(v) \;\text{ für } 0 \leq \mu < \nu.$$

Mit

$$\varphi(T) = T^k - \sum_{i=0}^{k-1} \gamma_i T^i$$

ist

$$g = \varphi(f) = f^k - \sum_{i=0}^{k-1} \gamma_i f^i,$$

also

$$f^k = g + \sum_{i=0}^{k-1} \gamma_i f^i.$$

Damit ist

$$f(f^{k-1}g^\mu(v)) = f^k g^\mu(v) = (g + \sum_{i=0}^{k-1} \gamma_i f^i)g^\mu(v) = g^{\mu+1}(v) + \sum_{i=0}^{k-1} \gamma_i f^i g^\mu(v),$$

wobei $g^\mu(v)$ das nächste Basis-Element ist, sofern $\mu+1 < \nu$ bzw. $g^{\mu+1}(v) = g^\nu(v) = \varphi^\nu(f)(v) = 0$, im Falle $\mu + 1 = \nu$. Also taucht hier eine Spalte auf von der Form

$$\begin{pmatrix} \vdots \\ 0 \\ \gamma_0 \\ \vdots \\ \gamma_{k-1} \\ 1 \\ \vdots \end{pmatrix},$$

wobei die Eintragung 1 wieder genau unter der Diagonalen steht. (Bei dem letzten Term taucht die 1 nicht mehr auf.)

Dies bedeutet:

Satz Z.34 *Hat der f-zyklische Unterraum das Minimalpolynom $\psi(T) = \varphi^\nu(T)$ und ist*

$$G = \begin{pmatrix} 0\,0\,0\,... & 0 & \gamma_0 \\ 1\,0\,0\,... & 0 & \gamma_1 \\ 0\,1\,0\,... & 0 & \gamma_2 \\ .\ .\ .\ ... & . & . \\ .\ .\ .\ ... & . & . \\ 0\,0\,0\,... & 1 & \gamma_{k-1} \end{pmatrix}_{(k,k)}$$

die Begleitmatrix zu $\varphi(T)$, so läßt sich f auf U_v darstellen durch

$$f \mathrel{\hat{=}} F = \begin{pmatrix} G & 0 & 0 & ... & 0 & 0 \\ E & G & 0 & ... & 0 & 0 \\ 0 & E & G & ... & 0 & 0 \\ . & . & . & ... & . & . \\ . & . & . & ... & . & . \\ 0 & 0 & 0 & ... & E & G \end{pmatrix}_{(\nu,\nu)},$$

wobei

$$E = \begin{pmatrix} 0\,0\,...\,0\,1 \\ 0\,0\,...\,0\,0 \\ .\ .\ ...\ .\ . \\ .\ .\ ...\ .\ . \\ 0\,0\,...\,0\,0 \end{pmatrix}$$

und dabei insgesamt ν viele Blöcke G auftreten.

Bemerkung Z.35 *Numeriert man die hier benutzte Basis rückwärts, so erhält man eine analoge Darstellung, wobei zu ersetzen sind*

- die Matrix G durch die andere Begleitmatrix G' zu φ,

- die Matrix E durch eine Matrix E', die ihre einzige 1 nun in der linken unteren Ecke hat, und

- diese Matrizen E' nun oberhalb der Diagonale anzusiedeln sind.

Von besonderem Interesse ist der Spezialfall, daß der Körper algebraisch abgeschlossen ist. Dann hat jedes normierte irreduzible Polynom die Form

$$\varphi(T) = T - \lambda,$$

das Minimalpolynom eines irreduziblen f-zyklischen Raumes also die Gestalt

$$\psi(T) = (T - \lambda)^k.$$

Die Begleitmatrix zu φ ist dann eine 1×1 Matrix, nämlich

$$G = (\lambda),$$

und die Matrix F nach Satz Z.34 lautet

$$F = \begin{pmatrix} \lambda & 0 & 0 & \dots & 0 & 0 \\ 1 & \lambda & 0 & \dots & 0 & 0 \\ 0 & 1 & \lambda & \dots & 0 & 0 \\ . & . & . & \dots & . & . \\ . & . & . & \dots & . & . \\ 0 & 0 & 0 & \dots & 1 & \lambda \end{pmatrix}$$

oder wenn man anders numeriert

$$F' = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ . & . & . & \dots & . & . \\ . & . & . & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix},$$

was uns als ein Kasten der JORDANschen Normalform bekannt ist. Wir nennen daher die Darstellung aus Satz Z.34 auch verallgemeinerte JORDANmatrix.

Fassen wir nochmal zusammen, was bisher erreicht ist:

Satz Z.36 *(i) Einer Zerlegung $U = U_1 \oplus \dots \oplus U_n$ eines f-invarianten Raumes in f-invariante Unterräume entspricht eine Darstellung von f als eine Blockdiagonalmatrix $F = \operatorname{diag}(F_1, \dots, F_n)$, wobei die F_i Matrixdarstellungen für $f_{|U_i}$ sind.*

(ii) Ein f-invarianter f-zyklischer Raum U besitzt eine im wesentlichen eindeutig bestimmte Zerlegung $U = U_1 \oplus \dots \oplus U_n$ in irreduzible, f-invariante, f-zyklische Räume. Deren Minimalpolynome sind von der Form $\psi_i = \varphi_i^{\nu_i}$ mit paarweise verschiedenen irreduziblen Polynomen φ_i. Die dazu gemäß 1. gehörenden Blockmatrizen F_i können in verallgemeinerter JORDANform (Satz Z.34) gewählt werden.

Unsere bisherigen Ergebnisse sichern uns aber *noch nicht*, daß *jeder* f-invariante Raum eine Darstellung als direkte Summe f-zyklischer Räume besitzt, die man dann entsprechend 2. verfeinern könnte.

Wir könnten dies mit den bisher benutzten Methoden beweisen, wollen aber einen anderen Zugang wählen, der die Situation von einem allgemeineren Standpunkt aus betrachtet. Dazu gehen wir von Vektorräumen zu "Moduln" über, was bedeutet, daß der Skalarbereich jetzt ein Ring sein kann und nicht notwendig ein Körper sein muß, sodaß wir also bei den Skalaren auf das Dividieren verzichten.

Etwas über Ringe

Definition Z.37 (Kommutativer Ring mit 1) *Ein Ring besteht aus einer Menge R mit zwei ausgezeichneten Elementen $0, 1$ und zwei Operationen, genannt Addition und Multiplikation: $+, \cdot : R \times R \to R$, sodaß alle Gesetze für einen Körper gelten mit Ausnahme der Existenz der Inversen bezüglich der Multiplikation.*

Da wir für die Multiplikation die Kommutativität und die Existenz eines 1-Elementes gefordert haben, sprechen wir genauer von einem "kommutativen Ring mit 1".

Man spricht auch unter schwächeren Bedingungen an die Multiplikation noch von Ringen, doch sollen uns die hier nicht interesssieren.

Beispiele kommutativer Ringe mit 1 sind etwa:

1. Jeder Körper K.
2. Jede K-Algebra, etwa $K[T]$. Man redet daher auch vom Polynomring statt von der Polynomalgebra.
3. Die ganzen Zahlen $\mathbb{Z}$
4. Die Restklassenringe $\mathbb{Z}_{/n\mathbb{Z}}$.

Ein Beispiel für einen nicht-kommutativen Ring sind etwa die $n \times n$ Matrizen, einen Ring ohne 1-Element bilden etwa die geraden ganzen Zahlen. Solche Situationen hatten wir ausgeschlossen.

Im Ring $\mathbb{Z}_{/6\mathbb{Z}}$ sind 2 und 3 von Null verschiedene Elemente, ihr Produkt ist aber $= 0$. Sie sind sogenannte "Nullteiler". Ringe, in denen es so etwas gibt, seien auch noch ausgeschlossen, ebenso der triviale Ring, der nur aus dem Nullelement besteht.

Definition Z.38 (Integritäts-Ring) *Ein kommutativer Ring R mit 1 heißt ein "Integritäts-Ring", wenn*

(i) $1 \neq 0$ *und*

(ii) R ohne Nullteiler ist, d.h. ist $ab = 0$, so ist $a = 0$ oder $b = 0$. Hierfür sagt man auch "nullteilerfrei".

Vereinbarung Z.39 *Im weiteren stehe "Ring" stets für "Integritäts-Ring", sofern nichts anderes gesagt ist.*

Etwa $\mathbb{Z}$ und $K[T]$ für beliebige Körper K sind Integritätsringe. In diesen beiden Ringen kennen wir schon einiges über Teilbarkeit. Dies sei nun verallgemeinert.

Definition Z.40 (Teiler,Vielfache) *Es seien a, b, c Elemente eines Ringes R.*

(i) Gibt es ein c, sodaß $ac = b$, so sagt man"a teilt b","a ist Teiler von b" oder auch "b ist Vielfaches von a" und schreibt $a \mid b$.

(ii) Die Teiler von 1 nennt man "Einheiten". $R^\times$ bezeichnet die Gesamtheit der Einheiten von R.

(iii) a und b heißen "assoziiert", geschrieben $a \hat{=} b$, wenn $a \mid b$ und $b \mid a$.

(iv) a heißt "echter Teiler" von b, wenn a Teiler von b, aber a keine Einheit ist und auch nicht zu b assoziiert.

Die Einheiten in $\mathbb{Z}$ sind offenbar $\{+1, -1\}$, für einen Körper K ist $K^\times = K \setminus 0$, ferner ist $(K[T])^\times = K^\times$, d.h. die Einheiten in Polynomring sind genau die im Grundkörper.

Wir notieren einige Eigenschaften von Teilbarkeit als

Lemma Z.41 *Für die Teilbarkeit in (Integritäts-)Ringen gelten*

(i) $a \mid a$ *(reflexiv)*

(ii) *Aus* $a \mid b$ *und* $b \mid c$ *folgt* $a \mid c$. *(transitiv)*

(iii) $1 \mid a$ *und* $a \mid 0$.

(iv) *Aus* $a \mid b$ *folgt* $ac \mid bc$.

Zu diesen Aussagen war die Nullteilerfreiheit nicht nötig. Wir brauchen sie aber für

(v) *Aus* $ac \mid bc$ *folgt* $a \mid b$, *sofern* $c \neq 0$.

(vi) a *und* b *sind genau dann assoziiert, wenn* $a = \epsilon b$, *wobei* ϵ *eine Einheit ist.*

Wir **beweisen** nur die letzten beiden Aussagen.

Zu (v) Wird bc von ac geteilt, so gibt es ein $r \in R$, sodaß $bc = rac$, d.h. $0 = bc - rac = (b - ra)c$. Da $c \neq 0$ und der Ring ohne Nullteiler ist, muß $b - ra = 0$ sein, sodaß $b = ra$, also a Teiler von b.

Zu (vi) Zu einer Einheit ϵ gibt es nach Definition eine weitere Einheit ϵ', sodaß $\epsilon\epsilon' = 1$, beide also zueinander multiplikativ invers sind.

Sei $a = \epsilon b$: Dann gilt also $b \mid a$. Wir haben aber auch $\epsilon' a = \epsilon' \epsilon b = b$, also $a \mid b$. Somit sind sie assoziiert.

Sind umgekehrt a, b assoziiert, so gibt es Elemente $r, s \in R$ mit $ar = b, \quad bs = a$. Also ist $ars = bs = a = a \cdot 1$ und wegen der Nullteilerfreiheit ist also $rs = 1$, d.h. r und s sind Einheiten. Hierbei ist $a \neq 0$ vorausgesetzt. Behandeln Sie den Fall $a = 0$ selbst.

Wie bei den ganzen Zahlen und den Polynomen können wir zu zwei oder mehr Elementen gemeinsame Teiler und gemeinsame Vielfache betrachten und insbesondere nach "größten" gemeinsamen Teilern und nach "kleinsten" gemeinsamen Vielfachen fragen. Doch muß man damit rechnen, daß so etwas nicht unbedingt zu existieren braucht.

Definition Z.42 (ggT, kgV) *Es seien* $a_1, ..., a_n \in R$.

(i) $d \in R$ *heißt ein größter gemeinsamer Teiler (ggT) von* $a_1, ..., a_n$, *wenn*

(a) d *jedes* a_i *teilt und*

(b) *jeder gemeinsame Teiler* t *von* $a_1, ..., a_n$ *auch Teiler von* d *ist.*

(ii) $v \in R$ *heißt kleinstes gemeinsames Vielfaches (kgV) von* $a_1, ..., a_n$, *wenn*

(a) v *Vielfaches von jedem* a_i *ist und*

(b) *jedes gemeinsame Vielfache* w *von* $a_1, ..., a_n$ *auch Vielfaches von* v *ist.*

Lemma Z.43 *Es sei* d *ein ggT von* $a_1, ..., a_n$. *Dann gilt:* d' *ist ggT von* $a_1, ..., a_n$ *genau, wenn* d *und* d' *assoziiert sind.*

Analoges gilt für das kgV.

Beweis: Ist d' auch ein ggT, so gelten nach Definition Z.42 die beiden Teibarkeitsrelationen $d \mid d'$ und $d' \mid d$, sodaß d und d' assoziiert sind.

Sind d und d' assoziiert, so gelten beide Teilbarkeitsrelationen. Dann ist wegen $d' \mid d$ auch d' ein gemeinsamer Teiler. Andrerseits ist jeder gemeinsame Teiler von $a_1, ..., a_n$ auch Teiler von d, somit wegen $d \mid d'$ auch Teiler von d'. Folglich ist auch d' ein ggT. □

Die Menge

$$(a) := R \cdot a := \{ra \mid r \in R\}$$

aller Vielfachen von a liefert uns ein Beispiel für die folgende Definition:

Definition Z.44 (Ideal) *Eine nichtleere Teilmenge $\mathcal{I} \subset R$ heißt ein "Ideal" in dem kommutativen Ring R, wenn gelten*

(i) Sind $a, b \in \mathcal{I}$ so ist auch $a + b \in \mathcal{I}$.

(ii) Ist $a \in \mathcal{I}$ und $r \in R$ so ist auch $ra \in \mathcal{I}$.

Gibt es Elemente $a_1, ..., a_n \in \mathcal{I}$, sodaß $\mathcal{I} = \{\sum_1^n r_i a_i \mid r_i \in R\}$, so heißt $\mathcal{I}$ endlich erzeugt.

Das von einem Element $a \in R$ erzeugte Ideal wird mit (a) bezeichnet und heißt "Hauptideal" zu a.

Definition Z.45 (Hauptideal-Ring) *Ein (Integritäts-)Ring, in dem jedes Ideal ein Hauptideal ist, heißt "Hauptideal-Ring".*

Wir werden sehen, daß $\mathbb{Z}$ und die Polynomringe $K[T]$ solche Hauptideal-Ringe sind. In ihnen gelten Teilbarkeitseigenschaften, die recht ähnlich sind zu dem, was wir von $\mathbb{Z}$ kennen bzw. für $K[T]$ gezeigt haben. Daher sind solche Ringe für die Weiterführung unserer allgemeinen Fragestellung dieses Kapitels brauchbar.

Lemma Z.46 *Sind $\mathcal{I}_1, \mathcal{I}_2$ Ideale in R, so auch $\mathcal{I}_1 \cap \mathcal{I}_2$ und $\mathcal{I}_1 + \mathcal{I}_2 := \{a_1 + a_2 \mid a_i \in \mathcal{I}_i\}$*

Der triviale **Beweis** sei übergangen.

Übersetzen wir die Aussagen über Teilbarkeit in die Sprache der Ideale:

Lemma Z.47 *Über einem (Integritäts-)Ring gelten:*

(i) a teilt b genau, wenn $(b) \subset (a)$.

(ii) a und b sind assoziiert genau, wenn $(b) = (a)$.

(iii) v ist gemeinsames Vielfaches von a und b genau, wenn

$$(v) \subset (a) \cap (b)$$

und damit: v ist ein kgV von a und b genau, wenn

$$(v) = (a) \cap (b).$$

(iv) d ist gemeinsamer Teiler von a und b genau, wenn $(a) + (b) \subset (d)$.
(Die naheliegende Aussage über den ggT braucht nicht zu gelten, siehe aber Satz Z.48.)

