

6. Normalformentheorie

Vorbemerkungen

In diesem Kapitel geht es darum, für endlich-dimensionale K -Vektorräume V die Struktur der Endomorphismen von V zu klären. Was aber hat man unter der *Struktur* eines Endomorphismus $f: V \rightarrow V$ zu verstehen? Man kann beispielsweise die folgenden Fragen stellen:

1. Gibt es nicht-triviale Untervektorräume $U \subset V$ mit $f(U) \subset U$, auf denen $f|_U$ von besonders einfacher Gestalt ist, z. B. $f|_U = \lambda \operatorname{id}_U$ mit einem Skalar $\lambda \in K$?
2. Um f auf ganz V zu beschreiben: Kann man V in eine direkte Summe nicht-trivialer Untervektorräume $V = \bigoplus_{i=1}^r U_i$ zerlegen mit $f(U_i) \subset U_i$, so dass sich $f|_{U_i}$ in signifikanter Weise charakterisieren lässt? Gibt es eine feinste Zerlegung dieses Typs, und ist diese in irgendeiner Weise eindeutig charakterisiert?

Dies sind bereits die wichtigsten Fragen, die wir untersuchen wollen. Zunächst ist die Beantwortung der ersten Frage relativ einfach. Wir betrachten für $\lambda \in K$ den Endomorphismus $f - \lambda \operatorname{id}$ von V . Sein Kern gibt genau denjenigen (maximalen) Untervektorraum von V an, auf dem sich f wie $\lambda \operatorname{id}$ verhält, und dieser Unterraum ist genau dann nicht-trivial, wenn der Kern von $f - \lambda \operatorname{id}$ nicht-trivial ist, also gemäß 2.1/11 genau dann, wenn $f - \lambda \operatorname{id}$ nicht invertierbar ist, und damit nach 4.3/3 genau dann, wenn $\det(f - \lambda \operatorname{id}) = 0$ gilt. Es ist also die Gleichung $\det(f - \lambda \operatorname{id}) = 0$ für $\lambda \in K$ zu lösen, und wir werden damit automatisch dazu veranlasst, das so genannte *charakteristische Polynom* $\chi_f \in K[T]$ zu f zu betrachten, das entsteht, wenn wir auf $\det(\lambda \operatorname{id} - f)$ die Definition der Determinante einer beschreibenden Matrix anwenden, dabei jedoch anstelle von λ die Variable T vorsehen.

Die Nullstellen von χ_f in K werden als *Eigenwerte* zu f bezeichnet. Für einen solchen Eigenwert λ heißt $V_\lambda = \ker(\lambda \operatorname{id} - f)$ der zu λ gehörige *Eigenraum*, und die Elemente von $V_\lambda - \{0\}$ werden als *Eigenvektoren* zum Eigenwert λ bezeichnet. Wir werden zeigen, dass Eigenvektoren zu verschiedenen Eigenwerten stets linear unabhängig sind, und daraus folgern, dass die Summe der Eigenräume zu den verschiedenen Eigenwerten von f stets direkt ist. Stimmt diese Summe bereits mit V überein, so ist f *diagonalisierbar*, womit wir meinen,

dass V eine Basis bestehend aus Eigenvektoren zu f besitzt und was zur Folge hat, dass die zugehörige beschreibende Matrix von f eine Diagonalmatrix ist. Diese Situation ist beispielsweise gegeben, wenn die Anzahl der verschiedenen Eigenwerte von f gleich der Dimension von V ist. Wir erhalten damit auch eine erste (partielle) Antwort auf die eingangs gestellte Frage 2.

Es ist relativ leicht einzusehen, dass lineare Abbildungen $f: V \rightarrow V$ im Allgemeinen nicht diagonalisierbar sind. Beispielsweise ist eine Drehung um 0 im $\mathbb{R}^2$ nicht diagonalisierbar, es sei denn, der Drehwinkel beträgt 0° oder 180° . So wird man zur Beantwortung der Frage 2 noch nach anderen Möglichkeiten suchen müssen, um Untervektorräume $U \subset V$ mit $f(U) \subset U$, d. h. so genannte *f-invariante* Untervektorräume, zu konstruieren. Folgende Beobachtung ist hierbei grundlegend: Man betrachte zu einem Vektor $x \in V$ den Untervektorraum $U \subset V$, der von den Elementen $x, f(x), f^2(x), \dots$ erzeugt wird. Dann ist U offenbar ein *f-invarianter* Untervektorraum in V , offenbar der kleinste, der x enthält. Wir nennen U den von x erzeugten *f-zyklischen* Untervektorraum von V . Seine Struktur lässt sich leicht beschreiben, wenn man Ergebnisse aus Kapitel 5 über den Polynomring $K[T]$ verwendet. Man kann nämlich die K -lineare Abbildung

$$\varphi: K[T] \rightarrow U, \quad \sum_{i \in \mathbb{N}} c_i T^i \mapsto \sum_{i \in \mathbb{N}} c_i f^i(x),$$

betrachten und stellt dabei fest, dass $\ker \varphi$ ein Ideal in $K[T]$ ist. Denn für

$$p = \sum_{i=0}^r c_i T^i \in K[T], \quad q = \sum_{j=0}^s d_j T^j \in \ker \varphi$$

gilt

$$\begin{aligned} \varphi(pq) &= \varphi\left(\sum_{k=0}^{r+s} \left(\sum_{i+j=k} c_i d_j\right) T^k\right) = \sum_{k=0}^{r+s} \left(\sum_{i+j=k} c_i d_j\right) f^k(x) \\ &= \left(\sum_{i=0}^r c_i f^i\right)\left(\sum_{j=0}^s d_j f^j\right)(x) = \left(\sum_{i=0}^r c_i f^i\right)(0) = 0, \end{aligned}$$

d. h. $\ker \varphi$ ist insbesondere abgeschlossen unter Multiplikation mit Elementen aus $K[T]$. Nun wissen wir aber, dass $K[T]$ ein Hauptidealring ist, dass es folglich ein eindeutig bestimmtes normiertes Polynom kleinsten Grades

$$p(T) = T^r + c_1 T^{r-1} + \dots + c_r \in \ker \varphi$$

gibt und dass dieses das Ideal erzeugt; denn $\ker \varphi$ besteht nicht nur aus dem Nullpolynom, da $K[T]$ als K -Vektorraum von unendlicher, V aber von endlicher Dimension ist. Hieraus gewinnt man die Gleichung

$$p(f)(x) = f^r(x) + c_1 f^{r-1}(x) + \dots + c_r x = 0,$$

und diese zeigt in induktiver Weise, dass U bereits von $x, f^1(x), \dots, f^{r-1}(x)$ erzeugt wird und, da es in $\ker \varphi$ kein nicht-triviales Polynom vom Grade $< r$

gibt, dass diese Elemente sogar eine Basis von U bilden. Bezüglich dieser Basis wird $f|_U$ dann durch die so genannte *Begleitmatrix*

$$\begin{pmatrix} 0 & & & & -c_0 \\ 1 & 0 & & & -c_1 \\ & 1 & . & & -c_2 \\ & & . & . & \dots \\ & & & . & . & \dots \\ & & & & 0 & -c_{n-2} \\ & & & & 1 & -c_{n-1} \end{pmatrix}$$

des Polynoms p beschrieben.

Um die eingangs gestellte Frage 2 vollständig zu klären, ist noch zu untersuchen, in wie weit sich V als eine direkte Summe f -zyklischer Untervektorräume darstellen lässt. Dass dies in der Tat stets möglich ist, und zwar mit so genannten f -*unzerlegbaren* (ebenfalls f -zyklischen) Untervektorräumen, die keine weitere Zerlegung in eine direkte Summe f -invarianter Unterräume mehr zulassen, und mit zugehörigen normierten Polynomen aus $K[T]$, die eindeutig durch f bestimmt sind, ist ein tief liegendes Resultat, dessen Beweis einigen Aufwand erfordert. Um die eigentlichen Gründe für das Zustandekommen dieses Resultats aufzudecken, werden wir die so genannte *Elementarteilertheorie* behandeln, und zwar für *Moduln* über Hauptidealringen. Ein Modul über einem Ring ist formal genauso definiert wie ein Vektorraum über einem Körper, nur dass man als Skalarenbereich anstelle eines Körpers einen Ring vorsieht. Dass man beispielsweise einen K -Vektorraum mit einem Endomorphismus $f: V \rightarrow V$ als einen Modul über dem Polynomring $K[T]$ auffassen sollte, wobei man für $x \in V$ das Produkt Tx durch $f(x)$ definiert, wird durch die obige Betrachtung f -zyklischer Untervektorräume nahe gelegt. In diesem Sinne ist der von einem Vektor $x \in V$ erzeugte f -zyklische Untervektorraum $U \subset V$ zu sehen als der von x erzeugte $K[T]$ -Untermodul von V .

Obwohl Moduln als "Vektorräume" über Ringen interpretiert werden können, gibt es dennoch gravierende Unterschiede zu Vektorräumen über Körpern, die durch das Phänomen der so genannten *Torsion* verursacht sind. Für einen Modul M über einem Ring R gibt es nämlich im Allgemeinen von Null verschiedene Elemente $r \in R$ und $m \in M$ mit $rm = 0$, wobei dann r natürlich keine Einheit sein kann, da ansonsten $m = r^{-1}(rm) = 0$ folgen würde. Insbesondere kann ein solcher Modul keine Basis besitzen. Gibt es zu jedem $m \neq 0$ aus M ein $r \neq 0$ in R mit $rm = 0$, so bezeichnet man M als einen *Torsionsmodul*. Beispielsweise ist in der obigen Situation auch V als $K[T]$ -Modul ein Torsionsmodul. Da V von endlicher Dimension ist, existieren nämlich nicht-triviale Polynome $p \in K[T]$ mit $p \cdot V = 0$. Wir werden insbesondere sehen, dass es wiederum ein eindeutig bestimmtes normiertes Polynom kleinsten Grades $p_f \in K[T]$ mit $p_f \cdot V = 0$ gibt. Man nennt p_f das *Minimalpolynom* zu f , und wir werden mit dem Satz von Cayley-Hamilton zeigen, dass p_f stets ein Teiler des charakteristischen Polynoms χ_f ist und damit einen Grad $\leq n$ besitzt.

Die mittels der Elementarteilertheorie gewonnene Zerlegung von V in f -zyklische bzw. f -invariante Unterräume werden wir schließlich dazu verwenden, um f mittels kanonisch zugeordneter Matrizen zu beschreiben. Ist beispielsweise $V = \bigoplus_{i=1}^r U_i$ eine Zerlegung in f -zyklische Untervektorräume, wobei $U_i \simeq K[T]/(p_i)$ mit normierten Polynomen $p_i \in K[T]$ im Sinne von $K[T]$ -Moduln gelte, so kann man für jedes U_i eine K -Basis wählen, derart dass $f|_{U_i}$ bezüglich dieser Basis durch die Begleitmatrix $A(p_i)$ zu p_i dargestellt wird. Setzt man die einzelnen Basen der U_i zu einer Gesamtbasis von V zusammen, so ist die zugehörige f beschreibende Matrix von der Form $\text{Diag}(A(p_1), \dots, A(p_r))$, d. h. eine Art "Diagonalmatrix", auf deren Diagonalen die Kästchen $A(p_i)$ angeordnet sind. Geht man von irgendeiner Matrix $A \in K^{n \times n}$ aus, die f beschreibt, und sind die Polynome p_i Primpotenzen, so ist $A' = \text{Diag}(A(p_1), \dots, A(p_r))$ bereits die so genannte *allgemeine Normalform* von A . Im Übrigen werden wir auch die *Jordansche Normalform* von A betrachten (sofern das charakteristische Polynom χ_f in lineare Faktoren zerfällt) und abschließend zeigen, wie man Normalformen explizit berechnen kann, indem man die Elementarteiler der Matrix $TE - A \in K[T]^{n \times n}$ bestimmt; dabei sei $E \in K^{n \times n}$ die Einheitsmatrix.