Beweis:

(i) $a \mid b$ ist äquivalent zu : Es existiert ein c, sodaß $ac = b$. Haben wir dies, so folgt

$$b = ac \in (a), \text{ also } (b) \subset (a).$$

Ist umgekehrt $(b) \subset (a)$, so ist $b = ra$ mit einem $r \in R$, also a Teiler von b.

(ii) Trivial nach 1.

(iii) (a) und (b) enthalten genau die Vielfachen von a bzw. b. Also besteht $(a) \cap (b)$ genau aus den gemeinsamen Vielfachen von a und b. Insbesondere ist für jedes gemeinsame Vielfache v dann $(v) \subset (a) \cap (b)$.

Ist v ein kgV von a, b, so ist also $(v) \subset (a) \cap (b)$. Andrerseits ist ein gemeinsames Vielfaches w von a, b auch Vielfaches von v, woraus $(a) \cap (b) \subset (v)$ folgt und dann die Gleichheit erhalten ist.

Ist $(a) \cap (b) = (v)$ also insbesondere ein Hauptideal, von v erzeugt, so ist v ein gemeinsames Vielfaches, das trivialerweise jedes andere gemeinsame Vielfache teilt.

(iv) Ist d gemeinsamer Teiler, so ist $(a) \subset (d)$ und $(b) \subset (d)$ und damit dann auch $(a) + (b) \subset (d)$. Ist umgekehrt

$$(a) + (b) = \{ra + sb \mid r, s \in R\} \subset (d),$$

so ist auch $(a) \subset (d)$ und $(b) \subset (d)$, folglich d ein gemeinsamer Teiler. □

Achtung: Auch in schön aussehenden Ringen braucht es keine ggT. zu geben!

Es gibt sie jedoch in Hauptideal-Ringen.

Satz Z.48 *In einem Hauptideal-Ring R gibt es zu beliebigen Elementen $a_1, ..., a_n \in R$ stets einen ggT. Jeder ggT d besitzt eine Darstellung als*

$$d = r_1 a_1 + ... + r_n a_n \text{ mit } r_i \in R.$$

Beweis: Durch triviale Verallgemeinerung von Lemma Z.47.4 erhält man, daß t genau dann gemeinsamer Teiler ist, wenn

$$(a_1) + (a_2) + ... + (a_n) \subset (t.)$$

Da R ein Hauptideal-Ring ist, gibt es ein d, sodaß $(a_1) + (a_2) + ... + (a_n) = (d)$. Damit ist also d einmal ein gemeinsamer Teiler und für jeden weiteren t gilt, wie eben gezeigt, dann $(d) \subset (t)$, sodaß also d ein ggT ist.

Die Beziehung $(a_1) + (a_2) + ... + (a_n) = (d)$ bedeutet natürlich die Existenz der behaupteten Darstellung für den ggT.

□

Zur Darstellung vergleiche auch Satz Z.8.

Wir hatten oben gesagt, daß $\mathbb{Z}$ und die Polynomringe $K[T]$ Hauptideal-Ringe seien. Zeigen wir dies für die Polynomringe:

Lemma Z.49 *$K[T]$ ist Hauptideal-Ring.*

Beweis: Sei $\mathcal{I}$ ein Ideal in $K[T]$, Œ $\neq 0$. Wähle darin ein Polynom $\psi \neq 0$, von minimalem Grad. Dann ist $\mathcal{I} = (\psi)$: Denn für ein beliebiges Element $p \in \mathcal{I}$ können wir per Division mit Rest bilden

$$p = q\psi + r \quad \text{mit } \deg r < \deg \psi,$$

wobei, da $p, \psi \in \mathcal{I}$, auch $r = p - q\psi \in \mathcal{I}$ ist. Nun ist $\deg r < \deg \psi$ aber $\psi \in \mathcal{I}, \neq 0$ von minimalem Grad, sodaß notwendig $r = 0$ sein muß, also p ein Vielfaches von ψ ist, d.h. $p \in (\psi)$.

Folglich ist $\mathcal{I} = (\psi)$, also Hauptideal. □

In diesen Beweis ging von den spezifischen Eigenschaften des Polynomrings nur ein, daß wir eine Division mit Rest durchführen können, wobei der Rest kleineren Grad hat als der Divisor. Sowas geht auch in gewissen anderen Ringen:

Definition Z.50 (Euklidischer Ring) *Ein Integritätsring R heißt "euklidischer Ring", wenn eine Abbildung $\nu : R \to \mathbb{N}$ existiert mit $\nu(0) = 0$, sodaß gilt: Zu $a, b \in R, a \neq 0$ existieren $q, r \in R$, sodaß wir Division mit Rest machen können, d.h.*

$$b = qa + r \quad \text{wobei } \nu(r) < \nu(a).$$

ν heißt "euklidische Normfunktion" auf R.

Beispiele sind etwa $\mathbb{Z}$ mit dem gewöhnlichen Betrag als Normfunktion, oder der Polynomring $K[T]$ mit $\nu(0) = 0, \nu(p) = \deg p + 1$ für $p \neq 0$.

Für solche Ringe kann man offenbar den eben für die Polynome geführten Beweis sinngemäß wiederholen und erhält dann

Satz Z.51 *Jeder euklidische Ring ist Hauptideal-Ring.*

Bemerkung Z.52 *In euklidischen Ringen kann man den für Polynomringe studierten euklidischen (Divisions-)Algorithmus nachbauen und damit etwa den ggT berechnen und darstellen, etc.*

Wir wollen nun noch die von $\mathbb{Z}$ und den Polynomringen her bekannte Möglichkeit der Primfaktor-Zerlegung auf Hauptideal-Ringe ausdehnen. (Siehe Definition Z.13)

Definition Z.53 (irreduzibel, prim) *R sei ein (Integritäts-)Ring.*

(i) Ein Element $p \in R$ heißt "irreduzibel", wenn p nur triviale Teiler hat, d.h. wenn $p \notin R^\times$ und wenn aus $p = p_1 p_2$ folgt, daß $p_1 \in R^\times$ oder $p_2 \in R^\times$.

(ii) Ein Element $p \in R$ heißt "prim", wenn gilt: $p \notin R^\times$ und teilt p das Produkt $p_1 p_2$ so auch schon einen der Faktoren.

Über $\mathbb{Z}$ und $K[T]$ fallen die Begriffe "prim" und "irreduzibel" zusammen. In allgemeineren Ringen ist dies nicht so. Es gilt aber

Lemma Z.54 *(i) Über beliebigen (Integritäts-)Ringen sind Primelemente irreduzibel.*

(ii) Über Hauptideal-Ringen sind irreduzible Elemente auch prim. Hier fallen also die beiden Begriffe zusammen.

Beweis:

(i) sei wegen Trivialität übergangen.

(ii) Sei p irreduzibel, Teiler von ab, aber kein Teiler von a. Es sind dann p, a teilerfremd. Nach Satz Z.48 gibt es dann eine Darstellung $1 = rp + sa$. Damit gilt auch $b = rpb + sab$. Die rechte Seite wird hier von p geteilt, also auch die linke, d.h. $p \mid b$. □

Haben Sie diesen Beweis schon mal gesehen?

Definition Z.55 (Zerlegung in irreduzible Faktoren) *(i) Eine Darstellung*

$$a = \epsilon q_1 q_2 ... q_n \quad \text{mit } \epsilon \in R^\times, \quad q_i \text{ irreduzibel,}$$

nennen wir "Zerlegung von a in irreduzible Faktoren."

(ii) Gilt überdies: Sind für jede weitere solche Zerlegung von a

$$a = \epsilon' q_1' q_2' ... q_{n'}'$$

notwendig $n = n'$ und nach geeigneter Numerierung für alle i jeweils q_i und q_i' assoziiert, so sagen wir " a besitzt eine eindeutige Zerlegung in irreduzible Elemente."

(iii) Ein (Integritäts-)Ring, in dem jedes Element eine eindeutige Zerlegung in irreduzible Faktoren besitzt, heißt "faktoriell".

Über diese Zerlegung zeigen wir zunächst die folgende Eindeutigkeitsaussage:

Satz Z.56 *Es sei R ein Integritätsring, in dem jedes Element $a \neq 0$ eine Zerlegung in irreduzible Faktoren besitzt. Dann sind äquivalent*

(i) R ist faktoriell d.h. die Zerlegung ist eindeutig,

(ii) Jedes irreduzible Element ist prim.

Beweis:

$(i) \Rightarrow (ii)$: Sei p irreduzibel und Teiler von ab. Wir zerlegen

$$a = \alpha q_1 ... q_n, \quad b = \beta p_1 ... p_m,$$

Dann ist

$$ab = \alpha\beta q_1 ... q_n p_1 ... p_m$$

Zerlegung von ab in irreduzible Faktoren. Wegen der vorausgestzten Eindeutigkeit muß dann auch p zu einem dieser q_i oder p_i assoziiert sein, ist somit Teiler von a oder von b.

$(ii) \Rightarrow (i)$: Seien

$$a = \epsilon q_1 ... q_n = \epsilon' p_1 ... p_m$$

zwei Zerlegungen in irreduzible Faktoren und Œ $n > 0$. Dann ist q_1 irreduzibel, also prim und Teiler von $p_1 ... p_m$, teilt also einen der Faktoren rechts. Der sei Œ p_1, ist dann irreduzibel und somit zu q_1 assoziiert. Dann kann man q_1 und p_1 herauskürzen und induktiv weiterschließen. □

Hier hatten wir noch die Existenz der Zerlegung mit vorausgesetzt. Die ergibt sich mit bei folgendem

Satz Z.57 *Ein (Integritäts-)Ring R ist genau dann faktoriell, wenn die folgenden beiden Bedingungen erfüllt sind:*

(i) Jede Kette

$$(a_1) \subset (a_2) \subset ... \subset (a_n) \subset (a_{n+1}) \subset ...$$

von Hauptidealen ist schließlich stationär, d.h. es existiert ein Index n, sodaß $a_n = a_j$ für alle $j \geq n$. Man nennt dies die "aufsteigende Teilerketten-Bedingung".

(ii) Jedes irreduzible Element ist prim.

Beweis: Wir arbeiten dahin, Satz Z.56 anwenden zu können und zeigen : *Gilt (i), so hat jedes Element $a \neq 0$ eine Zerlegung in irreduzible Elemente:* Dazu betrachte die Menge M aller Hauptideale (a), die erzeugt werden von Elementen a, die keine solche Zerlegung besitzen. Wegen der Teilerketten-Bedingung besitzt M ein maximales Element, das ist hier ein Hauptideal (m), für das kein a existiert, sodaß $(a) \in M$ und $(m) \subsetneqq (a)$. Denn andernfalls könnten wir eine nicht stationäre Teilerkette konstruieren. So ein maximales Element (m) liefert aber einen Widerspruch. Denn ist m eine Einheit oder irreduzibel, so hat es ja eine solche Zerlegung, sodaß (m) nicht zu M gehören kann. Bleibt nur, daß $m = ab$ mit echten Teilern a, b, für die dann aber $(m) \subsetneqq (a)$, $(m) \subsetneqq (b)$. Da (m) maximal ist, sind also (a) und (b) nicht in M, d.h. a und b haben beide eine Zerlegung in irreduzible Faktoren, womit dann auch m eine solche hat, was aber nicht sein darf. Damit ist in allen Fällen ein Widerspruch erhalten.

Zusammen mit Satz Z.56 ergibt sich also:

Aus (i) und (ii) folgt, daß R faktoriell.

Sei nun R als faktoriell vorausgesetzt. Nach Satz Z.56 gilt jedenfalls (ii). Betrachte ein Element a, für das $(0) \neq (a) \neq (1)$, mit der Darstellung

$$a = \epsilon q_1 ... q_n \text{ mit irreduziblen Elementen } q_i.$$

Dann hat ein t mit $(a) \subsetneqq (t)$ bei geeigneter Numerierung eine Darstellung als

$$t = \epsilon q_1 ... q_m \text{ mit } 0 \leq m \leq n,$$

woraus sich sofort die Teilerketten-Bedingung ergibt. □

Zusammen mit Lemma Z.54 erhalten wir daraus insbesondere den für den Rest des Kapitels zentralen

Satz Z.58 *Jeder Hauptideal-Ring ist faktoriell.*

Beweis: Nach Lemma Z.54 ist jedes irreduzible Element prim. Wir zeigen die Teilerketten-Bedingung: Sei $(a_1) \subset (a_2) \subset ...$ eine aufsteigende Kette von Hauptidealen in R. Betrachte $\mathcal{I} := \bigcup_{i \in \mathbb{N}} (a_i)$. Dies ist, wie man leicht nachprüft, wieder ein Ideal in R, also, da wir in einem Hauptideal-Ring sind, ein Hauptideal. Somit existiert ein $a \in \bigcup_{i \in \mathbb{N}} (a_i)$, sodaß

$$(a) = \bigcup_{i \in \mathbb{N}} (a_i).$$

Dann gibt es natürlich eine Nummer m, sodaß $a \in (a_m)$, d.h. daß also für jedes j

$$(a_j) \subset \bigcup_{i \in \mathbb{N}} (a_i) = (a) \subset (a_m)$$

und diese Kette also spätestens ab Stelle m stationär ist. □

Matrizen über Hauptideal-Ringen

Im Zusammenhang mit der Matrixdarstellung von Vektorraum-Homomorphismen hatten wir in Satz H.41 gezeigt:

Über einem Körper K gibt es zu jeder Matrix $F \in K^{n \times m}$ invertierbare Matrizen R, S, sodaß $F' := RFS$ die spezielle Gestalt

$$F' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

hat. Dabei ist I_r die $r \times r$ Einheitsmatrix.

Für den Fortgang des Hauptthemas dieses Kapitels brauchen wir etwas ähnliches wie diese Normalform von Matrizen über Ringen. Speziell für Hauptideal-Ringe werden wir eine schöne solche Verallgemeinerung finden.

Zunächst einiges vorbereitende für Matrizen über Ringen, worunter wir wieder Integritätsringe verstehen wollen.

Zu einem Ring R bezeichne $R^{n \times m}$ die $n \times m$ Matrizen mit Elementen aus R. Addition, Multiplikation mit Skalaren d.h. Ringelementen und Multiplikation von Matrizen jeweils passender Formate sind wie über Körpern erklärt. $I_n \in R^{n \times n}$ bezeichne wieder die Einheitsmatrix, eine Matrix $A \in R^{n \times n}$ heißt invertierbar, wenn es eine Matrix B gibt, sodaß $AB = BA = I_n$.

Im Einklang mit den schon für Ringe eingeführten Bezeichnungen nennt man invertierbare Matrizen auch "Einheiten", Matrizen, die sich nur um eine solche Einheit als Faktor unterscheiden auch "assoziiert". Etwa über die Formeln des Entwicklungssatzes von LAPLACE können wir auch für Matrizen in $R^{n \times n}$ eine Determinante

erklären und wie über Körpern hat eine invertierbare Matrix eine nicht verschwindende Determinante. Doch reicht dies hier nicht für die Invertierbarkeit.

Dazu ein **Beispiel:** Es sei $R = \mathbb{Z}, A := \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$ mit $\det A = 2$.

Für eine Matrix $B = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}$ mit $AB = I$ müßte gelten

$$I = AB = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} = \begin{pmatrix} 2\beta_{11} + \beta_{21} & 2\beta_{12} + \beta_{22} \\ \beta_{21} & \beta_{22} \end{pmatrix},$$

sodaß notwendig $1 = \beta_{22}, 0 = 2\beta_{12} + \beta_{22} = 2\beta_{12} + 1$, was (zwar über $\mathbb{Q}$, aber) nicht über $\mathbb{Z}$ lösbar ist.

Die richtige Bedingung lautet:

Satz Z.59 *$A \in R^{n\times n}$ ist genau dann invertierbar, wenn $\det A \in R^\times$, d.h. Einheit in R ist.*

Beweis: Ist A invertierbar, so ist $AA^{-1} = I$, also $\det A \det A^{-1} = \det I = 1$, sodaß $\det A$ und $\det A^{-1}$ Einheiten sind.

Daß dies auch hinreichend ist, entnimmt man der in Satz D.21 gegebenen Darstellung der Inversen mittels der Algebraischen Komplemente, in der lediglich durch die Determinante der Matrix dividiert wird. □

In den uns interessierenden Fällen werden wir allerdings die Invertierbarkeit durch Angeben der Inversen nachweisen.

Unser angestrebtes Hauptergebnis ist der folgende

Satz Z.60 (Invarianten-Teiler-Satz) *Über einem Hauptideal-Ring R gibt es zu jeder Matrix $F \in R^{n\times m}$ invertierbare Matrizen $B \in R^{n\times n}$ und $C \in R^{m\times m}$, ferner eine eindeutig bestimmte Zahl $k \leq \min\{n, m\}$ und bis auf Einheiten festgelegte Ringelemente $(\delta_i \mid i = 1, 2, ..., k)$, sodaß*

(i) alle $\delta_i \neq 0$ sind und die sukzessive Teilbarkeitsbeziehung

$$\delta_1 \mid \delta_2 \mid ... \mid \delta_k$$

gilt,

(ii) $F' := B^{-1}FC$ die spezielle Gestalt

$$F' = \begin{pmatrix} D_k & 0 \\ 0 & 0 \end{pmatrix}$$

mit $D_k = \mathrm{diag}(\delta_1, ..., \delta_k)$ hat.

Bezeichnung Z.61 *Die δ_i nennt man die Invarianten-Teiler der Matrix A.*

Über einem Körper sind ja alle Elemente $\neq 0$ schon Einheiten und mit solchen kann man trivialerweise auch die δ_i noch multiplizieren, sodaß über einem Körper dann die δ_i sämtlich $= 1$ gewählt werden können, womit wir dann Satz H.41 wiedergewonnen haben.

Zur Konstruktion dieser Normalform verschaffen wir uns zunächst eine Reihe von speziellen invertierbaren Matrizen, mit denen wir die gewünschte Darstellung konstruktiv gewinnen werden.

Lemma Z.62 *Beispiele für über einem Hauptideal-Ring R invertierbaren Matrizen sind:*

(i) *Für Einheiten* $\lambda_1, ..., \lambda_n \in R^\times$

$$\Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_n), \quad \Lambda^{-1} = \mathrm{diag}(\lambda_1^{-1}, ..., \lambda_n^{-1}).$$

(ii)
$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

(iii)
$$E_\alpha = \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix}, \quad E_\alpha^{-1} = E_{-\alpha}.$$

Dito für die transponierte Matrix.

(iv) *Zu* $\alpha, \beta \in R$, *beide* $\neq 0$ *mit* $\delta := \mathrm{ggT}(\alpha, \beta) =: \lambda\alpha + \mu\beta$ *sei* $\alpha = \alpha'\delta$, $\beta = \beta'\delta$. *Dann sind*
$$Q = \begin{pmatrix} \lambda & \mu \\ -\beta' & \alpha' \end{pmatrix}, \quad Q^{-1} = \begin{pmatrix} \alpha' & -\mu \\ \beta' & \lambda \end{pmatrix}$$
bzw. die jeweiligen transponierten invertierbar und je invers zueinander.

Beweis: Nur (iv) ist nicht (völlig) trivial. Es ist $\delta = \lambda\alpha + \mu\beta = \lambda\alpha'\delta + \mu\beta'\delta$ also, da wir kürzen dürfen, auch $1 = \lambda\alpha' + \mu\beta'$. Dann folgt die Behauptung durch Ausmultiplizieren. □

Wenn man diese $2 \times 2-$Matrizen symmetrisch zur Diagonalen in eine (große) Einheitsmatrix als Haupt-Minoren einbaut, erhält man bei (ii) bzw. (iii) Matrizen, die durch Multiplikation von links

— 2 Zeilen vertauschen (Transposition)

— das α-fache einer Zeile zu einer anderen addieren (Elimination),

bzw. das Analoge mit Spalten bei Multiplikation von rechts.

Wenden wir die Matrizen aus (iv) auf eine Matrix F an, die zwei Zeilen der Form
$$F = \begin{pmatrix} \alpha & \dots \\ \beta & \dots \end{pmatrix}$$
hat, so ist

$$F' = QF = \begin{pmatrix} \lambda & \mu \\ -\beta' & \alpha' \end{pmatrix} \begin{pmatrix} \alpha & \dots \\ \beta & \dots \end{pmatrix} = \begin{pmatrix} \lambda\alpha + \mu\beta & \dots \\ \cdot & \dots \end{pmatrix} = \begin{pmatrix} \mathrm{ggT}(\alpha, \beta) & \dots \\ \cdot & \dots \end{pmatrix}$$

Diese Operation erzeugt also den ggT der Elemente α und β in der Position von α in der ersten Zeile.

Analog kann man mit Q^T von rechts auf den Spalten operieren.

Wir haben somit als Werkzeug zum Beweisen von Satz Z.60 das Folgende zur Verfügung:

Lemma Z.63 *Über einem Hauptideal-Ring sind an einer Matrix durch Multiplikation mit invertierbaren Matrizen folgende Operationen ausführbar:*

(i) *Vertausche zwei Zeilen (Transponieren).*

(ii) *Addiere zu einer Zeile ein Vielfaches einer anderen (Elimination).*

(iii) *Ersetze eine Zeile durch eine solche Linearkombination mit einer anderen, daß an einer gegebenen Position der ggT der beiden entsprechenden Elemente entsteht.*

Analoges geht für die Spalten.

Wir beschreiben nun zunächst einen Algorithmus, der eine gegebene Matrix $F = (\varphi_{ij}) \in R^{n\times m}$ in die in Satz Z.60 genannte Diagonalform überführt, *sofern der Ring euklidisch ist.* Wir benutzen dabei zentral, daß die euklidische Normfunktion ν ihre Werte in den natürlichen Zahlen hat.

Algorithmus Z.64 *Gegeben sei eine Matrix $F = (\varphi_{ij}) \in R^{n\times m}$ über einem euklidischen Ring R.*

Schritt 1: Suche in der ganzen Matrix ein Element $\varphi_{ij} \neq 0$, sodaß $\nu(\varphi_{ij})$ minimal, und schaffe dieses durch Tauschen von Zeilen und Spalten an die Position $(1,1)$. Gibt es kein solches, so sind wir fertig.

Schritt 2: Betrachte die weiteren Elemente der ersten Zeile und ersten Spalte, also die Elemente $\varphi_{1j}, \varphi_{i1}$ $(i,j > 1)$.

(a) Sind alle $= 0$, so mache weiter bei Schritt 3.

(b) Ist etwa $\varphi_{1j} \neq 0$ (sinngemäß geht alles genauso für ein Element der ersten Spalte), so dividiere

$$\varphi_{1j} = q\varphi_{11} + \varphi'_{1j} \text{ wobei } \nu(\varphi'_{1j}) < \nu(\varphi_{11}).$$

Ein Eliminations-Schritt erlaubt es φ_{1j} durch φ'_{1j} zu ersetzen.

i. Ist jetzt $\varphi_{1j} \neq 0$ (gemeint ist das aktuell an dieser Position stehende Element), so ist jedenfalls $\nu(\varphi_{1j}) < \nu(\varphi_{11})$. Wir beginnen wieder neu mit Schritt 1, wobei danach die Norm des $(1,1)$-Elementes kleiner geworden ist.

ii. Ist jetzt $\varphi_{1j} = 0$, so beginne wieder Schritt 2.

Schritt 3: Jetzt ist $\varphi_{11} \neq 0$, aber alle anderen Elemente der ersten Zeile und der ersten Spalte sind alle $= 0$.

(a) Es gibt ein Element $\varphi_{ij}, (i,j \geq 2)$, das nicht von φ_{11} geteilt wird: Dann Addiere Zeile i zu Zeile 1. (Dabei ändert sich φ_{11} nicht!) Beginne wieder Schritt 2.

(b) Andernfalls haben wir die Situation: In der ersten Zeile und Spalte sind alle Elemente $= 0$ ausgenommen das Diagonalelement φ_{11}, was $\neq 0$ ist. Dieses teilt alle Elemente φ_{ij} mit $i,j \geq 2$. Nun streiche die erste Zeile und erste Spalte — wobei das Diagonalelement als ein δ für die gesuchte Diagonalmatrix aufzubewahren ist — und beginne den Algorithmus neu für die verkleinerte Matrix mit Schritt 1.

Dieser Algorithmus liefert offenbar die gewünschte Diagonalform, wenn sicher ist, daß er jemals fertig wird. Dabei ist allein der Rücksprung nach Schritt 1 problematisch, bei dem ja scheinbar fast alles Erreichte wieder durcheinander gebracht wird. Dies ist aber nicht so: denn jedesmal, wenn dieser Schritt 1 durchgeführt wird, bekommt das Element in der Position $(1,1)$ eine definitiv kleinere Norm, was nur endlich oft möglich ist, da es sich stets um natürliche Zahlen handelt. Also kommt der Algorithmus zum Ende.

Wir bemerken noch für das praktische Durchführen, daß man jederzeit die Elemente einer Zeile oder Spalte mit einer festen Einheit aus $R^\times$ multiplizieren darf.

Ist der Ring nur ein *Hauptideal-Ring,* so können wir die Primfaktorzerlegung heranziehen. Statt der euklidischen Norm-Funktion ν verwenden wir die Länge $\lambda(\alpha) :=$ Anzahl der Faktoren in der Primfaktorzerlegung zu α und ersetzen die Division mit Rest durch die Bildung des ggT.

Haben α und $\mathrm{ggT}(\alpha,\beta)$ die selbe Anzahl von Primfaktoren, so ist α ein Teiler von β, also $\beta - \gamma\alpha = 0$ für ein geeignetes γ. Andernfalls hat $\delta := \mathrm{ggT}(\alpha,\beta)$ weniger Primfaktoren als α, ferner können wir nach Lemma Z.63.3 dann α durch diesen

ggT ersetzen. Folglich lassen sich alle Schritte des Algorithmus auch in diesem Fall nachbauen.

Damit ist der Existenz-Teil von Satz Z.60 nachgewiesen.

Es bleibt noch die Eindeutigkeit der Invariantenteiler zu zeigen.

Bezeichnung Z.65 *Zu einer Matrix A über R sei*
$d_j(A)$ *der ggT aller j-reihigen Unterdeterminanten von A.*
$d_j(A)$ *heißt der "j-te Determinanten-Teiler von A."*

Dafür gilt

Lemma Z.66 *Mit Matrizen A, B, C über R ist*
$d_j(A)$ *Teiler von* $d_j(BA)$,
$d_j(A)$ *Teiler von* $d_j(AC)$.

Beweis: Die Spalten von AC sind Linearkombinationen der Spalten von A. Folglich ist jede j-reihige Unterdeterminante von AC eine Linearkombination von (vielen) j-reihigen Unterdeterminanten von A. Diese werden aber alle von $d_j(A)$ geteilt, somit teilt dies auch jede Linearkombination, also auch die j-reihigen Unterdeterminanten von AC.

Analog schließt man über die Zeilen für die andere Aussage. □

Damit bekommen wir

Satz Z.67 *Es seien R ein Hauptideal-Ring, $A \in R^{n\times m}$, $B \in R^{n\times n}$, $C \in R^{m\times m}$, dabei B, C invertierbar, $A' := B^{-1}AC$. Dann haben A und A' die selben Determinanten-Teiler.*

Beweis: Mit Lemma Z.66 und der Invertierbarkeit der Matrizen B, C bekommen wir die Teilbarkeitsbeziehungen

$$d_j(A) \mid d_j(B^{-1}AC) \quad \text{und} \quad d_j(B^{-1}AC) \mid d_j(BB^{-1}ACC^{-1}) = d_j(A). \qquad \square$$

Für die Matrix $F' = \begin{pmatrix} D_k & 0 \\ 0 & 0 \end{pmatrix}$ mit $D_k = \operatorname{diag}(\delta_1, ..., \delta_k)$, wobei die $\delta_i \neq 0$ und $\delta_1 \mid \delta_2 \mid ... \mid \delta_k$ haben wir offenbar

$$d_j(F') = \delta_1 \cdot ... \cdot \delta_j \quad \text{für } j \leq k,$$

während für $j > k$ alle j-reihigen Unterdeterminanten von F' verschwinden, was dann auch $d_j(F') = 0$ bedeutet. Für die Determinanten-Teiler ist es nun nach Satz Z.67 unerheblich ob wir sie für die Matrix F' oder für die "unbehandelte" Matrix F betrachten. Also folgt

Satz Z.68 *Hat F gemäß Satz Z.60 eine Darstellung als F' mit Invarianten-Teilern $\delta_1, ..., \delta_k$ (alle $\neq 0$), so gilt für die Determinanten-Teiler von F :*

$$d_j(F) = \begin{cases} \delta_1 \cdot ... \cdot \delta_j & \text{für } j \leq k, \\ 0 & \text{für } j > k. \end{cases}$$

Damit sind k und die δ_j (letztere nur bis auf Multiplikation mit Einheiten) eindeutig bestimmt.

Dies war nun gerade der noch fehlende Eindeutigkeits-Teil zum Invarianten-Teiler-Satz Z.60, der damit vollständig bewiesen ist.

Behandeln wir über dem Polynomring $K[T]$ ein paar für das Weitere wichtige Beispiele:

Satz Z.69 *Ist F die Begleitmatrix zu einem normierten Polynom p vom Grad n, so hat deren charakteristische Matrix $A(T) := F - T \cdot I$ die Invarianten-Teiler*

$$(\delta_1, ..., \delta_n) = (1, 1, ..., 1, p(T)),$$

wobei also genau $n-1$ mal eine 1 auftritt.

Beweis: Die charakteristische Matrix $A(T)$ hat die Gestalt

$$A(T) = \begin{pmatrix} -T & 0 & & \dots & 0 & \alpha_0 \\ 1 & -T & 0 & \dots & 0 & \alpha_1 \\ & 1 & -T & \dots & 0 & \alpha_3 \\ & & & \ddots & \vdots & \vdots \\ & & & & -T & \alpha_{n-2} \\ & & & & 1 & \alpha_{n-1} - T \end{pmatrix}_{n,n}.$$

Streicht man die erste Zeile und die letzte Spalte, so entsteht eine obere Dreiecks-Matrix mit Diagonal-Elementen 1, deren Determinante also 1 ist. Damit folgt sofort:

Für $j < n$ sind alle Determinanten-Teiler $d_j(A) = 1$. Ferner ist nach Satz Z.32 $d_n(A) = \pm p(T)$. Über Satz Z.68 ergibt sich daraus unmittelbar die Behauptung für die Invarianten-Teiler. □

Dies kann man verallgemeinern.

Satz Z.70 *Es seien $p_1, ..., p_m \in K[T]$ normierte Polynome, F_i deren Begleitmatrizen, $A_i(T) := F_i - T \cdot I$ die entsprechenden charakteristischen Matrizen. Ferner bilden wir die Block-Diagonal-Matrix*

$$A := \text{diag}(A_1, ..., A_m).$$

Dann gibt es invertierbare Matrizen B, C über $K[T]$, sodaß

$$A' := B^{-1}AC = \text{diag}(1, ..., 1, p_1, ..., p_m).$$

Erfüllen insbesondere die p_i die Teilbarkeitsbedingung $p_1 \mid p_2 \mid ... \mid p_m$, so sind sie genau die nichttrivialen Invarianten-Teiler der charakteristischen Matrix A.

Beweis: Jede einzelne Matrix A_i ist nach dem vorigen Satz zu einer speziellen Diagonalmatrix assoziiert, d.h. es gibt über $K[T]$ passende invertierbare Matrizen B_i, C_i, sodaß

$$A_i' := B_i^{-1} A_i C_i = \text{diag}(1, ..., 1, p_i).$$

Führt man diese Transformationen simultan für alle Blöcke der Block-Diagonal-Matrix $A = \text{diag}(A_1, ..., A_m)$ aus, so erhält man als die dazu asoziierte Matrix $\text{diag}(A_1', ..., A_m')$. Dies ist schon eine echte Diagonal-Matrix, sogar mit den selben Diagonal-Elementen wie A', allerdings noch in anderer Reihenfolge notiert. Dies bringt man mit Permutations-Matrizen in Ordnung, womit dann die gewünschte Form erhalten ist.