6.1 Eigenwerte und Eigenvektoren

Wir kehren nunmehr zur Theorie der Vektorräume über einem Körper K zurück und betrachten zunächst eine K -lineare Abbildung $f: V \longrightarrow W$ zwischen endlich-dimensionalen K -Vektorräumen V und W . Ist dann $X = (x_1, \dots, x_n)$ eine Basis von V und $Y = (y_1, \dots, y_m)$ eine Basis von W , so lässt sich f durch eine zugehörige Matrix $A_{f,X,Y}$ beschreiben; vgl. 3.1/3. Durch geschickte Wahl von X und Y kann man erreichen, dass $A_{f,X,Y}$ von möglichst einfacher Gestalt ist. So hatten wir in 3.4/7 gesehen, dass es Basen X' von V und Y' von W gibt mit

$$A_{f,X',Y'} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix};$$

dabei bezeichnet $E_r \in K^{r \times r}$ die Einheitsmatrix und r den Rang von f . Weiter besteht die Relation

$$A_{f,X',Y'} = (A_{\text{id},Y',Y})^{-1} \cdot A_{f,X,Y} \cdot A_{\text{id},X',X}$$

mit den Basiswechsellmatrizen $A_{\text{id},X',X}$ und $A_{\text{id},Y',Y}$; vgl. 3.4/4. Unter Benutzung der bijektiven Korrespondenz zwischen linearen Abbildungen und Matrizen können wir daher auch sagen, dass es zu einer Matrix $A \in K^{m \times n}$ vom Rang r stets invertierbare Matrizen $S \in \text{Gl}(m, K)$ und $T \in \text{Gl}(n, K)$ mit

$$S^{-1} \cdot A \cdot T = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

gibt; vgl. hierzu auch 3.4/8.

Wir wollen im Weiteren ein entsprechendes Problem für Endomorphismen $f: V \longrightarrow V$ studieren. Genauer soll durch geeignete Wahl einer Basis X von

V erreicht werden, dass die Matrix $A_{f,X,X}$ von möglichst einfacher Gestalt ist. Übersetzt in die Sprache der Matrizen bedeutet dies: Ausgehend von einer Matrix $A \in K^{n \times n}$ ist eine invertierbare Matrix $S \in \text{Gl}(n, K)$ zu konstruieren, derart dass die Matrix $S^{-1} \cdot A \cdot S$ von möglichst einfacher Gestalt ist, beispielsweise eine Diagonalmatrix ist:

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \\ & & & \lambda_n \end{pmatrix}$$

Dabei sei erwähnt, dass eine solche Diagonalgestalt allerdings nicht in allen Fällen zu erreichen ist. Um eine bequeme Sprechweise für unser Problem zu haben, sagen wir:

Definition 1. Zwei Matrizen $A, B \in K^{n \times n}$ heißen *ähnlich*, wenn es eine invertierbare Matrix $S \in \text{Gl}(n, K)$ mit $B = S^{-1} \cdot A \cdot S$ gibt.

Man sieht unmittelbar, dass die Ähnlichkeit von Matrizen aus $K^{n \times n}$ eine Äquivalenzrelation darstellt. Somit zerfällt $K^{n \times n}$ in disjunkte Klassen ähnlicher Matrizen.

Bemerkung 2. Zu einer Matrix $A \in K^{n \times n}$ betrachte man einen n -dimensionalen K -Vektorraum V mit einer Basis X und den (eindeutig bestimmten) Endomorphismus $f: V \rightarrow V$ mit $A_{f,X,X} = A$ vgl. 3.3/2. Für eine weitere Matrix $B \in K^{n \times n}$ ist dann äquivalent:

- (i) A und B sind ähnlich.
- (ii) Es existiert eine Basis X' von V mit $A_{f,X',X'} = B$.

Beweis. Seien zunächst A und B ähnlich, gelte also $B = S^{-1} \cdot A \cdot S$ mit $S \in \text{Gl}(n, K)$. Fassen wir dann die Matrix S gemäß 3.4/1 (und den sich daran anschließenden Erläuterungen) als Basiswechselmatrix auf, so erhalten wir eine Basis X' von V mit $S = A_{\text{id},X',X}$, und es folgt mit 3.4/4

$$A_{f,X',X'} = (A_{\text{id},X',X})^{-1} \cdot A_{f,X,X} \cdot A_{\text{id},X',X} = S^{-1} \cdot A \cdot S = B,$$

d. h. Bedingung (ii) ist erfüllt.

Ist umgekehrt Bedingung (ii) gegeben, so zeigt die Gleichung

$$A_{f,X',X'} = (A_{\text{id},X',X})^{-1} \cdot A_{f,X,X} \cdot A_{\text{id},X',X},$$

dass A und B ähnlich sind. □

Wir wollen uns nun mit der Frage beschäftigen, wann eine gegebene Matrix $A \in K^{n \times n}$ zu einer Diagonalmatrix ähnlich ist. Dazu führen wir folgende Sprechweise ein:

Definition 3. Eine Matrix $A \in K^{n \times n}$ heißt *diagonalisierbar*, wenn sie zu einer Diagonalmatrix ähnlich ist.

Ein Endomorphismus $f: V \longrightarrow V$ eines endlich-dimensionalen K -Vektorraums V heißt diagonalisierbar, wenn die beschreibende Matrix $A_{f,X,X}$ für eine Basis X von V (und damit für alle Basen von V) diagonalisierbar ist.

Aus Bemerkung 2 kann man ablesen, dass ein Endomorphismus $f: V \longrightarrow V$ eines endlich-dimensionalen K -Vektorraums V genau dann diagonalisierbar ist, wenn es eine Basis $X = (x_1, \dots, x_n)$ von V mit

$$A_{f,X,X} = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \\ & & & \lambda_n \end{pmatrix}$$

gibt, so dass also $f(x_i) = \lambda_i x_i$ für gewisse Konstanten $\lambda_i \in K$ gilt, $i = 1, \dots, n$. Wir werden in diesem Zusammenhang folgende Terminologie verwenden:

Definition 4. Es sei $f: V \longrightarrow V$ ein Endomorphismus eines K -Vektorraums V . Eine Konstante $\lambda \in K$ heißt Eigenwert zu f , wenn es einen Vektor $a \in V - \{0\}$ mit $f(a) = \lambda a$ gibt. Man nennt in diesem Falle a einen Eigenvektor von f zum Eigenwert λ . Für eine Matrix $A \in K^{n \times n}$ seien Eigenwerte und -Vektoren erklärt als Eigenwerte und -Vektoren der zugehörigen linearen Abbildung $K^n \longrightarrow K^n$, $x \longmapsto Ax$.

Eigenvektoren sind definitionsgemäß immer von 0 verschieden, und wir können formulieren:

Bemerkung 5. Ein Endomorphismus $f: V \longrightarrow V$ eines endlich-dimensionalen K -Vektorraums V ist genau dann diagonalisierbar, wenn es in V eine Basis aus Eigenvektoren bezüglich f gibt.

Als Anwendung der Beschreibung linearer Abbildungen mittels Matrizen, vgl. 3.1/7, ergibt sich:

Bemerkung 6. Es sei $f: V \longrightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V mit Basis X . Für $\lambda \in K$ ist dann äquivalent:

- (i) λ ist Eigenwert von f .
- (ii) λ ist Eigenwert von $A_{f,X,X}$.

Beweis. Sei $\dim_K V = n$. Wir benutzen das kommutative Diagramm

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ \kappa_X \downarrow & & \downarrow \kappa_X \\ K^n & \xrightarrow{\tilde{f}} & K^n \end{array}$$

aus 3.1/8. Dabei ist κ_X derjenige Isomorphismus, der einem Vektor $v \in V$ den zugehörigen Koordinatenspaltenvektor $v_X \in K^n$ zuordnet, sowie $\tilde{f}: K^n \longrightarrow K^n$

die durch $u \mapsto A_{f,X,X} \cdot u$ erklärte Abbildung. Ist nun $a \in V$ ein Eigenvektor zu f mit Eigenwert $\lambda \in K$, so gilt insbesondere $a \neq 0$ und damit auch $\kappa_X(a) \neq 0$. Weiter folgt aufgrund der Kommutativität des obigen Diagramms

$$\tilde{f}(\kappa_X(a)) = \kappa_X(f(a)) = \kappa_X(\lambda a) = \lambda \kappa_X(a),$$

d. h. $\kappa_X(a)$ ist Eigenvektor zu $\tilde{f}$, ebenfalls zum Eigenwert λ . Ist umgekehrt $b \in K^n$ ein Eigenvektor zu $\tilde{f}$ zum Eigenwert λ , so folgt entsprechend, dass $\kappa_X^{-1}(b) \in V$ ein Eigenvektor zu f zum Eigenwert λ ist. $\square$

Insbesondere sieht man mit Bemerkung 2 oder auch mittels einfacher direkter Rechnung:

Bemerkung 7. *Ähnliche Matrizen besitzen dieselben Eigenwerte.*

Als Beispiel zeigen wir, dass die Matrizen

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}, \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \in \mathbb{C}^{2 \times 2}$$

nicht diagonalisierbar sind. Da das Gleichungssystem

$$\alpha_2 = \lambda \alpha_1, \quad -\alpha_1 = \lambda \alpha_2$$

für einen nicht-trivialen Vektor $(\alpha_1, \alpha_2)^t \in \mathbb{R}^2$ stets auf die Gleichung $\lambda^2 = -1$ führt, die in $\mathbb{R}$ nicht lösbar ist, sehen wir, dass A in $\mathbb{R}^{2 \times 2}$ nicht diagonalisierbar sein kann, da die zugehörige lineare Abbildung $\mathbb{R}^2 \rightarrow \mathbb{R}^2$, $x \mapsto Ax$, keinen Eigenwert besitzt. Das Bild ändert sich jedoch, wenn wir A als Matrix in $\mathbb{C}^{2 \times 2}$ auffassen, denn die durch A gegebene $\mathbb{C}$ -lineare Abbildung $\mathbb{C}^2 \rightarrow \mathbb{C}^2$, $x \mapsto Ax$, wird bezüglich der Basis $\begin{pmatrix} 1 \\ i \end{pmatrix}, \begin{pmatrix} 1 \\ -i \end{pmatrix}$ durch eine Diagonalmatrix beschrieben. Weiter zeigt das Gleichungssystem

$$\alpha_1 = \lambda \alpha_1, \quad \alpha_1 + \alpha_2 = \lambda \alpha_2,$$

dass die Matrix B höchstens $\lambda = 1$ als Eigenwert besitzt. Wäre B also diagonalisierbar, so müsste B zur Einheitsmatrix ähnlich sein und dann schon mit dieser übereinstimmen, da die Einheitsmatrix aus trivialen Gründen nur zu sich selbst ähnlich ist.

Satz 8. *Es sei $f: V \rightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V . Sind $a_1, \dots, a_r \in V$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r$, so sind $a_1, \dots, a_r$ linear unabhängig.*

Beweis. Wir schließen mit Induktion nach r , wobei wir $r \geq 1$ annehmen dürfen. Der Fall $r = 1$ ist klar, denn ein Eigenvektor ist nach Definition stets von 0 verschieden. Sei also $r > 1$, und gelte

$$\sum_{i=1}^r \alpha_i a_i = 0$$

mit Koeffizienten $\alpha_1, \dots, \alpha_r \in K$. Man hat dann

$$\sum_{i=1}^r \lambda_i \alpha_i a_i = \sum_{i=1}^r \alpha_i f(a_i) = f\left(\sum_{i=1}^r \alpha_i a_i\right) = f(0) = 0,$$

aber auch

$$\sum_{i=1}^r \lambda_1 \alpha_i a_i = 0$$

und folglich

$$\sum_{i=2}^r (\lambda_i - \lambda_1) \alpha_i a_i = 0.$$

Nun sind $a_2, \dots, a_r$ insgesamt $r - 1$ Eigenvektoren zu paarweise verschiedenen Eigenwerten und somit nach Induktionsvoraussetzung linear unabhängig. Es ergibt sich daher $(\lambda_i - \lambda_1) \alpha_i = 0$ und damit $\alpha_i = 0$ für $i = 2, \dots, r$. Dann zeigt die Gleichung $\sum_{i=1}^r \alpha_i a_i = 0$, dass auch der Term $\alpha_1 a_1$ verschwindet und wegen $a_1 \neq 0$ sogar der Koeffizient α_1 . Die Vektoren $a_1, \dots, a_r$ sind also wie behauptet linear unabhängig. $\square$

Korollar 9. *Ein Endomorphismus $f: V \longrightarrow V$ eines endlich-dimensionalen K -Vektorraums V hat höchstens $n = \dim_K V$ verschiedene Eigenwerte.*

Korollar 10. *Besitzt ein Endomorphismus $f: V \longrightarrow V$ eines endlich-dimensionalen K -Vektorraums V genau $n = \dim_K V$ verschiedene Eigenwerte, so ist f diagonalisierbar.*

Beweis. Seien $\lambda_1, \dots, \lambda_n \in K$ paarweise verschiedene Eigenwerte zu f , und seien $a_1, \dots, a_n \in V$ zugehörige Eigenvektoren. Dann sind diese gemäß Satz 8 linear unabhängig, bilden also wegen $n = \dim_K V$ eine Basis X von V . Die zugehörige Matrix $A_{f,X,X}$ ist eine Diagonalmatrix mit $\lambda_1, \dots, \lambda_n$ als Diagonalelementen. $\square$

Wir wollen Satz 8 noch etwas verallgemeinern.

Definition 11. *Es sei $\lambda \in K$ Eigenwert eines Vektorraumendomorphismus $f: V \longrightarrow V$. Dann heißt*

$$V_\lambda := \ker(f - \lambda \text{id}) = \{a \in V; f(a) = \lambda a\}$$

der Eigenraum von f zum Eigenwert λ .

Korollar 12. *Sei $f: V \longrightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V , und seien $\lambda_1, \dots, \lambda_r$ die sämtlichen Eigenwerte von f . Ist dann $V' = \sum_{i=1}^r V_{\lambda_i}$ der von den zugehörigen Eigenräumen erzeugte Untervektorraum von V , so gilt*

$$V' = \bigoplus_{i=1}^r V_{\lambda_i};$$

die Summe ist also direkt. Im Übrigen ist f genau dann diagonalisierbar, wenn $V = V'$ gilt, wenn also V von den Eigenräumen zu f erzeugt wird.

Beweis. Da Eigenvektoren zu paarweise verschiedenen Eigenwerten linear unabhängig sind, kann eine Summe $\sum_{i=1}^r v_i$ mit $v_i \in V_{\lambda_i}$ nur dann verschwinden, wenn alle v_i verschwinden. Dies bedeutet aber, dass V' die direkte Summe der Eigenräume V_{λ_i} ist. Ist nun f diagonalisierbar, so besitzt V eine Basis aus Eigenvektoren, und es gilt $V = V'$. Umgekehrt, ist V darstellbar als direkte Summe der Eigenräume V_{λ_i} , so wähle man in jedem dieser Eigenräume eine Basis. Das System aller dieser Elemente bildet dann eine Basis von V , die aus lauter Eigenvektoren von f besteht, d. h. f ist diagonalisierbar. $\square$

Korollar 13. Für eine Diagonalmatrix

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \\ & & & \lambda_n \end{pmatrix}$$

sind $\lambda_1, \dots, \lambda_n$ die einzigen Eigenwerte von A .

Beweis. Wir betrachten die lineare Abbildung $f: K^n \rightarrow K^n, x \mapsto Ax$. Es ist klar, dass es sich bei den $\lambda_1, \dots, \lambda_n$ um Eigenwerte von A bzw. f handelt. Um Wiederholungen zu vermeiden, schreibe man $\{\lambda_1, \dots, \lambda_n\} = \{\lambda'_1, \dots, \lambda'_s\}$, wobei die Elemente $\lambda'_1, \dots, \lambda'_s$ paarweise verschieden sind. Die Diagonalgestalt von A besagt, dass es in K^n eine Basis gibt, nämlich die kanonische Basis $e_1, \dots, e_n$, so dass e_i jeweils Eigenvektor von f zum Eigenwert λ_i ist. Für $j = 1, \dots, s$ sei nun $U_{\lambda'_j} \subset V$ derjenige lineare Unterraum, der erzeugt wird von allen e_i mit Indizes i , für die $\lambda_i = \lambda'_j$ gilt. Es besteht dann die Zerlegung

$$V = \bigoplus_{j=1}^s U_{\lambda'_j}, \quad U_{\lambda'_j} \subset V_{\lambda'_j},$$

wobei $V_{\lambda'_j}$ jeweils der Eigenraum von f zum Eigenwert λ'_j ist. Ein Vergleich mit der Zerlegung aus Korollar 12 ergibt $U_{\lambda'_j} = V_{\lambda'_j}$ für $j = 1, \dots, s$ und zeigt außerdem, dass es neben $\lambda'_1, \dots, \lambda'_s$ keine weiteren Eigenwerte von A geben kann. $\square$

Aufgaben

V sei stets ein Vektorraum endlicher Dimension über einem Körper K .

1. Es seien $a, b \in V$ Eigenvektoren eines Endomorphismus $f: V \rightarrow V$. Man untersuche, in welchen Fällen auch $a - b$ ein Eigenvektor von f ist.

2. Es sei $\lambda \in K$ Eigenwert eines Endomorphismus $f: V \rightarrow V$. Man zeige, dass für Polynome $q \in K[T]$ jeweils $q(\lambda)$ Eigenwert von $q(f)$ ist.
3. Für die Matrix

$$A = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathbb{R}^{4 \times 4}$$

berechne man alle Eigenwerte und die zugehörigen Eigenräume. Ist A diagonalisierbar?

4. Die Matrizen $A, B \in K^{n \times n}$ seien ähnlich. Man zeige in direkter Weise:
- (i) Ein Element $\lambda \in K$ ist genau dann ein Eigenwert von A , wenn es Eigenwert von B ist.
 - (ii) Für Eigenwerte $\lambda \in K$ von A bzw. B gilt $\dim_K V_{A,\lambda} = \dim_K V_{B,\lambda}$, wobei $V_{A,\lambda}$ den Eigenraum zu λ bezüglich der linearen Abbildung $K^n \rightarrow K^n$, $x \mapsto Ax$ bezeichne; entsprechend für $V_{B,\lambda}$.
5. Es seien $A, B \in K^{n \times n}$ ähnlich. Dann sind für Polynome $q \in K[T]$ auch die Matrizen $q(A)$ und $q(B)$ ähnlich.
6. Zwei Endomorphismen $f, g: V \rightarrow V$ heißen ähnlich, wenn es einen Automorphismus $h: V \rightarrow V$ mit $g = h^{-1} \circ f \circ h$ gibt. Man zeige: f und g sind genau dann ähnlich, wenn für eine gegebene Basis X von V die beschreibenden Matrizen $A_{f,X,X}$ und $A_{g,X,X}$ ähnlich sind.
7. Für ein kommutatives Diagramm linearer Abbildungen zwischen K -Vektorräumen

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

zeige man:

- (i) Ist h injektiv, so ist jeder Eigenwert von f auch Eigenwert von g .
- (ii) Ist h surjektiv, so ist jeder Eigenwert von g auch Eigenwert von f .

Man konstruiere einfache Beispiele, die zeigen, dass in den vorstehenden Aussagen die Voraussetzungen “injektiv” bzw. “surjektiv” nicht entbehrlich sind.

6.2 Minimalpolynom und charakteristisches Polynom

Wie wir im vorigen Abschnitt gesehen haben, steht das Problem der Diagonalisierbarkeit von Endomorphismen oder Matrizen in engem Zusammenhang mit dem Problem, die zugehörigen Eigenwerte und Eigenvektoren zu bestimmen. Wir beschäftigen uns daher zunächst mit der Berechnung von Eigenwerten. Generell sei V in diesem Abschnitt ein K -Vektorraum *endlicher* Dimension n .

Satz 1. Sei $f: V \longrightarrow V$ ein Endomorphismus. Für $\lambda \in K$ ist dann äquivalent:

- (i) λ ist Eigenwert zu f .
- (ii) $\ker(\lambda \operatorname{id} - f) \neq 0$.
- (iii) $\lambda \operatorname{id} - f$ ist nicht invertierbar.
- (iv) $\det(\lambda \operatorname{id} - f) = 0$.

Beweis. Sei λ ein Eigenwert zu f . Dann existiert ein Eigenvektor zu λ , d. h. ein $a \in V - \{0\}$ mit $f(a) = \lambda a$. Hieraus ergibt sich $a \in \ker(\lambda \operatorname{id} - f)$ und damit insbesondere $\ker(\lambda \operatorname{id} - f) \neq 0$. Umgekehrt ist jeder von Null verschiedene Vektor $a \in \ker(\lambda \operatorname{id} - f)$ ein Eigenvektor zum Eigenwert λ . Bedingungen (i) und (ii) sind also äquivalent.

Weiter ergibt sich die Äquivalenz (ii) $\iff$ (iii) aus 2.1/11 und die Äquivalenz (iii) $\iff$ (iv) aus 4.3/3 (iv). $\square$

Als Beispiel wollen wir alle Eigenwerte der Matrix

$$A = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

bestimmen. Bezeichnet E die Einheitsmatrix in $\mathbb{R}^{2 \times 2}$, so gilt für $\lambda \in \mathbb{R}$

$$\det(\lambda E - A) = \det \begin{pmatrix} \lambda - 1 & -4 \\ -1 & \lambda - 1 \end{pmatrix} = \lambda^2 - 2\lambda - 3.$$

Die Gleichung $\det(\lambda E - A) = 0$ ist daher äquivalent zu $\lambda = 3$ oder $\lambda = -1$. Daher sind 3, -1 die Eigenwerte von A , und man sieht mittels 6.1/10, dass A diagonalisierbar ist. Und zwar ist A ähnlich zu der Matrix

$$\begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}.$$

Man kann nun leicht die zu den Eigenwerten 3, -1 gehörigen Eigenräume bestimmen, indem man die linearen Gleichungssysteme

$$(3E - A)x = 0 \quad \text{bzw.} \quad (-E - A)x = 0$$

löst.

Wir wollen $\det(\lambda E - A)$ für eine Matrix $A = (\alpha_{ij})_{ij} \in K^{n \times n}$ und eine Konstante $\lambda \in K$ genauer auswerten; E sei nunmehr die Einheitsmatrix in $K^{n \times n}$. Wie in 4.2/4 definiert, gilt

$$\begin{aligned} \det(\lambda E - A) &= \det((\lambda \delta_{ij} - \alpha_{ij})_{ij}) \\ &= \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn} \pi \cdot \prod_{i=1}^n (\lambda \delta_{\pi(i), i} - \alpha_{\pi(i), i}), \end{aligned}$$

und wir können folgende Definition treffen:

Definition 2. Sei $A = (\alpha_{ij})_{ij} \in K^{n \times n}$. Dann heit

$$\chi_A = \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn} \pi \cdot \prod_{i=1}^n (T\delta_{\pi(i),i} - \alpha_{\pi(i),i}) \in K[T]$$

das charakteristische Polynom von A .

Insbesondere gilt $\chi_A(\lambda) = \det(\lambda E - A)$ fr alle $\lambda \in K$, und die Nullstellen von χ_A in K sind gerade die Eigenwerte von A .

Satz 3. Das charakteristische Polynom $\chi_A \in K[T]$ zu einer Matrix $A \in K^{n \times n}$ ist normiert vom Grad n . Es gilt

$$\chi_A = \sum_{i=0}^n c_{n-i} T^i, \quad c_i \in K,$$

mit $c_0 = 1$, $-c_1 = \operatorname{Spur} A = \sum_{i=1}^n \alpha_{ii}$ und $(-1)^n c_n = \det A$, wobei die Summe der Diagonalelemente α_{ii} als Spur von A bezeichnet wird.

Beweis. Als Summe n -facher Produkte linearer Polynome in T ist χ_A vom Grad $\leq n$, und es gilt

$$c_n = \chi_A(0) = \det(-A) = (-1)^n \det A,$$

also $(-1)^n c_n = \det A$. Weiter besitzt der zweite Term in der Zerlegung

$$\chi_A = \prod_{i=1}^n (T - \alpha_{ii}) + \sum_{\substack{\pi \in \mathfrak{S}_n \\ \pi \neq \operatorname{id}}} \prod_{i=1}^n (T\delta_{\pi(i),i} - \alpha_{\pi(i),i})$$

einen Grad $\leq n-2$, denn es wird nur ber Permutationen $\pi \neq \operatorname{id}$ summiert. Fr $\pi \neq \operatorname{id}$ gibt es nmlich mindestens zwei verschiedene Indizes $i, j \in \{1, \dots, n\}$ mit $\pi(i) \neq i$, $\pi(j) \neq j$, so dass folglich die Ausdrcke $T\delta_{\pi(i),i}$, $T\delta_{\pi(j),j}$ fr diese Indizes verschwinden. Die Koeffizienten vom Grad n und $n-1$ in χ_A , also c_0 und c_1 , stimmen daher berein mit den Koeffizienten vom Grad n und $n-1$ in $\prod_{i=1}^n (T - \alpha_{ii})$, und es folgt $c_0 = 1$, sowie $c_1 = -\sum_{i=1}^n \alpha_{ii}$, wie behauptet. $\square$

Man kann das charakteristische Polynom $\chi_A \in K[T]$ zu einer Matrix $A \in K^{n \times n}$ auch durch die Gleichung $\chi_A = \det(TE - A)$ erklren, wobei man dann allerdings die Determinante einer Matrix mit Eintrgen aus dem Polynomring $K[T]$ zu bilden hat. Da wir bisher nur Determinanten von Matrizen mit Koeffizienten aus einem Krper betrachtet und auch nur fr diese Situation Rechenregeln fr Determinanten bewiesen haben, greifen wir zu einem Trick. hnlich wie man den Krper $\mathbb{Q}$ der rationalen Zahlen als Krper aller Brche ganzer Zahlen bildet, konstruiert man zu $K[T]$ den so genannten *rationalen Funktionenkrper* $K(T)$ aller Brche von Polynomen aus $K[T]$. Es ist dann $K[T]$ ein Unterring des Krpers $K(T)$, und man kann $TE - A$ als Matrix in $K(T)^{n \times n}$ auffassen. Insbesondere ist $\det(TE - A)$ wohldefiniert, und man darf

zur Berechnung dieser Determinante die bekannten Entwicklungssätze oder andere Rechenregeln für Determinanten anwenden. Erwähnt sei aber auch, dass sich alternativ die Determinantentheorie über beliebigen kommutativen Ringen entwickeln lässt, worauf wir hier aber nicht weiter eingehen wollen.