Mit der Eindeutigkeits-Aussage des Invarianten-Teiler-Satzes folgt dann die Behauptung. □

Moduln über Hauptideal-Ringen

Im weiteren sei R wieder ein Hauptideal-Ring, also insbesondere kommutativ mit Einselement und ohne Nullteiler.

Definition Z.71 (R-Modul) *Ersetzen wir in der Definition des Vektorraumes den Körper K durch unseren Ring R, so erhalten wir den Begriff des "R-Moduls".*

Auf natürliche Weise übertragen sich darauf die Begriffe Untermodul (statt Unterraum), Erzeugendensystem, linear abhängig, linear unabhängig, Basis. Ebenso erhalten wir aus den Vektorraum-Homomorphismen den entsprechenden Begriff der R-Modul-Homomorphismen, bzw. entsprechender Isomorphismen. Kerne und Bilder von Homomorphismen sind offenbar wieder Untermoduln.

Beispiele für Moduln sind etwa

- jeder K-Vektorraum als Modul über dem Ring $R := K$.
- jede abelsche Gruppe als Modul über dem Ring $\mathbb{Z}$ der ganzen Zahlen.
- ein K-Vektorraum als $K[T]$-Modul:

  Fixiere einen Vektorraum-Homomorphismus $f \in \mathrm{Hom}_K(V,V)$. Erkläre eine Operation
  $$\cdot : K[T] \times V \to V$$
  durch
  $$(p(T), v) \mapsto p(f)(v).$$
  Damit wird V ein Modul über dem Polynomring $K[T]$.

Definition Z.72 *Ein Modul heißt*
- endlich erzeugt, wenn er ein endliches Erzeugendensystem besitzt,
- zyklisch, wenn er von einem Element erzeugt wird,
- frei, wenn er eine Basis besitzt.

Hier werden nun schon Unterschiede zum Vektorraum deutlich.

Nehmen wir wieder einen endlich-dimensionalen K-Vektorraum V aufgefaßt als $K[T]$-Modul, wobei die "Skalar-Multiplikation" über den Vektorraum-Homomorphismus f vermittelt werde. Die "konstanten" Polynome in $K[T]$, d.h. die vom Grad $= 0$ operieren dann wie die Körperelemente auf V. Damit folgt sofort:

Erzeugen $(v_1, ..., v_n)$ den Raum V als K-Vektorraum, so auch als $K[T]$-Modul.
Andrerseits gilt:

Der $K[T]$-Modul V besitzt keine linear unabhängige Familie, insbesondere keine Basis; denn ist $\psi(T) \in K[T]$ das Minimalpolynom zu f, so ist $\psi(f) = 0$, also auch $\psi(f)(v) = 0$ für *jedes* Element $v \in V$, sodaß wegen $\psi \neq 0$ also stets (v) eine einelementige Familie ist, die linear abhängig ist, und damit natürlich in keiner Basis vorkommen kann.

Ferner erkennt man die aus einem Element v entstehenden f-zyklischen Unterräume $U_v := \{p(f)(v) \mid p \in K[T]\}$ als genau die zyklischen Untermoduln des $K[T]$-Moduls V.

Daß nicht alles anders wird, zeigen die nächsten Aussagen:

Lemma Z.73 *Für jeden Ring R und jedes m ist*
$$R^m := \left\{ \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} \mid r_i \in R \right\}$$
mit der naheliegenden Struktur ein freier R-Modul. Die kanonischen Einheitsvektoren - sie seien wieder mit e_i bezeichnet - bilden eine Basis.

Satz Z.74 *(i) Sind V, W R-Moduln, dabei V frei mit Basis $(v_1, ..., v_n)$ und ist $(w_1, .., w_n)$ in W, so existiert genau ein R-Modul-Homomorphismus*

$$\varphi : V \to W \text{ mit } v_i \mapsto w_i \ (i = 1, ..., n).$$

(ii) φ ist genau dann ein Isomorphismus, wenn auch $(w_1, .., w_n)$ eine Basis ist.

Die **Beweise** verlaufen wie bei Vektorräumen und seien deshalb übergangen.

Wir hatten gesehen, daß ein Modul keine Basis zu besitzen braucht. Ist der Ring nicht so schön, wie wir dies hier voraussetzen, so kann es sogar geschehen, daß ein Untermodul eines endlich erzeugten Moduls nicht mehr endlich erzeugt ist. Über Hauptideal-Ringen (und auch noch über etwas allgemeineren Ringen) kann dies allerdings nicht passieren. Dies wollen wir nun herleiten. Zunächst

Lemma Z.75 *Ist $\varphi : V \to W$ ein Modulhomomorphismus und sind* $\ker \varphi$ *und* $\text{im } \varphi$ *endlich erzeugt, so ist auch V endlich erzeugt.*

Beweis: Sei $\ker \varphi = \text{span}(v_1, ..., v_k)$, $\text{im } \varphi = \text{span}(w_1, ..., w_m)$. Dann existieren Elemente $(v'_1, ..., v'_m) \in V$, sodaß $w_i = \varphi(v'_i)$ $(i = 1, ..., m)$.

Wir zeigen: $(v_1, ..., v_k, v'_1, ..., v'_m)$ erzeugen V:

Dazu wähle beliebig $v \in V$. Dann ist $\varphi(v) \in \text{im } \varphi = \text{span}(w_1, ..., w_m)$, also $\varphi(v) = \sum_{i=1}^{m} \alpha'_i w_i$, d.h. mit $v' := \sum_{i=1}^{m} \alpha'_i v'_i$ ist $\varphi(v) = \varphi(v')$, oder $\varphi(v - v') = 0$, somit

$$v'' := v - v' \in \ker \varphi = \text{span}(v_1, ..., v_k).$$

Also gibt es eine Darstellung

$$v'' = \sum_{i=1}^{k} \alpha''_i v_i$$

und damit

$$v = v'' + v' = \sum_{i=1}^{k} \alpha''_i v_i + \sum_{i=1}^{m} \alpha'_i v'_i,$$

was zu zeigen war. □

Satz Z.76 *Jeder Untermodul U eines endlich erzeugten Moduls V über einem (Hauptideal-)Ring R ist selbst endlich erzeugt.*

Beweis: *1. Fall : $V = R^m$:*

Wir führen Induktion über m.

$m = 1$: Dann ist $V = R$, $U \subset R$ ist als Untermodul ein Ideal, was nach Voraussetzung über den Ring sogar Hauptideal ist, also endlich erzeugt. (Hier würde eine schwächere Ring-Eigenschaft reichen.)

$m > 1$: Betrachte die Abbildung

$$\pi : R^m \to R^{m-1} : \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{m-1} \\ \alpha_m \end{pmatrix} \mapsto \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{m-1} \end{pmatrix}$$

mit

$$\ker \pi = R \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = R \cdot e_m \hat{=} R, \quad \text{im } \pi = R^{m-1}.$$

Es ist $U \subset R^m$. Betrachte

$$\varphi : U \to R^{m-1},\ \varphi := \pi_{|U}.$$

Dann ist

$$\ker \varphi = \ker \pi \cap U = U \cap (R \cdot e_m),$$

also Untermodul von R und damit endlich erzeugt. Ferner ist im φ Untermodul von R^{m-1}, also (nach Induktionsannahme) endlich erzeugt. Folglich ist nach Lemma Z.75 auch U selbst endlich erzeugt.

2. Fall : $V = \text{span}(v_1, ..., v_m)$ ist allgemeiner endlich erzeugter Modul mit Untermodul U: Dann existiert nach Satz Z.74 ein Modulhomomorphismus

$$\psi : R^m \to V, \quad e_i \to v_i\ (i = 1, ..., m),$$

der also surjektiv ist. Betrachte zu $U \subset V$:

$$\psi^{-1}(U) := \{w \in R^m \mid \psi(w) \in U\} \subset R^m.$$

Dies ist trivialerweise ein Untermodul von R^m, also nach dem schon behandelten Fall endlich erzeugt: $\psi^{-1}(U) = \text{span}(w_1, ..., w_n)$. Dann gilt aber für U selbst: $U = \text{span}(\psi(w_1), ..., \psi(w_n))$, d.h. U ist endlich erzeugt. □

Eine analoge Aussage gilt für freie Moduln. Ehe wir diese genauer formulieren, ein paar technische Vorüberlegungen:

Bezeichnung Z.77 *Sind $(v_1, ..., v_n)$, $(w_1, ..., w_m)$ Familien in einem R-Modul V und gilt mit Koeffizienten α_{ij} einer Matrix $A = (\alpha_{ij}) \in R^{n \times m}$*

$$w_j = \sum_{i=1}^{n} \alpha_{ij} v_i,$$

so notieren wir

$$(w_1, ..., w_m) = (v_1, ..., v_n) \cdot A.$$

Lemma Z.78 *(i) Die eben eingeführte Notation ist kompatibel mit der Matrixmultiplikation.*

(ii) Ist $(v_1, ..., v_n)$ erzeugend und A invertierbar, so ist auch $(v_1, ..., v_n) \cdot A$ erzeugend.

(iii) Ist $(v_1, ..., v_n)$ eine Basis, so ist $(v_1, ..., v_n) \cdot A$ genau dann Basis, wenn A invertierbar ist.

Beweis als Übung.

Satz Z.79 *Es sei V frei mit Basis $(v_1, .., v_n)$, $U \subset V$ ein Untermodul. Dann existieren eine Basis $(v'_1, .., v'_n)$ von V und Ringelemente $\delta_1, ..., \delta_k$ $(k \leq n)$ wobei alle $\delta_i \neq 0$ und jeweils δ_i Teiler von δ_j, sofern $i < j$, sodaß $(\delta_1 v'_1, ..., \delta_k v'_k)$ Basis von U ist.*

Insbesondere ist U selbst frei.

Beweis: V ist endlich erzeugt. Wie oben gezeigt hat dann auch der Untermodul U ein endliches Erzeugenden-System: $U = \text{span}(u_1, ..., u_m)$. Dies sind alles Elemente von V selbst. Mit einer geeigneten Koeffizientenmatrix $A = (\alpha_{ij}) \in R^{n \times m}$ ist also

$$(u_1, ..., u_m) = (v_1, ..., v_n) \cdot A.$$

Nach dem Diagonalisierungs - Satz Z.60 gibt es dann invertierbare Matrizen $B \in R^{n\times n}$ und $C \in R^{m\times m}$, sodaß $A' := B^{-1}AC$ die Gestalt

$$A' = \text{diag}(\delta_1, ..., \delta_k, 0, ..., 0)$$

hat, wobei $k \leq \min\{n, m\}, \delta_i \neq 0$ und die Teilbarkeits-Bedingung $\delta_1 \mid \delta_2 \mid ... \mid \delta_k$ gilt.

Da B, C invertierbar sind, folgt aus dem Lemma Z.78:
$(v'_1, ..., v'_n) := (v_1, ..., v_n)B$ ist Basis von V und
$(u'_1, ..., u'_m) := (u_1, ..., u_m)C$ ist erzeugend für U.

Es ist dann

$$(u'_1, ..., u'_m) = (u_1, ..., u_m)C = (v_1, ..., v_n)AC = (v'_1, ..., v'_n)B^{-1}AC = (v'_1, ..., v'_n)A'$$

$$= (v'_1, ..., v'_n) \cdot \text{diag}(\delta_1, ..., \delta_k, 0, ..., 0) = (\delta_1 v'_1, ..., \delta_k v'_k, 0..., 0).$$

Die $(v'_1, ..., v'_n)$ bilden eine Basis von V. Sie sind also linear unabhängig. Da die $\delta_i \neq 0, (1 \leq i \leq k)$, sind dann auch die $(\delta_1 v'_1, ..., \delta_k v'_k)$ linear unabhängig, also auch die $(u'_1, ..., u'_k)$. Ferner ist $u'_i = 0$ für $k+1 \leq i \leq m$, sodaß schon die $(u'_1, ..., u'_k)$ den Untermodul U erzeugen, damit also eine Basis der gewünschten Art von U bilden. □

Bemerkung Z.80 *Wären wir in einem Vektorraum, d.h. wäre der Ring R sogar ein Körper, so könnte man die Faktoren δ_i herausdividieren und wir hätten eben eine Basis des ganzen Raumes V erhalten, die als Teilfamilie eine Basis von U enthält.*

Dies braucht im Modul-Fall nicht so zu sein.

Als ein **Beispiel** betrachten wir über $R := \mathbb{Z}$ den freien Modul

$$\mathbb{Z}^2 = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \mid z_i \in \mathbb{Z} \right\} \quad \text{mit der Basis } v_1 := e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, v_2 := e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Dazu betrachte den Untermodul

$$U = \left\{ \begin{pmatrix} 2z_1 \\ 3z_2 \end{pmatrix} \mid z_i \in \mathbb{Z} \right\} \quad \text{erzeugt von } u_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}.$$

Offenbar sind diese beiden Erzeugenden sogar unabhängig, also eine Basis von U, aber sie erzeugen *nicht* den ganzen Modul V (e_1, e_2 gehören nicht zu U!)

Es kann also sein, daß ein freier Modul und ein echter Untermodul von ihm gleich lange Basen besitzen! In Vektorräumen ist das unmöglich.

Bestimmt man die Basen nach Satz Z.79 so erhält man:

$$u'_1 = \begin{pmatrix} -2 \\ 3 \end{pmatrix}, u'_2 = \begin{pmatrix} 6 \\ -6 \end{pmatrix}, v'_1 = \begin{pmatrix} -2 \\ 3 \end{pmatrix}, v'_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ mit } \delta_1 = 1, \delta_2 = 6.$$

Sind wir in einem K-Vektorraum V und ist $\alpha v = 0$, so ist entweder $\alpha = 0 \in K$ oder $v = 0 \in V$. Bei Moduln über Ringen (selbst wenn die schön sind!) braucht dies nicht der Fall zu sein

Beispiel: Es sei $V := \mathbb{Z}_{/10\mathbb{Z}}$, d.h. die abelsche Gruppe mit den Elementen $\{0, 1, ..., 9\}$ und der Addition durch Rest modulo 10 nach Addition als ganze Zahlen. Hier ist offenbar 0 das neutrale Element der Addition, ferner ist etwa $2 \in V, \neq 0$, ebenso sind $2 \cdot 2 := 2 + 2$, $3 \cdot 2 := 2 + 2 + 2$, $4 \cdot 2 := 2 + 2 + 2 + 2$ alle $\neq 0$, aber $5 \cdot 2 := 2 + 2 + 2 + 2 + 2 = 10 \equiv 0$. In diesem $\mathbb{Z}$-Modul $V := \mathbb{Z}_{/10\mathbb{Z}}$ sind also die Elemente $1 \cdot 2$, $2 \cdot 2$, $3 \cdot 2$, $4 \cdot 2$ sämtlich $\neq 0$, während $5 \cdot 2 = 0$ ist.

Man sagt: Das Element 2 hat die Ordnung 5. Offenbar ist für eine ganze Zahl $z \in \mathbb{Z}$ hier $z \cdot 2 = 0$ in $\mathbb{Z}_{/10\mathbb{Z}}$, genau wenn 5 ein Teiler von z, d.h. $z \in (5)$, dem von

5 in $\mathbb{Z}$ erzeugten Ideal ist. Dies nennt man das Ordnungs-Ideal von $2 \in \mathbb{Z}_{/10\mathbb{Z}}$ im Ring $\mathbb{Z}$.

Beispiel: V sei wieder ein K-Vektorraum, $f \in \mathrm{Hom}(V,V)$ und wir betrachten wieder V als $K[T]$-Modul mittels

$$p(T) \cdot v := p(f)(v) \qquad p \in K[T],\ v \in V.$$

Wir wissen: Es gibt dann zu jedem $v \in V$ nichttriviale Polynome $q(T) \in K[T]$, sodaß $q(f)(v) = 0$ oder mit der Modul-Operation $q(T) \cdot v = 0$. Wie man leicht nachrechnet bilden diese Polynome ein Ideal in $K[T]$, das Ordnungs-Ideal zu dem Element v.

Satz und Definition Z.81 *Es sei V ein Modul über einem (Hauptideal-)Ring R. Dann ist für jedes $v \in V$ die Menge $\mathcal{J}_v := \{s \in R \mid s \cdot v = 0\}$ ein Ideal in R, das sogenannte "Ordnungs-Ideal" oder den "Annihilator" von v. Da R ein Hauptideal-Ring ist, ist $\mathcal{J}_v$ ein Hauptideal, also von einem Element ψ_v erzeugt. Jedes solche ψ_v heißt eine "Ordnung" von v. Genau das Nullelement hat die Ordnung 1, d.h. den ganzen Ring als Ordnungsideal.*

Beweis: Es ist nur die Idealeigenschaft zu zeigen: Seien $s_1, s_2 \in \mathcal{J}_v, r_1, r_2 \in R$ beliebig; dann ist

$$(r_1 s_1 + r_2 s_2)v = r_1(s_1 v) + r_2(s_2 v) = r_1 0 + r_2 0 = 0 + 0 = 0,$$

sodaß also auch $r_1 s_1 + r_2 s_2 \in \mathcal{J}_v$, was die Ideal-Eigenschaft beweist. □

Für Moduln erklären wir auch wieder den Begriff der Direkten Summe:

Definition Z.82 *Es sei V ein R-Modul mit Untermoduln $V_1, ..., V_n$. Dann ist V die "direkte Summe" der V_i,*

$$V = V_1 \oplus ... \oplus V_n,$$

wenn $V = V_1 + ... + V_n$, d.h. jedes $v \in V$ eine Darstellung $v = v_1 + ... + v_n$ mit $v_i \in V_i$ hat, und dies überdies direkt ist, d.h. wenn $0 = v_1 + ... + v_n$ mit $v_i \in V_i$, so notwendig $v_1 = ... = v_n = 0$.

Damit haben wir nun alles zusammen für den zentralen

Satz Z.83 (Struktursatz für endl. erz. Moduln über Hauptideal-Ringen) *Jeder endlich erzeugte Modul V über einem Hauptideal-Ring R ist eine direkte Summe von zyklischen Moduln und einem freien Modul.*

Genauer existiert eine Darstellung

$$V = V_1 \oplus ... \oplus V_k \oplus V',$$

wobei V' frei (evtl. trivial), die V_i zyklisch, d.h. von der Form $V_i = \mathrm{span}(v_i)$ und die v_i von Ordnungen $\delta_i \neq 0$, die zudem die Teilbarkeitsbedingung $\delta_1 \mid \delta_2 \mid ... \mid \delta_k$ erfüllen.

Darüber hinaus gilt folgende Eindeutigkeitsaussage:

Die Dimension des freien Anteiles ist eindeutig, die Ordnungen δ_i sind bis auf Einheiten eindeutig bestimmt.

Beweis: Wir beschränken uns darauf, die Existenzaussage zu beweisen.

Es sei $V = \mathrm{span}(w_1, ..., w_n)$. Betrachte den surjektiven Modul-Homomorphismus

$$\varphi : R^n \to V, \quad e_i \to w_i,$$

und setze $U := \ker\varphi \subset R^n$. Dies ist ein Untermodul des freien Moduls R^n, somit gibt es nach Satz Z.79 eine Basis $(b_1, ..., b_n)$ von R^n und dazu Ringelemente $\delta_1, .., \delta_k \neq 0$ mit $\delta_i \mid \delta_j$ $(1 \leq i < j \leq k)$, sodaß $(\delta_1 b_1, ..., \delta_k b_k)$ Basis von U ist.

Setze

$$v_i := \varphi(b_i) \ (i = 1, ..., n)$$

und

$$V_i := \text{span}(v_i) \ (i = 1, ..., k), \quad V' := \text{span}(v_{k+1}, ..., v_n).$$

Da $(b_1, ..., b_n)$ eine Basis von R^n bilden und $\varphi : R^n \to V$ surjektiv ist, ist damit $(v_1, ..., v_n)$ erzeugend für V, d.h.

$$V = V_1 + ... + V_k + V',$$

wobei $V_1, ..., V_k$ zyklisch sind.

Es bleibt zu zeigen, daß die v_i $(i \leq k)$ die richtige Ordnung haben, daß V' frei ist und die Summe direkt. All dies ergibt sich aus den beiden folgenden Aussagen, die wir zunächst beweisen werden. Es gelten

1. $\delta_i v_i = 0 \quad (i = 1, ..., k)$
2. $\sum_{i=1}^n \alpha_i v_i = 0 \ \Leftrightarrow \ (\alpha_i = \lambda_i \delta_i \text{ mit } \lambda_i \in R, \ (i = 1, ..., k) \text{ und } \alpha_i = 0 \ (i > k))$.

Beweis:

1. Für $i \leq k$ ist $\delta_i b_i \in \ker\varphi$, also

$$0 = \varphi(\delta_i b_i) = \delta_i \varphi(b_i) = \delta_i v_i.$$

2. Haben die Koeffizienten α_i die angegebene Gestalt, so folgt mit 1. sofort, daß $\sum_{i=1}^n \alpha_i v_i = 0$.

   Sei umgekehrt

$$0 = \sum_{i=1}^n \alpha_i v_i = \sum_{i=1}^n \alpha_i \varphi(b_i) = \varphi(\sum_{i=1}^n \alpha_i b_i).$$

   Dann ist

$$u := \sum_{i=1}^n \alpha_i b_i \in \ker\varphi = U,$$

   hat also die Basisdarstellung

$$u := \sum_{i=1}^k \lambda_i \delta_i b_i,$$

   sodaß wegen der Eindeutigkeit folgt, daß

$$\alpha_i = \lambda_i \delta_i \ (i = 1, ..., k)$$

   und

$$\alpha_i = 0 \ (i > k),$$

   womit die Zwischen-Behauptung gezeigt ist.

Wenden wir dies an: Ist mit $\alpha \in R$ für ein $i \leq k$: $\alpha v_i = 0$, so ist nach 2. notwendig $\alpha \in (\delta_i)$, dem von δ_i erzeugten Hauptideal, sodaß wegen $\delta_i v_i = 0$ dies genau das Ordnungsideal ist. Alle v_i haben also jeweils die Ordnung δ_i $(i \leq k)$.

Ist mit $\alpha_i \in R$: $\sum_{i=k+1}^{n} \alpha_i v_i = 0$, so ist nach 2. notwendig $\alpha_{k+1} = ... = \alpha_n = 0$, sodaß $(v_{k+1}, ..., v_n)$ linear unabhängig sind, also eine Basis von V'. Damit ist V' frei.

Sind schließlich $w_i \in V_i$ $(i = 1, ..., k)$, $w' \in V'$ und $0 = \sum_{i=1}^{k} w_i + w'$, so haben wir ja Darstellungen

$$w_i = \alpha_i v_i \quad (i = 1, ...k),$$

$$w' = \sum_{i=k+1}^{n} \alpha_i v_i,$$

sodaß

$$0 = \sum_{i=1}^{k} \alpha_i v_i + \sum_{i=k+1}^{n} \alpha_i v_i,$$

also nach 2.

$$\alpha_i = \lambda_i \delta_i \text{ mit } \lambda_i \in R, (i = 1, ..., k) \text{ und } \alpha_i = 0 \ (i > k).$$

Damit ist $w' = 0$ und, da $\delta_i v_i = 0$, sind auch alle $w_i = 0$ $(i = 1, ..., k)$, sodaß die Summe direkt ist. □

Ist eines der δ_i eine Einheit in R, so ist notwendig $v_i = 0$, somit der entsprechende zyklische Modul trivial (d.h. besteht nur aus dem Nullelement). Sowas kann man getrost weglassen!

Um diese Zerlegung eines R-Moduls V zu erhalten sind also folgende Schritte zu tun.

Algorithmus Z.84 (Zerlegung eines Moduls)

(i) *Wähle ein Erzeugenden-System $(w_1, ..., w_n)$ von V.*

(ii) *Zu dem Modul-Homomorphismus*

$$\varphi : R^n \to V, \quad e_i \to w_i$$

bestimme ein Erzeugenden-System $(u_1, ..., u_m)$ von $U := \ker \varphi \subset R^n$. Dabei sind die e_i die kanonischen Einheitsvektoren im R^n.

(iii) *Bestimme die Koeffizienten-Matrix $A \in R^{n \times m}$, mit der*

$$(u_1, ..., u_m) = (e_1, ..., e_n)A$$

(Die Spalten von A sind einfach die u_j !)

(iv) *Transformiere A auf Diagonalform, d.h. bestimme invertierbare Matrizen B, C sodaß*

$$A' := B^{-1}AC = \text{diag}(\delta_1, ..., \delta_k, 0, ..., 0),$$

mit $\delta_i \neq 0$ und der Teilbarkeits-Bedingung.

(v) *Berechne die Basis*

$$(b_1, ..., b_n) = (e_1, ..., e_n)B.$$

(Wieder sind die b_i einfach die Spalten von B !)

(vi) *Die $v_i := \varphi(b_i)$ für $(i = 1, ..., k)$ erzeugen die zyklischen Moduln V_i mit den Ordnungen (δ_i), die restlichen v_i $(i > k)$ sind Basis für den freien Anteil V'.*

Struktursatz für endlich erzeugte abelsche Gruppen

Als erstes Anwendungsbeispiel betrachten wir abelsche Gruppen, die wir schon als Moduln über dem Ring $\mathbb{Z}$ der ganzen Zahlen kennengelernt hatten.

$\mathbb{Z}$ selbst, sowie die Restklassengruppen $\mathbb{Z}_{/d\mathbb{Z}}$ ($d \in \mathbb{Z},\ \geq 1$) sind offenbar zyklisch, nämlich von dem Element 1 erzeugt und haben Ordnungen 0 (Null) bei $\mathbb{Z}$ bzw. d bei $\mathbb{Z}_{/d\mathbb{Z}}$. Dies sind aber auch schon im wesentlichen alle nichttrivialen zyklischen abelschen Gruppen.

Satz Z.85 *(i) Jede zyklische (abelsche) Gruppe G der Ordnung $d \in \mathbb{Z}$ ist isomorph (als $\mathbb{Z}$-Modul) zur Gruppe $\mathbb{Z}_{/d\mathbb{Z}}$.*

(ii) Jede endlich erzeugte freie abelsche Gruppe ist isomorph zu $\mathbb{Z}^n$ für ein $n \in \mathbb{N}$.

Beweis:

(i) Es sei $G = (g) = \{zg \mid z \in \mathbb{Z}\}$ von Ordnung d.

(a) $d = 0$: Dann ist $zg = 0$ genau wenn $z \in (0)$, d.h. $z = 0$. Somit ist G frei mit Basis (g) und
$$\varphi : \mathbb{Z} \to G,\ 1 \to g$$
ist der gewünschte Isomorphismus.

(b) $d > 0$: Es ist $zg = 0$ genau wenn $z \in (d)$, d.h. wenn $d \mid z$. Betrachte den Homomorphismus
$$\varphi : \mathbb{Z}_{/d\mathbb{Z}} \to G,\ 1 \to g.$$
Dann ist im $\varphi = \{0, g, 2g, ..., (d-1)g\}$. Nach Definition der Ordnung ist $zg = z'g$ genau, wenn $z - z' \in (d)$, d.h. ein Vielfaches von d. Daraus erkennt man, daß die für im φ notierten Elemente alle verschieden sind und sich jedes Element von G darstellen läßt als zg mit $0 \leq z < d$. Somit ist wieder φ ein Isomorphismus.

(ii) Die Aussage zu den freien Gruppen zeigt man (simpel) über Basen. □

Damit können wir Satz Z.83 formulieren als

Satz Z.86 (Struktursatz für endlich erzeugte abelsche Gruppen)
Jede endlich erzeugte abelsche Gruppe G ist isomorph zu einer direkten Summe

$$\mathbb{Z}_{/d_1\mathbb{Z}} \oplus ... \oplus \mathbb{Z}_{/d_k\mathbb{Z}} \oplus \mathbb{Z}^n,$$

wobei $n \in \mathbb{Z},\ \geq 0$ und die Ordnungen $d_i \in \mathbb{Z},\ > 1$ sind und zudem die Teiler-Bedingung $d_1 \mid d_2 \mid ... \mid d_k$ gilt.

4

Zerlegung eines Raumes nach einem Endomorphismus II

Wir kommen nun wieder auf die Fragestellung vom Anfang des Kapitels zurück, einen K-Vektorraum V zu einem Endomorphismus $f \in \mathrm{Hom}(V, V)$ möglichst fein in f-invariante Unterräume zu zerlegen. Dazu fassen wir V auf als Modul über dem Hauptideal-Ring $K[T]$ mit der Skalar-Multiplikation $p(T) \cdot v = p(f)v$. Auf diese Situation passen dann die hergeleiteten Sätze.

Zunächst eine "Übersetzung":

Lemma Z.87 *Bei der Deutung von V als K-Vektorraum bzw. als $K[T]$- Modul entsprechen sich:*

(i) *"U ist f-invarianter Unterraum" und "U ist Untermodul".*

(ii) *"U_v ist f-zyklischer Unterraum" und "U_v ist zyklischer (von v erzeugter) Untermodul".*

(iii) *"ψ ist Minimalpolynom zu f auf U_v" und "Das Haupt-Ideal (ψ) ist Ordnungsideal zum Element v."*

Beweis als Übung.

Ferner sei an Satz Z.26 erinnert, wonach jedes Element $v \in V$ ein nichttriviales Ordnungs-Ideal besitzt und damit in V (als Modul) keine freie Familie existiert.

Auf diesen Modul können wir unseren Satz Z.83 anwenden. Dies führt zu

Satz Z.88 *Es sei V ein Vektorraum, $f \in \mathrm{Hom}(V,V)$ ein Endomorphismus. Dann gibt es eine Darstellung*

$$V = V_1 \oplus ... \oplus V_k$$

von V als direkte Summe f-zyklischer Unterräume V_i, für deren Minimalpolynome die Teilbarkeitsbeziehung

$$\psi_1 \mid \psi_2 \mid ... \mid \psi_k$$

gilt.

Diese Minimalpolynome sind dabei genau die (nichttrivialen) Invarianten-Teiler der charakteristischen Matrix $F - T \cdot I$ zu einer beliebigen Matrixdarstellung F von f und somit eindeutig bestimmt.

ψ_k ist das Minimalpolynom von f auf V.

Beweis: Die Existenz einer solchen Zerlegung einschließlich der Teilbarkeit der Minimalpolynome folgt direkt aus Satz Z.83.

Sei nun eine solche Zerlegung gegeben. In jedem Teilraum V_i kann man eine Basis wählen, bezüglich der $f_{|V_i}$ dargestellt wird durch die Begleitmatrix F_i zum Minimalpolynom ψ_i. Zusammen genommen ergeben diese Basen eine Basis $(w_1, ..., w_n)$ des Gesamt-Raumes V, bezüglich der dann f die Matrixdarstellung

$$F = \mathrm{diag}(F_1, ..., F_k)$$

hat. Dies ist nun genau die Situation von Satz Z.70, sodaß $\psi_1, ..., \psi_k$ genau die nichttrivialen Invarianten-Teiler der charakteristischen Matrix $F - T \cdot I$ sind. Die sind aber allein durch f bestimmt und nicht abhängig davon, mittels welcher Basis f durch eine Matrix dargestellt wird.