Satz 4. Sind $A, B \in K^{n \times n}$ ähnlich, so folgt $\chi_A = \chi_B$.

Beweis. Sei $S \in \text{Gl}(n, K)$ mit $B = S^{-1}AS$. Dann gilt aufgrund der Multiplizität der Determinante

$$\begin{aligned}\chi_B &= \det(TE - S^{-1}AS) = \det(S^{-1}(TE - A)S) \\ &= \det S^{-1} \cdot \det(TE - A) \cdot \det S = \det(TE - A) \\ &= \chi_A.\end{aligned}$$

□

Korollar 5. Ähnliche Matrizen besitzen die gleiche Spur.

Da die charakteristischen Polynome ähnlicher Matrizen übereinstimmen, kann man nunmehr, unter Benutzung von 6.1/2, auch das charakteristische Polynom eines Endomorphismus $f: V \rightarrow V$ erklären.

Definition 6. Sei $f: V \rightarrow V$ ein Endomorphismus und X eine Basis eines endlich-dimensionalen K -Vektorraums V . Dann bezeichnet man

$$\chi_f = \chi_{A_{f,X,X}} \in K[T]$$

als das charakteristische Polynom von f und

$$\text{Spur } f = \text{Spur } A_{f,X,X}$$

als die Spur von f . Für den trivialen Fall $V = 0$ gilt $\chi_f = 1$ und $\text{Spur } f = 0$.

Es folgt mit 6.1/2 und Satz 4, dass χ_f und $\text{Spur } f$ unabhängig von der speziellen Wahl von X sind. Weiter können wir mit Satz 1 feststellen:

Satz 7. Ein Element $\lambda \in K$ ist genau dann ein Eigenwert eines Endomorphismus $f: V \rightarrow V$, wenn λ eine Nullstelle des charakteristischen Polynoms χ_f ist.

Als Nächstes wollen wir einsehen, dass die Diagonalisierbarkeit eines Endomorphismus an gewissen Eigenschaften des zugehörigen charakteristischen Polynoms abzulesen ist.

Satz 8. Für einen Endomorphismus $f: V \rightarrow V$ ist äquivalent:

- (i) f ist diagonalisierbar.

(ii) χ_f zerfällt vollständig in Linearfaktoren, etwa

$$\chi_f = \prod_{i=1}^r (T - \lambda_i)^{n_i},$$

und für den Eigenraum V_{λ_i} zum Eigenwert λ_i gilt

$$\dim V_{\lambda_i} = n_i.$$

Beweis. Sei $\dim_K V = n$. Wir beginnen mit der Implikation (i) $\implies$ (ii) und nehmen f als diagonalisierbar an. Dann existiert eine Basis X von V , bestehend aus Eigenvektoren zu f , also mit

$$A_{f,X,X} = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & & \lambda_n \end{pmatrix}.$$

Insbesondere folgt $\chi_f = \prod_{i=1}^n (T - \lambda_i)$. Schreiben wir dies durch Zusammenfassen gleicher Faktoren zu Potenzen in der Form $\chi_f = \prod_{i=1}^r (T - \lambda_i)^{n_i}$, so folgt $\dim_K(V_{\lambda_i}) \geq n_i$. Wegen $V = \bigoplus_{i=1}^r V_{\lambda_i}$, vgl. 6.1/12 und 6.1/13, gilt dann

$$n = \sum_{i=1}^r \dim_K(V_{\lambda_i}) \geq \sum_{i=1}^r n_i = n$$

und damit $\dim_K(V_{\lambda_i}) = n_i$ für alle $i = 1, \dots, r$.

Ist umgekehrt Bedingung (ii) gegeben, so folgt

$$\sum_{i=1}^r \dim_K(V_{\lambda_i}) = \sum_{i=1}^r n_i = n.$$

Nach 6.1/12 gilt $V' = \bigoplus_{i=1}^r V_{\lambda_i}$ für den von den Eigenräumen V_{λ_i} erzeugten Unterraum $V' \subset V$, also $\dim_K V' = n = \dim_K V$ und damit $V' = V$. Dann ist f aber diagonalisierbar, wiederum nach 6.1/12. $\square$

Als Beispiel für die Anwendung von Satz 8 wollen wir einen Endomorphismus $f: V \longrightarrow V$ betrachten, der bezüglich einer geeigneten Basis durch eine Dreiecksmatrix der Form

$$A = \begin{pmatrix} \lambda & & * \\ & \lambda & \\ & & \ddots \\ 0 & & & \lambda \end{pmatrix} \in K^{n \times n}$$

beschrieben wird. Dann gilt $\chi_f = \chi_A = (T - \lambda)^n$, und λ ist der einzige Eigenwert zu f bzw. A . Ist nun A keine Diagonalmatrix, so ist $\lambda \operatorname{id} - f$ nicht die Nullabbildung und folglich der Eigenraum $V_\lambda = \ker(\lambda \operatorname{id} - f)$ echt in V enthalten. Nach

Satz 8 kann f bzw. A daher nicht diagonalisierbar sein. Wir können dies aber auch in direkter Weise sehen. Wenn A diagonalisierbar ist, so ist A ähnlich zu λE , wobei $E \in K^{n \times n}$ die Einheitsmatrix bezeichne. Da aber E und damit auch λE mit allen Matrizen in $K^{n \times n}$ vertauschbar ist, kann λE nur zu sich selbst ähnlich sein. Somit müsste schon $A = \lambda E$ gelten.

Neben dem charakteristischen Polynom χ_f zu einem Endomorphismus f eines K -Vektorraums V kann man auch noch ein weiteres Polynom zu f betrachten, nämlich das so genannte *Minimalpolynom*. Um dieses zu definieren, betrachten wir den Endomorphismenring $\text{End}_K(V)$ als K -Algebra unter dem Ringhomomorphismus

$$K \longrightarrow \text{End}_K(V), \quad c \longmapsto c \cdot \text{id}_V,$$

und verwenden folgendes Resultat:

Satz 9. *Zu einem Endomorphismus $f: V \longrightarrow V$ eines endlich-dimensionalen K -Vektorraums V bilde man den K -Algebrahomomorphismus*

$$\varphi_f: K[T] \longrightarrow \text{End}_K(V), \quad p \longmapsto p(f),$$

der f anstelle von T einsetzt; vgl. 5.1/7. Dann ist $\ker \varphi_f$ von Null verschieden, und es existiert ein normiertes Polynom $p_f \in K[T]$ mit $\ker \varphi_f = (p_f)$. Es ist p_f das eindeutig bestimmte normierte Polynom kleinsten Grades in $K[T]$ mit $p_f(f) = 0$.

Beweis. Indem wir die Ringmultiplikation vergessen, können wir φ_f auch als Homomorphismus zwischen K -Vektorräumen auffassen. Man hat dann $\dim_K(K[T]) = \infty$ sowie gemäß 3.3/2

$$\dim_K(\text{End}_K(V)) = n^2 < \infty$$

für $n = \dim_K(V)$. Letzteres hat $\ker \varphi_f \neq 0$ zur Folge, bzw. dass φ_f nicht injektiv sein kann. Genauer, die Elemente $\varphi_f(T^0), \dots, \varphi_f(T^{n^2})$ sind aus Dimensionsgründen linear abhängig in $\text{End}_K(V)$, und $\ker \varphi_f$ enthält daher ein nicht-triviales Polynom vom Grad $\leq n^2$. Nun ist aber $\ker \varphi_f$ ein Ideal in $K[T]$ und $K[T]$ ein Hauptidealring; vgl. 5.2/7. Es existiert daher ein nicht-triviales Polynom $p_f \in K[T]$ mit $\ker \varphi_f = p_f \cdot K[T]$. Als erzeugendes Element eines Hauptideals in einem Integritätsring ist p_f nach 5.2/9 eindeutig bestimmt bis auf eine Einheit. Da aber $K[T]^* = K^*$ gilt, ist p_f eindeutig, wenn wir dieses Polynom als normiert voraussetzen. Natürlich ist p_f dann das normierte Polynom kleinsten Grades, welches f annulliert. $\square$

Definition 10. *Ist $f: V \longrightarrow V$ ein Endomorphismus, so heißt das nach Satz 9 eindeutig in $K[T]$ existierende normierte Polynom kleinsten Grades, welches f annulliert, das Minimalpolynom von f ; dieses wird mit p_f bezeichnet. Entsprechend ist das Minimalpolynom p_A einer Matrix $A \in K^{n \times n}$ erklärt als das normierte Polynom kleinsten Grades in $K[T]$, welches A annulliert.*

Im Beweis zu Satz 9 wurde gezeigt, dass der Kern des Homomorphismus $\varphi_f: K[T] \rightarrow \text{End}_K(V)$ nicht-triviale Polynome vom Grad $\leq n^2$ enthält. Als Konsequenz ergibt sich $\text{grad } p_f \leq n^2$. Diese Abschätzung lässt sich aber noch erheblich verbessern.

Satz 11 (Cayley-Hamilton). *Das Minimalpolynom p_f eines Endomorphismus $f: V \rightarrow V$ ist stets ein Teiler des charakteristischen Polynoms χ_f . Insbesondere gilt $\chi_f(f) = 0$ und $\text{grad } p_f \leq \text{grad } \chi_f = n$.*

Beweis. In der Situation von Satz 9 ist nur $\chi_f \in \ker \varphi_f = (p_f)$, d. h. $\chi_f(f) = 0$ zu zeigen. Indem wir dies in ein Matrizenproblem übersetzen, genügt es $\chi_A(A) = 0$ für Matrizen $A \in K^{n \times n}$ zu zeigen. Um bequem rechnen zu können, betrachten wir wieder den rationalen Funktionenkörper $K(T)$, dessen Elemente Brüche von Polynomen aus $K[T]$ sind. Sodann können wir den Unterring $K[T]^{n \times n}$ des Matrizenrings $K(T)^{n \times n}$ betrachten, der aus allen $(n \times n)$ -Matrizen mit Einträgen aus $K[T]$ besteht. Der Homomorphismus $K(T) \rightarrow K(T)^{n \times n}$, der ein Element $q \in K(T)$ auf das q -fache der Einheitsmatrix $E \in K(T)^{n \times n}$ abbildet, beschränkt sich zu einem Ringhomomorphismus

$$K[T] \rightarrow K[T]^{n \times n}, \quad f \mapsto f \cdot E,$$

und definiert auf $K[T]^{n \times n}$ die Struktur einer $K[T]$ -Algebra. Indem wir die Variable $T \in K[T]$ mit ihrem Bild $T \cdot E \in K[T]^{n \times n}$ identifizieren, lässt sich jedes Element $M \in K[T]^{n \times n}$ in der Form $M = \sum_{i \in \mathbb{N}} M_i T^i$ schreiben, wobei die Koeffizienten $M_i \in K^{n \times n}$ eindeutig durch M bestimmt sind und natürlich für fast alle $i \in \mathbb{N}$ verschwinden. Für $M = (\sum_{i \in \mathbb{N}} m_{\mu\nu i} T^i)_{\mu, \nu=1, \dots, n}$ mit Koeffizienten $m_{\mu\nu i} \in K$ setze man nämlich $M_i = (m_{\mu\nu i})_{\mu, \nu=1, \dots, n}$. In dieser Weise können wir $K[T]^{n \times n}$ als “Polynomring” $K^{n \times n}[T]$ auffassen, wobei allerdings zu berücksichtigen ist, dass der Ring $K^{n \times n}$ für $n > 1$ nicht kommutativ ist. Man kann aber wie in Abschnitt 5.1 den Polynomring $R[T]$ auch für einen nicht-kommutativen Ring R erklären, muss sich dann aber darüber im Klaren sein, dass Einsetzungsabbildungen des Typs

$$R[T] \rightarrow R, \quad \sum a_i T^i \mapsto \sum a_i t^i,$$

für Elemente $t \in R$ im Allgemeinen *keine* Homomorphismen mehr darstellen. Denn T ist in $R[T]$ mit allen Elementen aus R vertauschbar, möglicherweise aber nicht $t \in R$ mit allen anderen Elementen aus R .