Ist ψ das Minimalpolynom von f auf ganz V, so ist nach Satz Z.22 das Polynom ψ_k als Minimalpolynom auf einem Teilraum selbst Teiler von ψ. Andrerseits läßt sich jedes $v \in V$ entsprechend der Zerlegung darstellen als

$$v = v_1 + ... + v_k,$$

wofür dann für jedes i gerade $\psi_i(f)v_i = 0$ und wegen $\psi_i \mid \psi_k$ dann auch $\psi_k(f)v_i = 0$ ist. Dann ist aber auch $\psi_k(f)v = 0$ und, da v beliebig war, also $\psi \mid \psi_k$, sodaß beide übereinstimmen müssen. □

Die in der Zerlegung von Satz Z.88 auftretenden Teilräume V_1 sind zwar f-zyklisch, brauchen aber keineswegs irreduzibel zu sein, sodaß sie evtl. noch nach Satz Z.30 weiter zerlegt werden können. Führen wir dies aus, so erhalten wir den angestrebten

Satz Z.89 (Zerlegung eines Raumes nach einem Endomorphismus)

Zu einem Vektorraum V mit einem Endomorphismus $f \in \mathrm{Hom}(V,V)$ gibt es eine Zerlegung

$$V = U_1 \oplus ... \oplus U_m$$

in irreduzible f-zyklische Unterräume. Das zugehörige System von Minimalpolynomen $\varphi_1, ..., \varphi_m$ ist bis auf die Reihenfolge und abgesehen von trivialen Teilen eindeutig bestimmt.

Es ist

$$\psi = \mathrm{kgV}(\varphi_1, ..., \varphi_m)$$

das Minimalpolynom von f auf V,

$$\chi = \varphi_1 \cdot ... \cdot \varphi_m$$

das charakteristische Polynom von f.

Insbesondere ist das Minimalpolynom ein Teiler des charakteristischen Polynoms.

Beweis: Nach Satz Z.30 läßt sich jeder f-zyklische Raum in eine direkte Summe irreduzibler f-invarianter Räume zerlegen, wobei diese Zerlegung der Darstellung des Minimalpolynoms als Potenzprodukt verschiedener irreduzibler Faktoren entspricht. Damit ergibt sich die Existenz der Zerlegung und die Eindeutigkeit des Systems der Minimalpolynome direkt aus Satz Z.88.

Als Minimalpolynome zu f auf Teilräumen von V sind die φ_i alles Teiler des Minimalpolynoms ψ von f auf ganz V. Zerlegen wir zunächst nur nach Satz Z.88, so ergibt sich bei der hier entstehenden Zerlegung $V = V_1 \oplus ... \oplus V_k$ mit den sich sukzessive teilenden Minimalpolynomen $\psi_1 \mid \psi_2 \mid ... \mid \psi_k$ ja einerseits $\psi_k = \psi$, andrerseits wird bei der anschließenden Verfeinerung nach Satz Z.30 hier ψ_k zerlegt in ein Produkt teilerfremder Polynome, die alle als Minimalpolynome φ_i bei der Zerlegung in irreduzible Räume auftreten. Damit ist dann $\psi = \mathrm{kgV}(\varphi_1, ..., \varphi_m)$.

Schließlich können wir für jedes U_i eine Basis so wählen, daß bezüglich dieser f durch die Begleitmatrix F_i zu φ_i dargestellt wird. Dies ergibt für f eine Darstellung als

$$f \hat{=} F = \mathrm{diag}(F_1, ..., F_m).$$

Dann ist

$$\chi_f(T) = \det(F - T \cdot I) = \prod_i \det(F_i - T \cdot I_i) = \prod_i \varphi_i(T),$$

da nach Satz Z.32 die Begleitmatrix zu einem Polynom $p \in K[T]$ genau p als charakteristisches Polynom hat.

Damit ist dann auch bewiesen, daß das Minimalpolynom ein Teiler des charakteristischen Polynoms ist. □

Diese letzte Aussage ist bekannt als

Satz Z.90 (HAMILTON-CAYLEY) *Ist $\chi_f(T)$ das charakteristische Polynom eines Vektorraum-Endomorphismus f, so ist*

$$\chi_f(f) = 0,$$

oder formuliert für Matrizen:

Für $\chi(T) := \det(F - T \cdot I)$ ist

$$\chi(F) = 0.$$

Wir beschließen dieses Kapitel, indem wir noch den Algorithmus Z.84 zur konstruktiven Ausführung von Satz Z.83 für die Situation des durch einen Endomorphismus f gegebenen $K[T]$-Moduls explizit durchführen.

Sei also V ein K-Vektorraum mit einer Basis $(w_1, ..., w_n)$, $f \in \mathrm{Hom}(V, V)$ ein Vektorraum-Endomorphismus. Trivialerweise ist die gegebene Vektorraum-Basis $(w_1, ..., w_n)$ auch $K[T]$-erzeugend für V.

Wir verfolgen nun Algorithmus Z.84:

Schritt 1: Als $K[T]$-Erzeugenden-System für V wähle die gegebene K-Basis $(w_1, ..., w_n)$.

Schritt 2: Als nächstes ist der Kern U des durch

$$\varphi : (K[T])^n \to V, \quad e_i \to w_i \quad (i = 1, ..., n)$$

gegebenen Modul-Homomorphismus zu studieren. Die e_i sind wieder die kanonischen Einheitsvektoren.

Dafür gilt

Lemma Z.91 *Ist $F = (\eta_{ij}) \in K^{n\times n}$ die Matrix-Darstellung von f bezüglich der Vektorraum-Basis $(w_1, ..., w_n)$ von V, d.h. ist*

$$f(w_j) = \sum_{i=1}^{n} \eta_{ij} w_i \qquad (j = 1, ..., n),$$

so bilden die Elemente (in $(K[T])^n$)

$$u_j(T) = \sum_{i=1}^{n} \eta_{ij} e_i - T e_j \qquad (j = 1, ..., n),$$

eine Basis des oben genannten $K[T]$-(Unter-)Moduls $U \subset (K[T])^n$.

Die $u_j(T)$ sind dabei genau die Spalten der charakteristischen Matrix zu F.

Beweis:

$u_j \in U$: Es ist $\varphi(u_j) = \varphi(\sum_{i=1}^{n} \eta_{ij} e_i - T e_j) = \sum_{i=1}^{n} \eta_{ij} w_i - f(w_j) = 0$.
Also liegen die u_j in $\ker \varphi = U$.

$(u_1, ..., u_n)$ *ist erzeugend für* U: Die e_i sind Basis von $(K[T])^n$, wovon U ein Untermodul ist. Somit hat jedes $u \in (K[T])^n$ eine Darstellung als

$$u(T) = \sum_{j=1}^{n} \theta_j(T) e_j \qquad \theta_j(T) \in K[T].$$

Nach Konstruktion ist

$$T e_j = \sum_{i=1}^{n} \eta_{ij} e_i - u_j(T) \qquad (j = 1, ..., n).$$

Zerlegen wir also

$$\theta_j(T) = \gamma_j(T) \cdot T + \rho_j \quad \text{mit } \rho_j \in K,$$

so ist

$$\begin{aligned} u(T) &= \sum_{j=1}^{n} \theta_j(T) e_j = \sum_{j=1}^{n} \gamma_j(T) \cdot T e_j + \sum_{j=1}^{n} \rho_j e_j \\ &= \sum_{j=1}^{n} \gamma_j(T) \left(\sum_{i=1}^{n} \eta_{ij} e_i - u_j(T) \right) + \sum_{j=1}^{n} \rho_j e_j \\ &= -\sum_{j=1}^{n} \gamma_j(T) u_j(T) + \sum_{j=1}^{n} \hat{\theta}_j(T) e_j, \end{aligned}$$

wobei der maximale Grad der $\hat{\theta}_j$ kleiner ist als der der θ_j. Dies kann man iterieren und erhält:

Jedes $u \in (K[T])^n$ hat eine Darstellung als

$$u(T) = \sum_{j=1}^{n} \gamma_j(T) u_j(T) + \sum_{j=1}^{n} \rho_j e_j \text{ mit } \rho_j \in K.$$

Ist jetzt $u \in U$, so ist $\varphi(u) = 0, \varphi(u_j) = 0$ für alle j, also

$$0 = \varphi(u) = \sum_{j=1}^{n} \gamma_j(f) \varphi(u_j) + \sum_{j=1}^{n} \rho_j \varphi(e_j) = \sum_{j=1}^{n} \rho_j w_j,$$

und da die w_j eine K-Basis von V bilden, sind alle $\rho_i = 0$.

Somit besitzt jedes $u \in U = \ker \varphi$ eine Darstellung

$$u(T) = \sum_{j=1}^{n} \gamma_j(T) u_j(T),$$

sodaß die u_j ganz U erzeugen.

$(u_1, ..., u_n)$ *ist* $K[T]$*-unabhängig:* Wähle $\theta_j(T) \in K[T]$, sodaß

$$0 = \sum_{j=1}^{n} \theta_j(T) u_j(T) = \sum_{i,j=1}^{n} \theta_j(T) \eta_{ij} e_i - \sum_{j=1}^{n} \theta_j(T) T e_j.$$

Dies ist eine Gleichung in $K[T]^n$, d.h. für jedes k gilt in der k-ten Komponente die Polynomgleichung

$$0 = \sum_{j=1}^{n} \theta_j(T) \eta_{kj} - \theta_k(T) T.$$

Für ein θ_k mit maximalem Grad ist dies aber aus Gradgründen ein Widerspruch, sofern es sich um ein nichttriviales Polynom handelt. Also müssen alle $\theta_j = 0$ sein, was die Unabhängigkeit zeigt. □

Fahren wir im Algorithmus Z.84 fort.

Schritt 3: Bestimme die Koeffizienten-Matrix A für die

$$(u_1, ..., u_n) = (e_1, ..., e_n) A.$$

Wegen

$$u_j(T) = \sum_{i=1}^{n} \eta_{ij} e_i - T e_j$$

ist dies die Matrix

$$A = F - T \cdot I,$$

also die "charakteristische Matrix" des zugrunde liegenden Endomorphismus f.

Schritt 4: Diese Matrix $A \in (K[T])^{n \times n}$ haben wir nun noch zu diagonalisieren und die zugehörigen Transformationsmatrizen $B, C \in (K[T])^{n \times n}$ zu bestimmen. Da U eine Basis der Länge n hat, entsteht hier die Diagonalmatrix

$$A' = B^{-1} A C = \text{diag}(\delta_1, ..., \delta_n), \text{ wobei alle } \delta_i \in K[T], \neq 0 \text{ und } \delta_1 \mid \delta_2 \mid ... \mid \delta_n.$$

Schritt 5 und 6: Die Spalten $(b_1, ..., b_n)$ von B sind dann die gesuchte Basis, sodaß

$$v_i := \varphi(b_i) \quad (i = 1, ..., n)$$

die Erzeugenden für die zyklischen Moduln mit den Ordnungen δ_i sind.

BZ Beispiele zur Modulzerlegung

1

BZ1: Auf der Schule hatten Sie (reelle) Polynome eigentlich nur unter dem Aspekt einer besonders bequemen Abbildungsvorschrift für Funktionen $\mathbb{R} \to \mathbb{R}$ angesehen. Dieser Standpunkt sollte jetzt endgültig verlassen werden. Für einen Körper K ist ein Polynom $p \in K[T]$ ein Element der eingeführten Polynom-Algebra, auf das man für jede beliebige weitere K-Algebra A einen Einsetzhomomorphismus

$$\varphi_a : K[T] \to A;\ \varphi_a(p) = p(a)$$

anwenden kann. Speziell kann man auch als Algebra A den Körper K selbst nehmen.

Nehmen wir etwa den Körper F_2 mit den beiden Elementen $\{0, 1\}$, so ist für das Polynom $p(T) := T(T-1)$ offenbar $p(0) = 0, p(1) = 0$, d.h. unter den beiden nach $A := K := F_2$ möglichen Einsetzhomomorphimsen $\varphi_0, \varphi_1 : F_2[T] \to F_2$ kommt hier stets der Wert 0 heraus, obwohl das Polynom selbst nicht das 0-Element in $F_2[T]$ ist.

Wählt man hier als Algebra A etwa die Matrixalgebra $K^{2\times 2}$, so kann man über den Einsetzhomomorphismus jetzt eine 2×2-Matrix in ein Polynom einsetzen und da kann dann evtl. auch wieder die Nullmatrix herauskommen. So ist beispielsweise für $p(T) = T^2 - 10T + 24,\ a := \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}$

$$\begin{aligned} p(a) &= a^2 - 10 \cdot a + 24 \cdot I \\ &= \begin{pmatrix} 26 & -10 \\ -10 & 26 \end{pmatrix} - 10 \cdot \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix} + 24 \cdot \begin{pmatrix} 1 & -0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Dieses Polynom hat also die Matrix $a := \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}$ als "Nullstelle".

BZ2: Zu Euklids Algorithmus:

Mit den beiden Polynomen $p_1(T) = T^6 - 1,\ p_2(T) = T^4 + T^3 + 2T^2 + T + 1$ erhalten wir aus $p_1(T) = q_1(T)p_2(T) + p_3(T)$:

$$T^6 - 1 = (T^2 - T - 1)(T^4 + T^3 + 2T^2 + T + 1) + 2(T^3 + T^2 + T),$$

d.h. $p_3(T) = 2(T^3 + T^2 + T)$.

Der nächste Schritt ist dann

$$p_2(T) = q_2(T)p_3(T) + p_4(T),$$

d.h.

$$T^4 + T^3 + 2T^2 + T + 1 = (\frac{1}{2}T)2(T^3 + T^2 + T) + (T^2 + T + 1),$$

was $p_4(T) = T^2 + T + 1$ ergibt.

Damit folgt dann schließlich

$$2(T^3 + T^2 + T) = (2T)(T^2 + T + 1),$$

wobei kein Rest bleibt. Also ist

$$d(T) := T^2 + T + 1$$

der letzte nicht verschwindende Rest und damit ein ggT(p_1, p_2). Er besitzt die Darstellung aus den Anfangspolynomen

$$T^2 + T + 1 = (-\frac{1}{2}T) \cdot p_1 + \frac{1}{2}(T^3 - T^2 - T + 2) \cdot p_2.$$

Weiter bis

2

2

BZ3: Wir betrachten den Homomorphismus $f : \mathbb{R}^4 \to \mathbb{R}^4$, der bezüglich der kanonischen Basis durch die Matrix $F := \begin{pmatrix} 0 & -2 & 4 & -1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 1 & -2 & -1 \end{pmatrix}$ vermittelt wird.

Dafür ist $F^2 = \begin{pmatrix} -1 & -5 & 10 & 1 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ -1 & -1 & 2 & 0 \end{pmatrix}$, $F^3 = \begin{pmatrix} 1 & -7 & 14 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Man verifiziert leicht, daß $F + F^2 = F^3 - 2I$ ist. Damit annulliert das Polynom

$$p(T) := T^3 - T^2 - T - 2 = (T - 2)(T^2 + T + 1)$$

den Homomorphismus f, d.h. es ist $p(f) = 0$.

Die zweiten Spalten von I, F, F^2, also die Vektoren $e_2, f(e_2) = Fe_2, f^2(e_2) = F^2e_2$ sind offenbar linear unabhängig, sodaß es kein nichttriviales Polynom $q(T)$ vom Grad ≤ 2 geben kann, für das $q(f) = 0$ ist. (Denn dafür wäre dann auch $q(f)(e_2) = \beta_2 f^2(e_2) + \beta_1 f(e_2) + \beta_0 e_2 = 0$, etc....)

Also haben wir in $p(T)$ das Minimalpolynom gefunden. Es ist also für unseren Homomorphismus f das Polynom

$$\psi_f(T) := (T - 2)(T^2 + T + 1)$$

das Minimalpolynom.

BZ4: Zu dem Homomorphismus f aus BZ3 lautet das charakteristische Polynom

$$\chi_f(T) := (T - 2)^2(T^2 + T + 1)$$

Es wird also vom Minimalpolynom geteilt, genauer ist

$$\chi_f(T) := (T - 2)\psi_f(T)$$

und damit auch

$$\chi_f(f) = (f - \mathrm{id}) \circ \psi_f(f) = (f - \mathrm{id}) \circ 0 = 0.$$

Das ist ein Beispiel für den Satz Z.90, der besagt, daß für jeden Homomrphismus f stets $\chi_f(f) = 0$ ist.

BZ5: Bleiben wir bei dem in BZ3 begonnenen Beispiel. Nach Satz Z.22 ist das Minimalpolynom ψ' zu einem f-invarianten Unterraum $U \subset \mathbb{R}^4$ ein Teiler von $\psi_f(T) = (T - 2)(T^2 + T + 1)$, sodaß hier an nichttrivialen Teilerpolynomen gerade

$$\psi_1(T) := T - 2 \text{ und } \psi_2(T) := T^2 + T + 1$$

infrage kommen. Ferner sind nach dem selben Satz die Räume $U_i := \ker \psi_i(f)$ selbst f-invariant und es gilt $R^4 = U_1 \oplus U_2$. Diese Räume kann man ausrechnen. Es ist

$$\begin{array}{lll} U_1 & := \ker \psi_1(f) & = \ker(f - 2\,\mathrm{id}) & = \ker(F - 2I) \\ U_2 & := \ker \psi_2(f) & = \ker(f^2 + f + \mathrm{id}) & = \ker(F^2 + F + I). \end{array}$$

Hier sind einfach homogene Geleichungssysteme zu lösen und man erhält

$$\begin{aligned} U_1 &= \mathrm{span}(e_1 - e_2, e_1 + e_2 + e_3), \\ U_2 &= \mathrm{span}(e_1, e_4). \end{aligned}$$

Die vier hierbei notierten Vektoren bilden eine Basis des $\mathbb{R}^4$, womit – wie oben behauptet – dann $U_{12} \oplus U_2 = \mathbb{R}^4$.

Als Kerne von Polynomen in f sind U_1 und U_2 je f-invariant. Rechnen wir dies auch direkt nach: Es ist

$$f(e_1 - e_2) = F(e_1 - e_2) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} -2 \\ 2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 0 \\ 0 \end{pmatrix} = 2(e_1 - e_2),$$

ferner

$$f(e_1+e_2+e_3) = F(e_1+e_2+e_3) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -2 \\ 2 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 4 \\ 0 \\ 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 0 \end{pmatrix} = 2(e_1+e_2+e_3).$$

Damit sind beides Eigenvektoren von f zum Eigenwert 2 und ganz U_1 besteht somit nur aus solchen, d.h. $f_{|U_1} = 2\,\mathrm{id}_{U_1}$. Also hat $f_{|U_1}$ bezüglich jeder Basis von U_1 die Matrixdarstellung

$$f_{|U_1} \hat{=} F_1 := \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Für die Basis (e_1, e_4) von U_2 haben wir

$$\begin{array}{lll} f(e_1) & = Fe_1 & = e_4, \\ f(e_4) & = Fe_4 & = -(e_1 + e_4). \end{array}$$

Beides liegt wieder in U_2, sodaß sich U_2 als f-invariant erweist. Bezüglch dieser Basis haben wir also die Matrixdarstellung

$$f_{|U_2} \hat{=} F_2 := \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}.$$

Daraus gewinnt man gemäß Satz Z.23 die Zerlegung

$$\mathbb{R}^4 = U_1 \oplus U_2$$

und

$$f \hat{=} F' := \left(\begin{array}{cc|cc} 2 & 0 & & \\ 0 & 2 & & \\ \hline & & 0 & -1 \\ & & 1 & -1 \end{array}\right).$$

BZ6: Der in BZ5 erhaltene Raum $U_2 = \mathrm{span}(e_1, e_4)$ ist f-invariant und wegen $e_4 = f(e_1)$ auch zyklisch, erzeugt von e_1; d.h. $U_2 = \mathrm{span}(e_1, f(e_1))$.

Ist $u_1 \in U_1, \neq 0$, d.h. ein Eigenvektor zum Eigenwert 2, so ist auch der Raum

$$U := \text{span}(u_1, e_1, e_4)$$

f-zyklisch; denn U ist f-invariant wegen

$$f(u_1) = 2u_1 \in U, f(e_1) = e_4 \in U, f(e_4) = -(e_1 + e_4) \in U$$

und darüber hinaus haben wir etwa in U die Vektoren

$$v := u_1 + e_1,\ f(v) = 2u_1 + e_4,\ f^2(v) = 4u_1 - (e_1 + e_4),$$

und, da u_1, e_1, e_4 unabhängig sind, sind auch die letztgenannten Vektoren unabhängig und damit eine Basis von U, sodaß dieser Raum als zyklisch nachgewiesen ist.

Der ganze Raum $\mathbb{R}^4$ ist aber nicht f-zyklisch, was man folgendermaßen sieht: Nach BZ5 bilden die Vektoren

$$(b_1, b_2, b_3, b_4) := (e_1 - e_2, e_1 + e_2 + e_3, e_1, e_4)$$

eine Basis des $\mathbb{R}^4$, wobei b_1, b_2 Eigenvektoren zum Eigenwert 2 sind und b_3, b_4 den f-invarianten Raum U_2 aufspannen. Jeder Vektor $v \in \mathbb{R}^4$ hat eine Darstellung als

$$v = \beta_1 b_1 + \beta_2 b_2 + \beta_3 b_3 + \beta_4 b_4.$$

Dafür ist

$$f(v) = 2\beta_1 b_1 + 2\beta_2 b_2 + \beta_3' b_3 + \beta_4' b_4$$

und allgemein

$$f^k(v) = 2^k \beta_1 b_1 + 2^k \beta_2 b_2 + \beta_3^{(k)} b_3 + \beta_4^{(k)} b_4$$

und alle diese Vektoren liegen in

$$\text{span}((\beta_1 b_1 + \beta_2 b_2), b_3, b_4).$$

Damit ist auch

$$U_v := \text{span}(v, f(v), f^2(v), \ldots) \subset \text{span}((\beta_1 b_1 + \beta_2 b_2), b_3, b_4) \subsetneqq \mathbb{R}^4.$$

BZ7: Betrachten wir nochmal den in BZ6 eingeführten Unterraum $U := \text{span}(u_1, e_1, e_4)$, wobei u_1 ein Eigenvektor zum Eigenwert 2 von f ist. Er ist f-zyklisch mit Basis

$$(b_1, b_2, b_3) := (v, f(v), f^2(v)) := (u_1 + e_1, 2u_1 + e_4, 4u_1 - (e_1 + e_4)).$$

Es ist

$$\begin{aligned} f(b_1) &= f(v) = b_2, \\ f(b_2) &= f(f(v)) = f^2(v) = b_3, \\ f(b_3) &= f(4u_1 - (e_1 + e_4)) = 8u_1 - e_4 - (-e_1 - e_4) \\ &= 8u_1 + e_1 = 2b_1 + b_2 + b_3. \end{aligned}$$

Also hat $f_{|U}$ bezüglich dieser Basis die Matrixdarstellung

$$f_{|U} \hat{=} F' := \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

Dies ist die Begleitmatrix zu dem Polynom

$$\psi(T) := T^3 - (2 + T + T^2) = (T - 2)(T^2 + T + 1).$$

Es ist dies das Minimalpolynom zu f auf dem f-zyklischen Raum U. Es ist nicht irreduzibel. Wenden Sie nun auf diese Situation den Satz Z.30 an und gewinnen Sie daraus die Zerlegung in irreduzible Räume!

Weiter bis

3

3

BZ8: Nach Definition Z.50 ist der Ring $K[T]$ der Polynome in einer Veränderlichen T über dem Körper K stets Hauptidealring. Ersetzt man hier K durch einen allgemeinen Ring oder nimmt man mehrere Veränderliche, so stimmt diese nicht mehr.

Betrachten wir etwa $\mathbb{R}[X, Y]$, d.h. die reellen Polynome in 2 Veränderlichen. Sie bilden einen kommutativen Ring mit 1, der *kein* Hauptidealring ist. Um letzteres zu sehen betrachten wir das von den beiden Veränderlichen X, Y erzeugte Ideal:

$$\mathcal{I} := (X, Y) := \{p(X, Y) \cdot X + q(X, Y) \cdot Y \mid p, q \in \mathbb{R}[X, Y]\}.$$

Daß dies ein Ideal im Sinne von Definition Z.44 ist, rechnen Sie mal selber nach. Es enthält genau alle Polynome, deren absoluter Term verschwindet.

Wäre $\mathcal{I}$ ein Hauptidealring, so gäbe es ein Polynom $a(X, Y) \in \mathbb{R}[X, Y]$, sodaß

$$\mathcal{I} = (a(X, Y)) = \{r(X, Y) \cdot a(X, Y) \mid r \in \mathbb{R}[X, Y]\}.$$

Wegen $X \in \mathcal{I}$ gäbe es dann ein r_1 sodaß $X = r_1 \cdot a(X, Y)$. Da $a(X, Y)$ kein Absolutglied hat, ginge dies (bis auf Trivialitäten) nur für $r_1 = 1, a(X, Y) = X$. Damit wäre aber $Y = r_2 \cdot X$, was unmöglich ist.

BZ9: Es bezeichne $\mathbb{Z}[i]$ die "ganzen GAUSSschen Zahlen", d.h. die komplexen Zahlen mit ganzzahligem Real- und Imaginärteil. Mit den von $\mathbb{C}$ ererbten Operationen ist dies ein kommutativer Ring mit 1. Wir definieren die Abbildung

$$\nu : \mathbb{Z}[i] \to \mathbb{N}; x := x_1 + ix_2 \mapsto \nu(x) := x_1^2 + x_2^2 = (x_1 + ix_2)(x_1 - ix_2) = x \cdot \overline{x}.$$

Mit dieser "Normfunktion" wird $\mathbb{Z}[i]$ ein EUKLID-Ring. Eine Division mit Rest, die den Anforderungen von Definition Z.50 genügt, erhält man so:

Zu $a := a_1 + ia_2 \neq 0, b := b_1 + ib_2$ gilt in $\mathbb{C}$

$$b = \frac{b}{a} \cdot a = \frac{\overline{a} \cdot b}{\overline{a} \cdot a} \cdot a = \frac{\overline{a} \cdot b}{\nu(a)} \cdot a.$$

Dabei ist $\overline{a} \cdot b = (a_1 - ia_2)(b_1 + ib_2) =: c_1 + c_2$, d.h. $\frac{\overline{a} \cdot b}{\nu(a)} = \frac{c_1}{\nu(a)} + i\frac{c_2}{\nu(a)}$.

Nun sind $\frac{c_1}{\nu(a)}, \frac{c_2}{\nu(a)}$ i.a. keine ganzen Zahlen, sondern nur rational. Aber wir haben eine Darstellung

$$\frac{\overline{a} \cdot b}{\nu(a)} = \frac{c_1}{\nu(a)} + i\frac{c_2}{\nu(a)} = (q_1 + \epsilon_1) + i(q_2 + \epsilon_2)$$

mit $q_1, q_2 \in \mathbb{Z}$ und $|\epsilon_1|, |\epsilon_2| \leq \frac{1}{2}$. Damit ist dann $b = (q_1 + iq_2)a + (\epsilon_1 + i\epsilon_2)a$, also $b - (q_1 + iq_2)a \in \mathbb{Z}[i]$, sodaß auch $r := (\epsilon_1 + i\epsilon_2)a \in \mathbb{Z}[i]$ und damit $b = (q_1 + iq_2)a + r$ in $\mathbb{Z}[i]$ gilt. Und hierfür ist

$$\nu(r) = (\epsilon_1^2 + \epsilon_2^2)\nu(a) \leq (\frac{1}{4} + \frac{1}{4})\nu(a) = \frac{1}{2}\nu(a) < \nu(a).$$

BZ10: Ein Beispiel für den Invariantenteilersatz über $\mathbb{Z}$ und den Algorithmus Z.64: Wir diagonalisieren die folgende Matrix

$$A := \begin{pmatrix} 2 & 2 & 0 & 2 & 2 \\ 2 & 0 & 0 & 2 & 3 \\ 6 & 0 & 6 & 0 & 6 \\ 4 & 2 & 0 & 4 & 5 \end{pmatrix}.$$

Ausräumen der ersten Zeile und Spalte nach Schritt 2 liefert

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 1 \\ 0 & -6 & 6 & -6 & 0 \\ 0 & -2 & 0 & 0 & 1 \end{pmatrix}.$$

Verwende Schritt 3a: addiere die zweite Zeile zur ersten:

$$\begin{pmatrix} 2 & -2 & 0 & 0 & 1 \\ 0 & -2 & 0 & 0 & 1 \\ 0 & -6 & 6 & -6 & 0 \\ 0 & -2 & 0 & 0 & 1 \end{pmatrix}.$$

Schritt 1: Tausche Spalten 1 und 5:

$$\begin{pmatrix} 1 & -2 & 0 & 0 & 2 \\ 1 & -2 & 0 & 0 & 0 \\ 0 & -6 & 6 & -6 & 0 \\ 1 & -2 & 0 & 0 & 0 \end{pmatrix}.$$

Schritt 2: Ausräumen der ersten Zeile und Spalte:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 \\ 0 & -6 & 6 & -6 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}.$$

Schritt 3b: Setze $\delta_1 := 1$ und mache weiter mit der Restmatrix

$$\begin{pmatrix} 0 & 0 & 0 & -2 \\ -6 & 6 & -6 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

Schritt 1: Tausche Spalten 1 und 4:

$$\begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 6 & -6 & -6 \\ -2 & 0 & 0 & 0 \end{pmatrix}.$$

Ausräumen der ersten Spalte

$$\begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 6 & -6 & -6 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Schritt 3b: Setze $\delta_2 := -2$ und mache weiter mit der Restmatrix

$$\begin{pmatrix} 6 & -6 & -6 \\ 0 & 0 & 0 \end{pmatrix}.$$

Schritt 2: Ausräumen der ersten Zeile:

$$\begin{pmatrix} 6 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Damit haben wir als Normalform

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

erhalten.

Bestimmen wir noch die Determinantenteiler

$$d_j(A) := \text{ggT}(j - \text{reihige Unterdeterminanten von} A).$$

Es ist $d_1(A)$ der ggT der Matrixelemente, also $d_1(A) = 1$.

Es ist $d_2(A) = 2$; denn die 2×2-Unterdeterminante in der rechten oberen Ecke hat den Wert 2 und in jeder 2×2-Untermatrix stehen in einer Zeile oder Spalte nur gerade Zahlen, sodaß jede 2×2-Unterdeterminante durch 2 teilbar ist.

Es ist $d_3(A) = 12$; denn die 3×3-Unterdeterminante in der rechten oberen Ecke hat den Wert 12. Bei den 3×3-Unterdeterminanten, die die dritte Zeile enthalten, liefert diese eine Faktor 6, ein weiterer Faktor 2 kann jeweils leicht gefunden werden. Somit sind diese alle durch 12 teilbar. Die 3×3-Unterdeterminanten, die die dritte Zeile nicht enthalten, sind alle $= 0$, da die vierte Zeile die Summe der ersten beiden ist.

Mit der letzten Bemerkung folgt dann auch, daß alle 4×4-Unterdeterminanten $= 0$ sind, d.h. $d_4(A) = 0$.

Vergleichen Sie dazu den Satz Z.68.

BZ11: Nun noch ein Beispiel zum Invariantenteilersatz über $\mathbb{R}[T]$:

Mit der Matrix F aus BZ3 betrachten wir

$$A := F - T \cdot I = \begin{pmatrix} -T & -2 & 4 & -1 \\ 0 & 2-T & 0 & 0 \\ 0 & 0 & 2-T & 0 \\ 1 & 1 & -2 & -1-T \end{pmatrix}$$

und bestimmen die Normalform.