Nach diesen Vorbereitungen betrachte man nun eine Matrix $A \in K^{n \times n}$ und ihr charakteristisches Polynom $\chi_A \in K[T]$, welches wir für die Zwecke dieses Beweises einmal in der Form $\chi_A = \sum_{i \in \mathbb{N}} c_i T^i$ schreiben wollen. Weiter fassen wir $TE - A$ als Matrix in $K[T]^{n \times n} \subset K(T)^{n \times n}$ auf und bilden deren adjungierte Matrix $(TE - A)^{\text{ad}} \in K(T)^{n \times n}$; vgl. 4.4/2. Aufgrund der Cramerschen Regel 4.4/3 besteht dann die Gleichung

$$(TE - A)^{\text{ad}} \cdot (TE - A) = (\det(TE - A)) \cdot E = \chi_A(T) \cdot E.$$

Dies ist zunächst eine Gleichung in $K(T)^{n \times n}$, sie gilt aber auch in $K[T]^{n \times n}$, da aufgrund der Definition der adjungierten Matrix mit $TE - A$ auch $(TE - A)^{\text{ad}}$ zu $K[T]^{n \times n}$ gehört. Gilt etwa $(TE - A)^{\text{ad}} = \sum_{i \in \mathbb{N}} A_i T^i$ mit Matrizen $A_i \in K^{n \times n}$, so folgt

$$\begin{aligned} \left(\sum_{i \in \mathbb{N}} A_i T^i \right) \cdot (TE - A) &= -A_0 T^0 A + \sum_{i \in \mathbb{N}} (A_i T^i TE - A_{i+1} T^{i+1} A) \\ &= -A_0 A + \sum_{i \in \mathbb{N}} (A_i - A_{i+1} A) T^{i+1} \\ &= \sum_{i \in \mathbb{N}} c_i T^i \cdot E = \chi_A(T) \cdot E \end{aligned}$$

mit

$$-A_0 A = c_0 E, \quad A_i - A_{i+1} A = c_{i+1} E \quad \text{für } i \in \mathbb{N}.$$

Dieselbe Rechnung lässt sich auch mit einer Matrix $B \in K^{n \times n}$ anstelle der Variablen T durchführen, sofern A mit B vertauschbar ist. Insbesondere dürfen wir in vorstehender Gleichung T durch $B := A$ ersetzen und erhalten dann wie gewünscht $\chi_A(A) = 0$, da der Term $(TE - A)$ nach Ersetzen von T durch A verschwindet. $\square$

Wir wollen noch zwei einfache Beispiele betrachten. Für die Einheitsmatrix $E \in K^{n \times n}$, $n > 0$, gilt $\chi_E = (T - 1)^n$, $p_A = T - 1$, und für die Nullmatrix $0 \in K^{n \times n}$ hat man $\chi_0 = T^n$, $p_0 = T$. Insbesondere sieht man, dass das Minimalpolynom im Allgemeinen nicht mit dem charakteristischen Polynom übereinstimmt.

Aufgaben

V sei stets ein Vektorraum endlicher Dimension n über einem Körper K .

1. Man bestimme Eigenwerte und zugehörige Eigenräume der folgenden Matrix:

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ -2 & 2 & 0 & 2 \\ 1 & 0 & 2 & 0 \\ 2 & -1 & 0 & -1 \end{pmatrix}$$

Ist A diagonalisierbar?

2. Man bestimme das Minimalpolynom p_f zu einem Endomorphismus $f: V \rightarrow V$ in folgenden Fällen:
 - (i) $V = 0$
 - (ii) $f = \text{id}$
 - (iii) $f = 0$
 - (iv) Es existieren lineare Unterräume $V_1, V_2 \subset V$ mit $V = V_1 \oplus V_2$, und es gilt $f(v_1 + v_2) = v_1$ für $v_i \in V_i$, $i = 1, 2$.

3. Es seien $U_1, U_2 \subset V$ lineare Unterräume und $f: V \rightarrow V$ ein Endomorphismus, der sich zu Endomorphismen $f_i: U_i \rightarrow U_i$, $i = 1, 2$, einschränkt. Man zeige:
 - (i) Gilt $V = U_1 + U_2$, so folgt $p_f = \text{kgV}(p_{f_1}, p_{f_2})$ für die Minimalpolynome von f, f_1, f_2 .
 - (ii) f schränkt sich zu einem Endomorphismus $f_{12}: U_1 \cap U_2 \rightarrow U_1 \cap U_2$ ein, und es gilt $p_{f_{12}} \mid \text{ggT}(p_{f_1}, p_{f_2})$ für die Minimalpolynome von f_{12}, f_1, f_2 . Gilt im Allgemeinen auch die Gleichheit $p_{f_{12}} = \text{ggT}(p_{f_1}, p_{f_2})$?
4. K sei algebraisch abgeschlossen. Man zeige, ein Endomorphismus $f: V \rightarrow V$ ist genau dann nilpotent (d. h. erfüllt eine Gleichung der Form $f^r = 0$), wenn f außer 0 keine weiteren Eigenwerte besitzt.
5. Es sei $f: V \rightarrow V$ ein Automorphismus. Man zeige, es existiert ein Polynom $q \in K[T]$ mit $f^{-1} = q(f)$.
6. Die Folge der *Fibonacci-Zahlen* $c_1, c_2, \dots \in \mathbb{N}$ ist definiert durch $c_1 = c_2 = 1$ und $c_{n+2} = c_{n+1} + c_n$ für $n \in \mathbb{N}$. Man gebe für c_n einen geschlossenen Ausdruck an, der nur von n abhängt. Hinweis: Man bestimme eine Matrix $A \in \mathbb{R}^{2 \times 2}$ mit $A \cdot (c_{n+1}, c_n)^t = (c_{n+2}, c_{n+1})^t$ für $n \geq 1$ und eine Basiswechselmatrix $S \in \text{Gl}(2, \mathbb{R})$, derart dass $S^{-1} \cdot A \cdot S$ Diagonalgestalt besitzt.

6.3 Der Elementarteilersatz

Es sei V ein Vektorraum über einem Körper K und $f: V \rightarrow V$ ein Endomorphismus. Dann setzt sich die auf V definierte skalare Multiplikation $K \times V \rightarrow V$, $(\alpha, v) \mapsto \alpha \cdot v$, fort zu einer äußeren Multiplikation

$$K[T] \times V \rightarrow V, \quad (p, v) \mapsto p \cdot v := p(f)(v).$$

Dabei ist für $p \in K[T]$ wie üblich $p(f)$ derjenige Ausdruck in $\text{End}_K(V)$, der aus p entsteht, indem man die Variable T durch f ersetzt. Weiter ist $p(f)(v)$ das Bild von v unter dem Endomorphismus $p(f): V \rightarrow V$. Man prüft leicht nach, dass V als additive abelsche Gruppe zusammen mit der äußeren Multiplikation $K[T] \times V \rightarrow V$ den in 1.4/1 aufgeführten Vektorraumaxiomen genügt, wenn man einmal davon absieht, dass $K[T]$ nur ein Ring und kein Körper ist; wir sagen, V sei ein $K[T]$ -Modul. (Man beachte: Im Unterschied zu anderem sprachlichen Gebrauch heißt es in der Mathematik “der Modul” bzw. “die Moduln”, mit Betonung auf der ersten Silbe.)

Will man Normalformen von Endomorphismen $f: V \rightarrow V$ studieren, so bedeutet dies, dass man die oben erklärte Struktur von V als $K[T]$ -Modul analysieren muss. Wir wollen daher zunächst ein paar Grundlagen über Moduln zusammenstellen.

Definition 1. Es sei R ein (kommutativer) Ring. Ein R -Modul ist eine Menge M mit einer inneren Verknüpfung $M \times M \rightarrow M$, $(a, b) \mapsto a + b$, genannt *Addition*, und einer äußeren Verknüpfung $R \times M \rightarrow M$, $(\alpha, a) \mapsto \alpha \cdot a$, genannt *skalare Multiplikation*, so dass gilt:

- (i) M ist eine abelsche Gruppe bezüglich der Addition “+”.
- (ii) $(\alpha + \beta) \cdot a = \alpha \cdot a + \beta \cdot a$, $\alpha \cdot (a + b) = \alpha \cdot a + \alpha \cdot b$ für alle $\alpha, \beta \in R$, $a, b \in M$, d. h. Addition und skalare Multiplikation verhalten sich distributiv.
- (iii) $(\alpha \cdot \beta) \cdot a = \alpha \cdot (\beta \cdot a)$ für alle $\alpha, \beta \in R$, $a \in M$, d. h. die skalare Multiplikation ist assoziativ.
- (iv) $1 \cdot a = a$ für das Einselement $1 \in R$ und alle $a \in M$.

Wir wollen einige Beispiele betrachten:

- (1) Ein Vektorraum über einem Körper K ist ein K -Modul.
- (2) Ein Ideal $\mathfrak{a} \subset R$ eines Ringes R ist ein R -Modul, insbesondere ist R selbst ein R -Modul. Für $n \in \mathbb{N} - \{0\}$ ist R^n in nahe liegender Weise ein R -Modul.
- (3) Jede abelsche Gruppe G ist ein $\mathbb{Z}$ -Modul; wie gewöhnlich erkläre man $n \cdot g$ für $n \in \mathbb{N}$ und $g \in G$ als n -fache Summe von g , sowie $(-n) \cdot g$ als $-(n \cdot g)$.
- (4) Jeder Endomorphismus $f: V \rightarrow V$ eines K -Vektorraums V induziert, wie oben erklärt, auf V die Struktur eines $K[T]$ -Moduls. Ist V endlich-dimensional und bezeichnet p_f das Minimalpolynom von f , so gilt $p_f \cdot v = 0$ für alle $v \in V$. Im Unterschied zu Vektorräumen kann man daher bei einem R -Modul M aus einer Gleichung $\alpha \cdot m = 0$ mit $\alpha \in R$, $m \in M$ nicht schließen, dass α oder m verschwinden.

Eine ganze Reihe von Begriffen, die bei Vektorräumen eine Rolle spielen, haben auch für Moduln ihre Bedeutung. Sei etwa M ein Modul über einem Ring R . Ein *Untermodul* von M ist eine nicht-leere Teilmenge $N \subset M$, so dass gilt:

$$\begin{aligned} a, b \in N &\implies a + b \in N, \\ \alpha \in R, a \in N &\implies \alpha \cdot a \in N \end{aligned}$$

Es ist N dann wieder ein R -Modul unter den von M ererbten Verknüpfungen. Betrachtet man einen Ring R als Modul über sich selbst, so stimmen die Ideale des Rings R mit den Untermoduln von R überein.

Ein *Homomorphismus* zwischen R -Moduln M und N ist eine Abbildung $\varphi: M \rightarrow N$, für die

$$\varphi(a + b) = \varphi(a) + \varphi(b), \quad \varphi(\alpha \cdot a) = \alpha \cdot \varphi(a),$$

für $a, b \in M$, $\alpha \in R$ gilt. Man spricht dabei auch von einer *R -linearen Abbildung*. Mono-, Epi- bzw. Isomorphismen von R -Moduln sind wie üblich als injektive, surjektive bzw. bijektive Homomorphismen erklärt.