Tausch der Zeilen 1 und 4 liefert

$$\begin{pmatrix} 1 & 1 & -2 & -1-T \\ 0 & 2-T & 0 & 0 \\ 0 & 0 & 2-T & 0 \\ -T & -2 & 4 & -1 \end{pmatrix}$$

Darin kann man nun die erste Zeile und erste Spalte ausräumen und erhält

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2-T & 0 & 0 \\ 0 & 0 & 2-T & 0 \\ 0 & -2+T & 4-2T & -T^2-T-1 \end{pmatrix}.$$

Also ist $\delta_1 = 1$ und wir können die erste Zeile und erste Spalte streichen. Dies liefert

$$\begin{pmatrix} 2-T & 0 & 0 \\ 0 & 2-T & 0 \\ -(2-T) & 2(2-T) & -(T^2+T+1) \end{pmatrix}.$$

Hierin ist $2-T$ ein Polynom niedrigsten Grades, sodaß wir damit Zeile 1 und Spalte 1 ausräumen können. Dies liefert

$$\begin{pmatrix} 2-T & 0 & 0 \\ 0 & 2-T & 0 \\ 0 & 2(2-T) & -(T^2+T+1) \end{pmatrix}.$$

Nun sind aber die Polynome $2-T$ und T^2+T+1 teilerfremd. Somit ist die dritte Zeile zu der ersten zu addieren:

$$\begin{pmatrix} 2-T & 2(2-T) & -(T^2+T+1) \\ 0 & 2-T & 0 \\ 0 & 2(2-T) & -(T^2+T+1) \end{pmatrix}.$$

Es ist $-(T^2+T+1) = (T+3)(2-T) - 7$, sodaß durch Subtraktion von Vielfachen der ersten Spalte sich

$$\begin{pmatrix} 2-T & 0 & -7 \\ 0 & 2-T & 0 \\ 0 & 2(2-T) & -(T^2+T+1) \end{pmatrix}$$

ergibt. Hierin ist -7 ein Polynom niedrigsten Grades. Wir haben also die Spalten 1 und 3 zu tauschen und erhalten

$$\begin{pmatrix} -7 & 0 & 2-T \\ 0 & 2-T & 0 \\ -(T^2+T+1) & 2(2-T) & 0 \end{pmatrix}.$$

Ausräumen der ersten Zeile und Spalte liefert

$$\begin{pmatrix} -7 & 0 & 0 \\ 0 & 2-T & 0 \\ 0 & 2(2-T) & -\frac{1}{7}(2-T)(T^2+T+1) \end{pmatrix}.$$

(-7) ist eine Einheit in $\mathbb{R}[T]$, somit ist $\delta_2 = -7$ und es bleibt

$$\begin{pmatrix} 2-T & 0 \\ 2(2-T) & -\frac{1}{7}(2-T)(T^2+T+1) \end{pmatrix}.$$

Hier ist alles durch $(2-T)$ teilbar, sodaß wir sofort auf

$$\begin{pmatrix} 2-T & 0 \\ 0 & -\frac{1}{7}(2-T)(T^2+T+1) \end{pmatrix}$$

kommen. Dies liefert

$$\delta_3 = (2-T),\ \delta_4 = -\frac{1}{7}(2-T)(T^2+T+1).$$

Da es auf die Multiplikation mit Einheiten nicht ankommt, können wir diese noch etwas glätten und erhalten als Normalform die Diagonalmatrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & (T-2) & 0 \\ 0 & 0 & 0 & (T-2)(T^2+T+1) \end{pmatrix}.$$

Berechnen wir noch die Determinantenteiler.

Trivialerweise ist $d_1(A) = 1$.

Unter den 2×2-Unterdeterminanten finden wir unter anderem

$$\det \begin{pmatrix} -T & -2 \\ 1 & 1 \end{pmatrix} = -T + 2$$
$$\det \begin{pmatrix} 4 & -1 \\ -2 & -1-T \end{pmatrix} = -4 - 4T - 2 = -2(2T+3),$$

die offenbar teilerfremd sind. Damit ist $d_2(A) = 1$.

An jeder 3×3-Unterdeterminante ist die zweite oder dritte Zeile beteiligt, sodaß sie sicher durch $T - 2$ teilbar ist. Spezielle solche Unterdeterminanten sind

$$\det \begin{pmatrix} -T & -2 & 4 \\ 0 & 0 & 2-T \\ 1 & 1 & -2 \end{pmatrix} = -(2-T)(-T+2)$$

und

$$\det \begin{pmatrix} -2 & 4 & -1 \\ 2-T & 0 & 0 \\ 1 & -2 & -1-T \end{pmatrix} = 2(2-T)(2T+3),$$

deren ggT gerade $T - 2$ ist. Also ist $d_3(A) = T - 2$.

Schließlich ist $d_4(A) = \det(A) = (T-2)^2(T^2+T+1)$.

Weiter bis

4

4

BZ12: Als die "KLEINsche Vierergruppe" bezeichnet man die Gruppe V mit vier Elementen $0, a, b, c$ für die die folgende Additionstabelle besteht:

$+$	0	a	b	c
0	0	a	b	c
a	a	0	c	b
b	b	c	0	a
c	c	b	a	0

Wir führen damit den Algorithmus Z.84 aus.

1.) Es ist (a, b) ein Erzeugendensystem, d.h. deren Linearkombinationen mit Koeffizienten in $\mathbb{Z}$ sind ganz V. Es hat zwei Elemente.

2.) Der $\mathbb{Z}$-Modul-Homomorphismus $\varphi : \mathbb{Z}^2 \to V$, $\varphi(e_1) = a, \varphi(e_2) = b$ hat als Kern genau $U := \operatorname{span}(2e_1, 2e_2)$. Denn jedes $u \in \mathbb{Z}^2$ hat eine Darstellung als

$$u = (2\alpha_1 + \epsilon_1)e_1 + (2\alpha_2 + \epsilon_2)e_2$$

mit $\alpha_1, \alpha_2 \in \mathbb{Z}, \epsilon_1, \epsilon_2 \in \{0, 1\}$. Dann ist aber

$$\begin{aligned} \varphi(u) &= (2\alpha_1 + \epsilon_1)\varphi(e_1) + (2\alpha_2 + \epsilon_2)\varphi(e_2) \\ &= (2\alpha_1 + \epsilon_1)a + (2\alpha_2 + \epsilon_2)b \\ &= \epsilon_1 a + \epsilon_2 b, \end{aligned}$$

da ja $2a = a + a = 0 = b + b = 2b$. Wegen $\epsilon_1, \epsilon_2 \in \{0, 1\}$ ist dann nach der Additionstabelle also $\varphi(u) = \epsilon_1 a + \epsilon_2 b = 0$ genau dann, wenn $\epsilon_1 = \epsilon_2 = 0$. Also ist $(u_1, u_2) := (2e_1, 2e_2)$ Basis von $U = \ker \varphi$.

3.) Es ist $(u_1, u_2) := (e_1, e_2)\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, d.h.

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

4.) Diese Matrix hat schon die gewünschte Normalform, d.h. wir können

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

wählen.

5.) Die gesuchte Basis ist dann $(b_1, b_2) = (e_1, e_2)$.

6.) Es ist

$$(v_1, v_2) := (\varphi(b_1), \varphi(b_2)) = (a, b),$$

d.h. die von a und b erzeugten Moduln V_a, V_b haben damit die Ordnungen 2 und ihre direkte Summe ist V. Ein freier Anteil tritt nicht auf. Mit Satz Z.86 ist also

$$V = \mathbb{Z}_{/2\mathbb{Z}} \oplus \mathbb{Z}_{/2\mathbb{Z}}.$$

(Für *dieses* Beispiel hätte scharfes Hinsehen natürlich auch ohne die ganze Theorie dieses Ergebnis gebracht.)

Weiter bis

5

5

BZ13: Wir demonstrieren die Zerlegung des Raumes nach einem Endomorphismus an dem schon in BZ3 und BZ11 behandelten Beispiel:

$$f : \mathbb{R}^4 \to \mathbb{R}^4 : x \mapsto f(x) := Fx \text{ mit}$$

$$F = \begin{pmatrix} 0 & -2 & 4 & -1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 1 & -2 & -1 \end{pmatrix}$$

Wir verwenden den Algorithmus Z.84 und arbeiten in dem $K[T]$-Modul $V := \mathbb{R}^4$ mit der skalaren Multiplikation $p(T) \cdot v := (p(f))(v)$.

Schritt (i) und (ii): Wir haben ein Erzeugendensystem $(w_1, \ldots, w_4)$ von V zu wählen und den durch

$$\varphi : (K[T])^4 \to V : e_i \mapsto w_i$$

gegebenen Homomorphismus zu bestimmen.

Wählt man als w_i die kanonischen Einheitsvektoren im $\mathbb{R}^4$, so bilden nach Lemma Z.91 die Spalten u_j der charakteristischen Matrix

$$(u_1, \ldots, u_4) := F - T \cdot I$$

eine Basis von $U := \ker \varphi \subset (K[T])^4$.

Schritt (iii): Die hier zu betrachtende Matrix ist also

$$A := F - T \cdot I.$$

Schritt (iv): Diese Matrix ist nun auf Diagonalform zu bringen. Die Diagonalform selbst haben wir schon in BZ11 bestimmmt. Wir brauchen aber auch die transformierenden Matrizen. Dafür beachten wir, daß der Tausch von Spalten oder das

Addieren eines Vielfachen einer Spalte zu einer anderen, wie man es zum Ausräumen von Zeilen braucht, die Multiplikation mit einer geeigneten Matrix von rechts bedeuten, und die entsprechenden Operationen mit Zeilen durch Multiplikation von links realisiert werden. Man braucht also nur die bei den jeweiligen Einzelschritten entstehenden Matrizen aufzumultiplizieren und erhält so die Transformationsmatrizen B^{-1} bzw. C.

Zur Abkürzung setzen wir

$$\alpha := 2 - T,\ \beta := -1 - T - T^2,$$

wofür dann

$$\beta + 7 = (T+3)\alpha$$

ist. Unsere Matrix lautet damit

$$\begin{pmatrix} -T & -2 & 4 & -1 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 1 & 1 & -2 & -(1+T) \end{pmatrix}$$

Aus den ersten Operationen ergibt sich dafür

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ T & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -T & -2 & 4 & -1 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 1 & 1 & -2 & -(1+T) \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 & 1+T \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & T \end{pmatrix} \begin{pmatrix} -T & -2 & 4 & -1 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 1 & 1 & -2 & -(1+T) \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 & 1+T \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 0 & -\alpha & 2\alpha & \beta \end{pmatrix} =: A_1.$$

Mit

$$B_1^{-1} := \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & T \end{pmatrix}$$

und

$$C_1 := \begin{pmatrix} 1 & -1 & 2 & 1+T \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

ist also $A_1 = B_1^{-1} A C_1$.

Nun kann man die erste Zeile und erste Spalte von A_1 zunächst weglassen und mit der Restmatrix

$$A_1' := \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ -\alpha & 2\alpha & \beta \end{pmatrix}$$

weitermachen. Die in BZ11 durchgeführten Schritte liefern

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{7}\beta & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ -\alpha & 2\alpha & \beta \end{pmatrix} \cdot$$
$$\cdot \begin{pmatrix} 1 & -2 & -(T+3) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & \frac{1}{7}\alpha \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ \frac{2}{7}\beta+1 & 0 & \frac{1}{7}\beta+1 \end{pmatrix} \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ -\alpha & 2\alpha & \beta \end{pmatrix} \begin{pmatrix} -(T+3) & -2 & -\frac{1}{7}\beta \\ 0 & 1 & 0 \\ 1 & 0 & \frac{1}{7}\alpha \end{pmatrix}$$
$$= \begin{pmatrix} -7 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 2\alpha & \frac{1}{7}\alpha\beta \end{pmatrix} =: A_2'.$$

Mit den aufmultiplizierten linken bzw. rechten Matrizen

$$(B_2')^{-1} := \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ \frac{2}{7}\beta+1 & 0 & \frac{1}{7}\beta+1 \end{pmatrix},$$

$$C_2' := \begin{pmatrix} -(T+3) & -2 & -\frac{1}{7}\beta \\ 0 & 1 & 0 \\ 1 & 0 & \frac{1}{7}\alpha \end{pmatrix}$$

ist also

$$A_2' = (B_2')^{-1} \cdot A_1' \cdot C_2'.$$

Die Normalform von A_2' erhält man offenbar, wenn man das Doppelte der zweiten Zeile von der letzten subtrahiert, d.h. von links mit der Matrix

$$(B_3')^{-1} := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}$$

multipliziert. Nimmt man alles zusammen, füllt die verkürzten Matrizen auf die volle Größe auf, so haben wir also: Mit

$$B^{-1} := \left(\begin{array}{c|c} 1 & \\ \hline & (B_3')^{-1} \end{array}\right) \left(\begin{array}{c|c} 1 & \\ \hline & (B_2')^{-1} \end{array}\right) \cdot B_1^{-1}$$

und

$$C := C_1 \cdot \left(\begin{array}{c|c} 1 & \\ \hline & C_2' \end{array}\right)$$

ist

$$B^{-1}AC = \operatorname{diag}(1, -7, \alpha, \frac{1}{7}\alpha\beta) = \begin{pmatrix} 1 & & & \\ & -7 & & \\ & & 2-T & \\ & & & -\frac{1}{7}(2-T)(T^2+T+1). \end{pmatrix}$$

Dabei ist

$$B^{-1} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 2 & 0 & T \\ 0 & 0 & 1 & 0 \\ \frac{1}{7}\beta+1 & \frac{2}{7}\beta+1 & -2 & T(\frac{1}{7}\beta+1) \end{pmatrix}$$

und

$$C = \begin{pmatrix} 1 & 2T+4 & 4 & \frac{1}{7}\beta + (T+1)\frac{1}{7}\alpha \\ 0 & -(T+3) & -2 & -\frac{1}{7}\beta \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & \frac{1}{7}\alpha \end{pmatrix}.$$

Hieraus bekommt man die Matrix B zu

$$B = \begin{pmatrix} -T & -\frac{2}{7}\beta - 1 & 4 & 2 \\ 0 & \frac{1}{7}\beta + 1 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Ihre Spalten seien $(b_1, \ldots, b_4)$.

Schritt (vi): Mit dem in Schritt (i) und (ii) notierten Homomorphismus φ haben wir zu berechnen $v_i := \varphi(b_i)$ und diese v_i erzeugen zyklische Moduln deren Ordnung gerade die δ_i, d.h. die Diagonalelemente der berechneten Normalform von $A = F - T \cdot I$ sind, also

$$\delta_1 := 1, \delta_2 := -7, \delta_3 := \alpha = 2 - T, \delta_4 := \frac{1}{7}\alpha\beta = -\frac{1}{7}(2-T)(T^2+T+1).$$

Es ist

$$\begin{aligned} \varphi(b_1) &= \varphi((-T)e_1 + e_4) = (-f)(e_1) + e_4 = 0 \\ \varphi(b_2) &= \varphi((-\frac{2}{7}\beta - 1)e_1 + (\frac{1}{7}\beta + 1)e_2) \\ &= \frac{2}{7}(f^2+f+1)(e_1) - e_1 - \frac{1}{7}(f^2+f+1)(e_2) + e_2 = 0. \end{aligned}$$

Daß hier jedesmal 0 herauskommt folgt natürlich theoretisch daraus, daß δ_1, δ_2 Einheiten sind, somit $\varphi(b_1), \varphi(b_2)$ als Ordnungsideal den ganzen Ring haben und damit $= 0$ sein müssen.

Es ist

$$v_3 := \varphi(b_3) = \varphi(\begin{pmatrix} 4 \\ -2 \\ 1 \\ 0 \end{pmatrix}) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

und damit (siehe BZ5) Eigenvektor von f zum Eigenwert 2. Hierzu gehört der recht simple zyklische Modul

$$V_3 := \mathbb{R}\text{-span}(v_3).$$

Ferner ist

$$v_4 := \varphi(b_4) = \varphi(\begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix}) = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

Hierzu gehört der zyklische Modul

$$V_4 := \mathbb{R}\text{-span}(v_4, f(v_4), \ldots) = \mathbb{R}\text{-span}(\begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -4 \\ 0 \\ -1 \end{pmatrix}).$$

Zum Modul V_3 gehört als Minimalpolynom die Ordnung von v_3, also das Polynom $T-2$ und zu dem Modul V_4 als Minimalpolynom die Ordnung von v_4, also das Polynom $(T-2)(T^2+T+1)$. Da dies reduzibel ist, kann man V_4 noch nach Satz Z.30 weiter zerlegen. Das auszuführen sei Ihnen überlassen.