Zu einem System $(a_i)_{i \in I}$ von Elementen aus M kann man den hiervon erzeugten Untermodul $M' \subset M$ betrachten, wobei letzterer durch

$$M' := \sum_{i \in I} R a_i := \left\{ \sum_{i \in I} \alpha_i a_i ; \alpha_i \in R, \alpha_i = 0 \text{ für fast alle } i \in I \right\}$$

definiert ist. Gilt $M = \sum_{i \in I} R a_i$, so nennt man $(a_i)_{i \in I}$ ein *Erzeugendensystem* von M . Man bezeichnet M als *endlich erzeugt* oder (in missbräuchlicher Sprechweise) als *endlich*, wenn M ein endliches Erzeugendensystem besitzt. Weiter

heißt ein System $(a_i)_{i \in I}$ von Elementen aus M *frei* (oder *linear unabhängig*), wenn aus einer Gleichung

$$\sum_{i \in I} \alpha_i a_i = 0$$

mit Koeffizienten $\alpha_i \in R$, die für fast alle $i \in I$ verschwinden, bereits $\alpha_i = 0$ für alle i folgt. Freie Erzeugendensysteme werden auch als *Basen* bezeichnet. Man beachte jedoch, dass Moduln im Unterschied zu Vektorräumen im Allgemeinen keine Basen besitzen; vgl. Beispiel (4) oben. Moduln, die eine Basis besitzen, heißen *frei*, bzw. *endlich frei*, wenn sie eine endliche Basis besitzen. Homomorphismen zwischen freien R -Moduln lassen sich wie gewöhnlich bezüglich gewählter Basen durch Matrizen mit Koeffizienten aus R beschreiben.

Sind $M_1, \dots, M_n$ Untermoduln eines R -Moduls M , so kann man deren *Summe*

$$M' = \left\{ \sum_{i=1}^n a_i; a_i \in M_i, i = 1, \dots, n \right\}$$

betrachten. Dabei ist M' wiederum ein Untermodul von M , und man schreibt $M' = \sum_{i=1}^n M_i$. Weiter sagt man, M' sei die *direkte Summe* der M_i , in Zeichen $M' = \bigoplus_{i=1}^n M_i$, wenn zusätzlich für jedes $a \in M'$ die jeweilige Darstellung $a = \sum_{i=1}^n a_i$ mit Elementen $a_i \in M_i$ eindeutig ist.

Wie im Falle von Vektorräumen lässt sich die direkte Summe $\bigoplus_{i=1}^n N_i$ von R -Moduln $N_1, \dots, N_n$, die nicht notwendig als Untermoduln eines R -Moduls N gegeben sind, auch *konstruieren*. Man setze nämlich

$$\bigoplus_{i=1}^n N_i = N_1 \times \dots \times N_n$$

und betrachte dieses kartesische Produkt wiederum als R -Modul, und zwar unter der komponentenweisen Addition bzw. skalaren Multiplikation. Es lässt sich dann N_i für $i = 1, \dots, n$ jeweils mit dem Untermodul

$$0 \times \dots \times 0 \times N_i \times 0 \times \dots \times 0 \subset N_1 \times \dots \times N_n$$

identifizieren, so dass der Modul $N_1 \times \dots \times N_n$ in der Tat als direkte Summe der Untermoduln $N_1, \dots, N_n$ aufzufassen ist.

Ist N ein Untermodul eines R -Moduls M , so kann man den *Restklassenmodul* M/N bilden. Wie im Falle von Vektorräumen gibt N nämlich Anlass zu einer Äquivalenzrelation auf M :

$$a \sim b \iff a - b \in N$$

Es besteht M/N aus den zugehörigen Äquivalenzklassen, d. h. aus den Nebenklassen $[a] = a + N$ zu Elementen $a \in M$. Die R -Modulstruktur auf M/N wird durch die Formeln

$$\begin{aligned} [a] + [b] &= [a + b], & a, b \in M, \\ \alpha \cdot [a] &= [\alpha \cdot a], & \alpha \in R, a \in M, \end{aligned}$$

gegeben, wobei natürlich die Wohldefiniertheit zu überprüfen ist. Die Homomorphiesätze 2.2/8 und 2.2/9 lassen sich ohne Änderungen übertragen. Insbesondere induziert ein R -Modulhomomorphismus $\varphi: M \rightarrow N$ stets einen injektiven R -Modulhomomorphismus $\bar{\varphi}: M/\ker \varphi \hookrightarrow N$.

Da Moduln im Allgemeinen keine Basen besitzen, lässt sich der Begriff der Dimension nicht ohne weiteres von Vektorräumen auf Moduln übertragen. Gewisse Aspekte des Dimensionsbegriffes werden durch die so genannte *Länge* eines Moduls abgedeckt. Hierunter versteht man für einen R -Modul M das Supremum $\ell_R(M)$ aller Längen ℓ von echt aufsteigenden Ketten von Untermoduln des Typs

$$0 \subsetneq M_1 \subsetneq M_2 \subsetneq \dots \subsetneq M_\ell = M.$$

Beispielsweise ist $\ell_R(M) = 0$ äquivalent zu $M = 0$. Weiter hat $\mathbb{Z}$ als freier Modul über sich selbst die Länge ∞ . Als Hilfsmittel für später benötigen wir zwei Lemmata.

Lemma 2. *Es sei R ein Hauptidealring und $a \in R$ ein Element mit Primfaktorzerlegung $a = p_1 \dots p_r$. Dann besitzt der Restklassenmodul R/aR die Länge $\ell_R(R/aR) = r$.*

Beweis. Es sei $\pi: R \rightarrow R/aR$ die kanonische Projektion. Da die Ideale $\bar{\mathfrak{a}} \subset R/aR$ unter der Zuordnung $\bar{\mathfrak{a}} \mapsto \pi^{-1}(\bar{\mathfrak{a}})$ bijektiv denjenigen Idealen in R entsprechen, die aR enthalten, stimmt die Länge von R/aR überein mit dem Supremum der Längen ℓ echt aufsteigender Idealketten des Typs

$$Ra \subsetneq \mathfrak{a}_1 \subsetneq \mathfrak{a}_2 \subsetneq \dots \subsetneq \mathfrak{a}_\ell = R.$$

Da R ein Hauptidealring ist, wird jedes $\mathfrak{a}_i$ von einem Element a_i erzeugt. Weiter ist eine echte Inklusion $\mathfrak{a}_{i-1} \subsetneq \mathfrak{a}_i$ gleichbedeutend damit, dass a_i ein echter Teiler von a_{i-1} ist. Die Länge von R/aR ist daher gleich dem Supremum aller $\ell \in \mathbb{N}$, so dass es $a_1, \dots, a_\ell \in R$ gibt mit der Eigenschaft, dass a_i jeweils ein echter Teiler von a_{i-1} ist, $i = 1, \dots, \ell$; dabei ist $a_0 = a$ zu setzen. Da in R der Satz von der eindeutigen Primfaktorzerlegung gilt und a_0 ein Produkt von r Primfaktoren ist, berechnet sich dieses Supremum zu r . $\square$

Lemma 3. *Ist ein R -Modul M die direkte Summe zweier Untermoduln M' und M'' , so gilt $\ell_R(M) = \ell_R(M') + \ell_R(M'')$.*

Beweis. Hat man echt aufsteigende Ketten von Untermoduln

$$\begin{aligned} 0 \subsetneq M'_1 \subsetneq M'_2 \subsetneq \dots \subsetneq M'_r = M', \\ 0 \subsetneq M''_1 \subsetneq M''_2 \subsetneq \dots \subsetneq M''_s = M'', \end{aligned}$$

so ist

$$\begin{aligned} 0 \subsetneq M'_1 \oplus 0 \subsetneq M'_2 \oplus 0 \subsetneq \dots \subsetneq M'_r \oplus 0 \\ \subsetneq M'_r \oplus M''_1 \subsetneq M'_r \oplus M''_2 \subsetneq \dots \subsetneq M'_r \oplus M''_s = M \end{aligned}$$

eine echt aufsteigende Kette der Länge $r + s$ in M . Also gilt

$$\ell_R(M) \geq \ell_R(M') + \ell_R(M'').$$

Zum Nachweis der umgekehrten Abschätzung betrachte man eine echt aufsteigende Kette von Untermoduln

$$0 = M_0 \subsetneq M_1 \subsetneq M_2 \subsetneq \dots \subsetneq M_\ell = M.$$

Es sei $\pi'': M' \oplus M'' \longrightarrow M''$ die Projektion auf den zweiten Summanden, so dass also $\ker \pi'' = M'$ gilt. Dann ist für $0 \leq \lambda < \ell$, wie wir sogleich sehen werden, jeweils $M_\lambda \cap M'$ echt enthalten in $M_{\lambda+1} \cap M'$ oder $\pi''(M_\lambda)$ echt enthalten in $\pi''(M_{\lambda+1})$. Hieraus folgt $\ell \leq \ell_R(M') + \ell_R(M'')$ und damit insgesamt wie gewünscht $\ell_R(M) = \ell_R(M') + \ell_R(M'')$.

Um das gerade behauptete Inklusionsverhalten zu rechtfertigen, nehmen wir einmal $M_\lambda \cap M' = M_{\lambda+1} \cap M'$ sowie $\pi''(M_\lambda) = \pi''(M_{\lambda+1})$ an und zeigen, dass dies bereits $M_\lambda = M_{\lambda+1}$ impliziert, im Widerspruch zu unserer Voraussetzung. In der Tat, zu $a \in M_{\lambda+1}$ gibt es wegen $\pi''(M_\lambda) = \pi''(M_{\lambda+1})$ ein $a' \in M_\lambda$ mit $\pi(a) = \pi(a')$, also mit $a - a' \in \ker \pi = M'$. Dann gilt wegen $a, a' \in M_{\lambda+1}$ sogar $a - a' \in M_{\lambda+1} \cap M' = M_\lambda \cap M'$ und damit $a = a' + (a - a') \in M_\lambda$. Es folgt daher $M_{\lambda+1} \subset M_\lambda$ bzw. $M_{\lambda+1} = M_\lambda$, was aber ausgeschlossen war. $\square$

Theorem 4 (Elementarteilersatz). *Es sei R ein Hauptidealring und F ein endlicher freier R -Modul. Weiter sei $M \subset F$ ein Untermodul. Dann existieren Elemente $x_1, \dots, x_s \in F$, die Teil einer Basis von F sind, sowie Koeffizienten $\alpha_1, \dots, \alpha_s \in R - \{0\}$, so dass gilt:*

- (i) $\alpha_1 x_1, \dots, \alpha_s x_s$ bilden eine Basis von M .
- (ii) $\alpha_i \mid \alpha_{i+1}$ für $1 \leq i < s$.

Dabei sind $\alpha_1, \dots, \alpha_s$ bis auf Assoziiertheit (d. h. bis auf Einheiten) eindeutig durch M bestimmt, unabhängig von der Wahl von $x_1, \dots, x_s$. Man nennt $\alpha_1, \dots, \alpha_s$ die Elementarteiler von $M \subset F$. Insbesondere ist deren Anzahl s eindeutig bestimmt.

Beweis der Existenzaussage von Theorem 4. Es sei $Y = (y_1, \dots, y_m)$ eine Basis von F . Wir zeigen zunächst per Induktion nach m , dass der Untermodul $M \subset F$ endlich erzeugt ist. Für $m = 1$ ist dies klar, denn F ist dann als R -Modul isomorph zu R , und M korrespondiert zu einem Ideal in R . Letzteres ist endlich erzeugt, da R ein Hauptidealring ist. Sei also $m > 1$. Man setze $F' = \sum_{i=1}^{m-1} R y_i$ und $F'' = R y_m$. Weiter betrachte man die Projektion $\pi: F \longrightarrow F''$, welche y_i für $i < m$ auf 0 und y_m auf y_m abbildet; es gilt dann $\ker \pi = F'$, und man hat — in der Sprache von Abschnitt 2.3 — eine kurze exakte Sequenz

$$0 \longrightarrow F' \longrightarrow F \longrightarrow F'' \longrightarrow 0.$$

Nun sind die Untermoduln $M \cap F' \subset F'$ und $\pi(M) \subset F''$ nach Induktionsvoraussetzung endlich erzeugt, und man zeigt wie üblich, z. B. wie im Beweis zu 2.1/10, dass ein Erzeugendensystem von $M \cap F'$ zusammen mit der Liftung

eines Erzeugendensystems von $\pi(M)$ insgesamt ein Erzeugendensystem von M bildet. M ist also endlich erzeugt.

Wir behalten $Y = (y_1, \dots, y_m)$ als Basis von F bei und wählen ein endliches Erzeugendensystem $z_1, \dots, z_n$ von M . Bezeichnet dann $e = (e_1, \dots, e_n)$ die kanonische Basis des R -Moduls R^n , so kann man die durch $e_j \mapsto z_j$ erklärte R -lineare Abbildung $f: R^n \rightarrow F$ betrachten, deren Bild M ergibt. Gilt dann

$$z_j = \sum_{i=1}^m \alpha_{ij} y_i, \quad j = 1, \dots, n,$$

so ist $A = (\alpha_{ij})_{i,j} \in R^{m \times n}$ die Matrix zu f bezüglich der Basen e und Y . Wir verwenden nun folgendes Hilfsresultat, das wir weiter unten beweisen werden:

Lemma 5. *Es sei R ein Hauptidealring und $A = (\alpha_{ij}) \in R^{m \times n}$ eine Matrix mit Koeffizienten aus R . Dann gibt es invertierbare Matrizen $S \in R^{m \times m}$ und $T \in R^{n \times n}$ mit*

$$S \cdot A \cdot T = \begin{pmatrix} \alpha_1 & 0 & & 0 & 0 & \dots & 0 \\ 0 & \alpha_2 & & 0 & 0 & \dots & 0 \\ & & \ddots & & & & \\ & & & \ddots & & & \\ 0 & 0 & & \alpha_s & 0 & \dots & 0 \\ 0 & 0 & & 0 & 0 & \dots & 0 \\ 0 & 0 & & 0 & 0 & \dots & 0 \end{pmatrix}$$

und mit Koeffizienten $\alpha_1, \dots, \alpha_s \in R - \{0\}$ (wobei $0 \leq s \leq \min(m, n)$ gilt), die für $1 \leq i < s$ die Bedingung $\alpha_i \mid \alpha_{i+1}$ erfüllen. Dabei sind $\alpha_1, \dots, \alpha_s$ bis auf Assoziiertheit eindeutig bestimmt; man nennt sie die Elementarteiler der Matrix A .

Indem man S und T als Basiswechselmatrizen auffasst, sieht man, dass die Matrix SAT ebenfalls die Abbildung f beschreibt, allerdings bezüglich geeigneter anderer Basen $e'_1, \dots, e'_n$ von R^n und $x_1, \dots, x_m$ von F . Insbesondere folgt, dass M als Bild von f durch $\alpha_1 x_1, \dots, \alpha_s x_s$ erzeugt wird. Da das System der $x_1, \dots, x_m$ frei ist und wir Koeffizienten aus einem Integritätsring R betrachten, bilden $\alpha_1 x_1, \dots, \alpha_s x_s$ sogar eine Basis von M . Damit haben wir die Existenz der Elementarteiler $\alpha_1, \dots, \alpha_s$ von $M \subset F$ auf die Existenzaussage von Lemma 5 zurückgeführt. $\square$

Beweis der Existenzaussage von Lemma 5. Wir nehmen zunächst R als euklidischen Ring an und zeigen anhand eines konstruktiven Verfahrens unter Verwendung der Division mit Rest, dass sich die Matrix $A = (\alpha_{ij})$ durch reversible elementare Zeilen- und Spaltenumformungen in die gewünschte Gestalt bringen lässt, nämlich durch Vertauschen von Zeilen (bzw. Spalten) sowie durch Addieren eines Vielfachen einer Zeile (bzw. Spalte) zu einer weiteren Zeile (bzw. Spalte). Wie im Fall einer Matrix mit Koeffizienten aus einem Körper sind elementare Umformungen dieses Typs als Multiplikation mit einer invertierbaren

Elementarmatrix von links (bzw. rechts) zu interpretieren. Die benötigten Zeilenumformungen korrespondieren daher insgesamt zur Multiplikation mit einer invertierbaren Matrix $S \in R^{m \times m}$ von links, die benötigten Spaltenumformungen entsprechend zur Multiplikation mit einer invertierbaren Matrix $T \in R^{n \times n}$ von rechts. Anschließend verallgemeinern wir das Verfahren, so dass es in modifizierter Version auch für Hauptidealringe anwendbar ist.

Wir betrachten im Folgenden also zunächst einen euklidischen Ring R mit Gradabbildung $\delta: R - \{0\} \rightarrow \mathbb{N}$. Für $A = 0$ ist nichts zu zeigen, so dass wir $A \neq 0$ annehmen dürfen. Es ist unsere Strategie, A mittels elementarer Umformungen so abzuändern, dass sich das Minimum

$$d(A) := \min\{\delta(\alpha) ; \alpha \text{ ist Koeffizient } \neq 0 \text{ von } A\}$$

schrittweise verringert. Da δ Werte in $\mathbb{N}$ annimmt, muss dieses Verfahren nach endlich vielen Schritten abbrechen. Ist dann $\alpha \neq 0$ ein Koeffizient der transformierten Matrix mit minimalem Grad $\delta(\alpha)$, so zeigen wir mittels Division mit Rest, dass α alle anderen Koeffizienten der Matrix teilt; α ist dann der erste Elementarteiler von A .

Im Einzelnen gehen wir wie folgt vor. Indem wir Zeilen und Spalten in A vertauschen, können wir $d(A) = \delta(\alpha_{11})$ annehmen, dass also $\delta(\alpha_{11})$ minimal ist unter allen $\delta(\alpha_{ij})$ mit $\alpha_{ij} \neq 0$. Diese Situation stellen wir zu Beginn eines jeden Schrittes her. Ist dann eines der Elemente der 1. Spalte, etwa α_{i1} , nicht durch α_{11} teilbar, so teile man α_{i1} mit Rest durch α_{11} , etwa $\alpha_{i1} = q\alpha_{11} + \beta$ mit $\delta(\beta) < \delta(\alpha_{11})$, und ziehe das q -fache der 1. Zeile von der i -ten Zeile ab. Als Resultat entsteht an der Position $(i, 1)$ das Element β . Das Minimum $d(A)$ der Grade von nichtverschwindenden Koeffizienten von A hat sich daher verringert, und man starte das Verfahren erneut mit einem weiteren Schritt. In gleicher Weise können wir die Elemente der 1. Zeile mittels elementarer Spaltenumformungen abändern. Da $d(A)$ Werte in $\mathbb{N}$ annimmt, also nicht beliebig oft verringert werden kann, ist nach endlich vielen Schritten jedes Element der 1. Spalte sowie der 1. Zeile ein Vielfaches von α_{11} , und wir können durch Addition von Vielfachen der 1. Zeile zu den restlichen Zeilen der Matrix annehmen, dass $\alpha_{i1} = 0$ für $i > 1$ gilt. Entsprechend können wir mit der 1. Zeile verfahren und auf diese Weise $\alpha_{i1} = \alpha_{1j} = 0$ für $i, j > 1$ erreichen. Dabei dürfen wir weiter annehmen, dass das Minimum $d(A)$ mit $\delta(\alpha_{11})$ übereinstimmt; ansonsten ist das Verfahren erneut zu beginnen und ein entsprechendes Element an der Stelle $(1, 1)$ neu zu positionieren. Existieren nun $i, j > 1$ mit $\alpha_{11} \nmid \alpha_{ij}$, so addiere man die j -te Spalte zur ersten, ein Prozess, der α_{11} unverändert lässt. Wie gerade beschrieben, lassen sich die Elemente unterhalb α_{11} erneut trivialisieren, und zwar unter Verringerung des Grades $d(A)$. Nach endlich vielen Schritten gelangt man so zu einer Matrix (α_{ij}) mit $\alpha_{i1} = \alpha_{1j} = 0$ für $i, j > 1$ sowie mit der Eigenschaft, dass α_{11} jedes andere Element α_{ij} mit $i, j > 1$ teilt. Man behandle dann in gleicher Weise die Untermatrix $(\alpha_{ij})_{i,j>1}$ von $A = (\alpha_{ij})$, sofern diese nicht bereits Null ist. Die hierfür benötigten Umformungen lassen die erste Zeile und Spalte von A invariant und erhalten insbesondere die Bedingung, dass α_{11} alle restlichen Koeffizienten von A teilt. Führt man dieses Verfahren in

induktiver Weise fort, so gelangt man schließlich nach endlich vielen Schritten zu einer Matrix, auf deren Hauptdiagonalen die gesuchten Elementarteiler mit der behaupteten Teilbarkeitseigenschaft stehen und deren sonstige Einträge alle verschwinden. Damit ist die Existenzaussage von Lemma 5 und insbesondere auch von Theorem 4 bewiesen, zumindest im Falle eines euklidischen Rings R .

Ist nun R lediglich als Hauptidealring bekannt, so benötigen wir elementare Matrizenumformungen eines etwas allgemeineren Typs, die wir zunächst erläutern wollen. Seien $\sigma, \tau, \sigma', \tau' \in R$ mit $\sigma\tau' - \tau\sigma' = 1$ gewählt. Dann sind die Matrizen

$$\begin{pmatrix} \sigma & \tau \\ \sigma' & \tau' \end{pmatrix}, \quad \begin{pmatrix} \tau' & -\tau \\ -\sigma' & \sigma \end{pmatrix} \in R^{2 \times 2}$$

invers zueinander, insbesondere also invertierbar. Entsprechend sieht man, dass für $\sigma\tau' - \tau\sigma' = \pm 1$ und $1 \leq i < j \leq m$ auch die Matrizen des Typs

$$E^{ij}(\sigma, \tau, \sigma', \tau') = \begin{pmatrix} 1 & & & \\ & \sigma & \tau & \\ & & 1 & \\ & \sigma' & \tau' & \\ & & & 1 \end{pmatrix} \in R^{m \times m}$$

invertierbar sind. Hierbei stehen die Elemente $\sigma, \tau, \sigma', \tau'$ jeweils an den Positionen $(i, i), (i, j), (j, i), (j, j)$, und mit "1" sind Serien von Elementen 1 auf der Diagonalen angedeutet. Im Übrigen ist die Matrix $E^{ij}(\sigma, \tau, \sigma', \tau')$ mit Elementen 0 aufgefüllt, die der Übersichtlichkeit halber aber nicht ausgedrückt sind. Multipliziert man nun A von links mit $E^{ij}(\sigma, \tau, \sigma', \tau')$, so hat dies folgenden Effekt: Als neue i -te Zeile erhält man die Summe des σ -fachen der alten i -ten Zeile und des τ -fachen der alten j -ten Zeile. Entsprechend ist die neue j -te Zeile die Summe des σ' -fachen der alten i -ten Zeile und des τ' -fachen der alten j -ten Zeile. Beispielsweise ergibt sich eine Vertauschung der i -ten und j -ten Zeile mit den Konstanten

$$\sigma = 0, \quad \tau = 1, \quad \sigma' = 1, \quad \tau' = 0,$$

sowie die Addition des ε -fachen der j -ten Zeile zur i -ten Zeile mit

$$\sigma = 1, \quad \tau = \varepsilon, \quad \sigma' = 0, \quad \tau' = 1.$$

Mittels Transponierens sieht man, dass analoge Spaltenumformungen von A durch Multiplikation von rechts mit Matrizen des Typs $E^{ij}(\sigma, \tau, \sigma', \tau') \in R^{n \times n}$ generiert werden können.

Wir bezeichnen nun für Elemente $\alpha \in R - \{0\}$ mit $\delta(\alpha)$ die Anzahl der Primfaktoren von α ; dies ist gemäß Lemma 2 gerade die Länge des Restklassenrings $R/\alpha R$. Weiter setzen wir, ähnlich wie im Falle euklidischer Ringe,

$$d(A) := \min\{\delta(\alpha) ; \alpha \text{ ist Koeffizient } \neq 0 \text{ von } A\}$$

mit dem Ziel, $d(A)$ schrittweise zu verringern, solange bis es einen Koeffizienten α von A gibt, der alle übrigen Koeffizienten teilt. Durch Vertauschen von Zeilen

und Spalten können wir wiederum $d(A) = \delta(\alpha_{11})$ annehmen. Ist nun eines der Elemente der 1. Spalte, etwa α_{i1} , kein Vielfaches von α_{11} , so bilde man den größten gemeinsamen Teiler β von α_{11} und α_{i1} . Für diesen gilt dann notwendig $\delta(\beta) < \delta(\alpha_{11})$, und es erzeugt β gemäß 5.2/16 das Ideal $R\alpha_{11} + R\alpha_{i1}$, d. h. es existiert eine Gleichung des Typs

$$\beta = \sigma\alpha_{11} + \tau\alpha_{i1},$$

wobei $\sigma, \tau \in R$ notwendig teilerfremd sind und damit eine Gleichung des Typs

$$\sigma\tau' - \tau\sigma' = 1$$

mit gewissen Elementen $\sigma', \tau' \in R$ erfüllen. Multipliziert man nun A von links mit $E^{1i}(\sigma, \tau, \sigma', \tau')$, so etabliert dieser Prozess in A an der Position $(1, 1)$ das Element β und verringert somit das Minimum $d(A)$. Iteriert man das Verfahren wie im Falle euklidischer Ringe, so kann man schließlich erreichen, dass die Elemente $\alpha_{21}, \dots, \alpha_{m1}$ durch α_{11} teilbar sind bzw., indem man geeignete Vielfache der 1. Zeile von den restlichen subtrahiert, dass $\alpha_{21} = \dots = \alpha_{m1} = 0$ gilt. In gleicher Weise kann man mittels entsprechender Spaltenumformungen die Elemente $\alpha_{12}, \dots, \alpha_{1n}$ trivialisieren usw. Wir sehen also, dass sich die Matrix A schrittweise wie im Falle euklidischer Ringe abändern lässt, bis schließlich die gewünschte Gestalt erreicht ist. $\square$

Wir wollen das für euklidische Ringe beschriebene Verfahren an einem einfachen Beispiel demonstrieren und betrachten hierzu die Matrix

$$A = \begin{pmatrix} 6 & 2 & 5 \\ 32 & 2 & 28 \\ 30 & 2 & 26 \end{pmatrix} \in \mathbb{Z}^{3 \times 3}.$$

Der Bequemlichkeit halber lassen wir zur Bestimmung der Elementarteiler von A neben den oben verwendeten elementaren Zeilen- und Spaltenumformungen auch noch die Multiplikation einer Zeile bzw. Spalte mit einer Einheit unseres Ringes $R = \mathbb{Z}$ zu. Dies ist erlaubt, denn auch diese Umformungen lassen sich als Multiplikation von links bzw. rechts mit invertierbaren Elementarmatrizen interpretieren, und zwar mit solchen, die aus der Einheitsmatrix hervorgehen, indem man einen der Diagonaleinträge 1 durch eine Einheit aus R ersetzt. Wir wollen uns ansonsten aber an das für euklidische Ringe geschilderte Verfahren halten, obwohl sich die Bestimmung der Elementarteiler von A durch eine geschicktere Wahl der elementaren Umformungen noch vereinfachen ließe.

$$\begin{aligned}
A &= \begin{pmatrix} 6 & 2 & 5 \\ 32 & 2 & 28 \\ 30 & 2 & 26 \end{pmatrix} \xrightarrow{(1)} \begin{pmatrix} 2 & 6 & 5 \\ 2 & 32 & 28 \\ 2 & 30 & 26 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 2 & 6 & 5 \\ 0 & 26 & 23 \\ 0 & 24 & 21 \end{pmatrix} \xrightarrow{(3)} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 26 & 23 \\ 0 & 24 & 21 \end{pmatrix} \\
&\xrightarrow{(4)} \begin{pmatrix} 1 & 0 & 2 \\ 23 & 26 & 0 \\ 21 & 24 & 0 \end{pmatrix} \xrightarrow{(5)} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 26 & -46 \\ 0 & 24 & -42 \end{pmatrix} \xrightarrow{(6)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 26 & 46 \\ 0 & 24 & 42 \end{pmatrix} \xrightarrow{(7)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 24 & 42 \\ 0 & 26 & 46 \end{pmatrix} \\
&\xrightarrow{(8)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 24 & 42 \\ 0 & 2 & 4 \end{pmatrix} \xrightarrow{(9)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 24 & 42 \end{pmatrix} \xrightarrow{(10)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 0 & -6 \end{pmatrix} \xrightarrow{(11)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix}
\end{aligned}$$

Es ergeben sich also 1, 2, 6 als die Elementarteiler von A , wobei im Einzelnen die folgenden elementaren Umformungen ausgeführt wurden:

- (1) Vertauschen von 1. und 2. Spalte
- (2) Subtrahieren der 1. von der 2. und der 3. Zeile
- (3) Subtrahieren des 3-fachen bzw. 2-fachen der 1. Spalte von der 2. bzw. 3. Spalte
- (4) Vertauschen von 1. und 3. Spalte
- (5) Subtrahieren des 23-fachen bzw. 21-fachen der 1. Zeile von der 2. bzw. 3. Zeile
- (6) Subtrahieren des 2-fachen der 1. Spalte von der 3. Spalte, Multiplikation der 3. Spalte mit -1
- (7) Vertauschen von 2. und 3. Zeile
- (8) Subtrahieren der 2. Zeile von der 3. Zeile
- (9) Vertauschen von 2. und 3. Zeile
- (10) Subtrahieren des 12-fachen der 2. Zeile von der 3. Zeile
- (11) Subtrahieren des 2-fachen der 2. Spalte von der 3. Spalte, Multiplikation der 3. Spalte mit -1

Als Nächstes wenden wir uns nun der Eindeutigkeitsaussage in Lemma 5 bzw. Theorem 4 zu und beginnen mit einem grundlegenden Lemma, welches auch später noch von Bedeutung sein wird.

Lemma 6. *Es sei R ein Hauptidealring und*

$$Q \simeq \bigoplus_{i=1}^s R/\alpha_i R$$

ein Isomorphismus von R -Moduln, wobei $\alpha_1, \dots, \alpha_s \in R - \{0\}$ Nichteinheiten mit $\alpha_i \mid \alpha_{i+1}$ für $1 \leq i < s$ sind und $\bigoplus_{i=1}^s R/\alpha_i R$ die konstruierte direkte Summe der R -Moduln $R/\alpha_i R$ bezeichne. Dann sind $\alpha_1, \dots, \alpha_s$ bis auf Assoziiertheit eindeutig durch Q bestimmt.

Beweis. Aus technischen Gründen invertieren wir die Nummerierung der α_i und betrachten zwei Zerlegungen

$$Q \simeq \bigoplus_{i=1}^s R/\alpha_i R \simeq \bigoplus_{j=1}^t R/\beta_j R$$

mit $\alpha_{i+1} \mid \alpha_i$ für $1 \leq i < s$ sowie $\beta_{j+1} \mid \beta_j$ für $1 \leq j < t$. Falls es einen Index $k \leq \min\{s, t\}$ mit $\alpha_k R \neq \beta_k R$ gibt, so wähle man k minimal mit dieser Eigenschaft. Da $\alpha_i R = \beta_i R$ für $1 \leq i < k$ und da $\alpha_{k+1}, \dots, \alpha_s$ sämtlich Teiler von α_k sind, zerlegt sich $\alpha_k Q$ zu

$$\begin{aligned} \alpha_k Q &\simeq \bigoplus_{i=1}^{k-1} \alpha_k \cdot (R/\alpha_i R) \\ &\simeq \bigoplus_{i=1}^{k-1} \alpha_k \cdot (R/\alpha_i R) \oplus \bigoplus_{j=k}^t \alpha_k \cdot (R/\beta_j R). \end{aligned}$$

Wir benutzen nun die Lemmata 2 und 3. Wegen $\ell_R(\alpha_k Q) \leq \ell_R(Q) < \infty$ ergibt sich durch Vergleich beider Seiten $\ell_R(\alpha_k \cdot (R/\beta_j R)) = 0$ für $j = k, \dots, t$. Letzteres bedeutet aber insbesondere $\alpha_k \cdot (R/\beta_k R) = 0$ bzw. $\alpha_k R \subset \beta_k R$. Entsprechend zeigt man $\beta_k R \subset \alpha_k R$ und somit $\alpha_k R = \beta_k R$, im Widerspruch zu unserer Annahme. Es gilt daher $\alpha_i R = \beta_i R$ für alle Indizes i mit $1 \leq i \leq \min\{s, t\}$. Hat man weiter $s \leq t$, so folgt, wiederum unter Benutzung von Lemma 3, dass $\bigoplus_{j=s+1}^t R/\beta_j R$ von der Länge 0 ist, also $R/\beta_j R$ für $j = s+1, \dots, t$ die Länge 0 hat. Andererseits ist aber β_j für $j = 1, \dots, t$ keine Einheit. Daher besitzt jeder Modul $R/\beta_j R$ gemäß Lemma 2 eine Länge größer als 0, und es folgt $s = t$. $\square$

Beweis der Eindeutigkeitsaussage von Theorem 4 und Lemma 5. Die Eindeutigkeit in Lemma 5 ist eine Konsequenz der Eindeutigkeitsaussage in Theorem 4. Es genügt daher, die Eindeutigkeit der Elementarteiler in Theorem 4 zu zeigen. Seien also $x_1, \dots, x_s$ Teil einer Basis von F , und seien $\alpha_1, \dots, \alpha_s \in R$ mit $\alpha_i \mid \alpha_{i+1}$ für $1 \leq i < s$, so dass $\alpha_1 x_1, \dots, \alpha_s x_s$ eine Basis von M bilden. Betrachtet man dann den Untermodul

$$F' = \bigoplus_{i=1}^s R x_i = \{a \in F; \text{ es existiert ein } \alpha \in R - \{0\} \text{ mit } \alpha a \in M\}$$

von F , so hängt F' nur von M und nicht von der speziellen Wahl der Elemente $x_1, \dots, x_s$ ab. Der kanonische Homomorphismus

$$F' \longrightarrow \bigoplus_{i=1}^s R/\alpha_i R, \quad \sum_{i=1}^s \gamma_i x_i \longmapsto (\overline{\gamma}_1, \dots, \overline{\gamma}_s),$$

wobei $\overline{\gamma}_i$ jeweils die Restklasse von γ_i in R/α_i bezeichne, ist dann surjektiv und besitzt M als Kern, induziert also aufgrund des Homomorphiesatzes einen Isomorphismus

$$F'/M \xrightarrow{\sim} \bigoplus_{i=1}^s R/\alpha_i R.$$

Aus der Eindeutigkeitsaussage in Lemma 6 ist dann zu folgern, dass jedenfalls die Nichteinheiten unter den $\alpha_1, \dots, \alpha_s$ bis auf Assoziiertheit eindeutig durch M bestimmt sind.

Um nun einzusehen, dass auch die Anzahl der Einheiten unter den $\alpha_1, \dots, \alpha_s$ eindeutig durch M bestimmt ist, wollen wir zeigen, dass s als Anzahl der Elemente einer Basis von M bzw. F' eindeutig bestimmt ist. Da jede Basis $z_1, \dots, z_s$ von M Anlass zu einem Isomorphismus $M \xrightarrow{\sim} R^s$ gibt, genügt es zu zeigen, dass die Existenz eines Isomorphismus $R^s \xrightarrow{\sim} R^{s'}$ bereits $s = s'$ nach sich zieht. Im Falle eines Körpers R ist dies klar aufgrund der Dimensionstheorie für Vektorräume; vgl. z. B. 2.1/8. Ist jedoch R kein Körper, so enthält $R - R^*$ mindestens ein von Null verschiedenes Element, und dieses lässt sich gemäß 5.2/13 als Produkt von Primelementen schreiben. Man findet daher in R mindestens ein Primelement p , und es ist leicht nachzuprüfen, dass jeder Isomorphismus von R -Moduln $R^s \xrightarrow{\sim} R^{s'}$ einen Isomorphismus von R/pR -Moduln

$$(R/pR)^s = R^s/pR^s \xrightarrow{\sim} R^{s'}/pR^{s'} = (R/pR)^{s'}$$

induziert. Da aber R/pR nach 5.2/17 ein Körper ist, können wir wiederum $s = s'$ schließen. Alternativ kann man an dieser Stelle auch die Eindeutigkeitsaussage von Lemma 6 ausnutzen. $\square$

Korollar 7. *Es sei R ein Hauptidealring und F ein endlicher freier R -Modul. Dann besitzen je zwei Basen von F gleiche Länge.¹ Diese Länge wird auch als der Rang von F bezeichnet.*

Beweis. Die Behauptung wurde bereits in obigem Beweis hergeleitet, um nachzuweisen, dass die Anzahl der Elementarteiler eines Untermoduls $M \subset F$ eindeutig bestimmt ist. Im Übrigen folgt die Aussage von Korollar 7 aber auch formal aus der Eindeutigkeitsaussage von Theorem 4, da die Anzahl der Elementarteiler des trivialen Untermoduls $F \subset F$ eindeutig bestimmt ist. $\square$

Für die Berechnung von Elementarteilern in der Situation von Theorem 4 ist es wichtig zu wissen, dass wir diese als Elementarteiler einer Matrix erhalten haben, denn die Elementarteiler einer Matrix können mit Hilfe des im Beweis zu Lemma 5 gegebenen praktischen Verfahrens bestimmt werden. Wir wollen die genauen Bedingungen hier noch einmal gesondert formulieren.

Korollar 8. *Es sei R ein Hauptidealring, F ein endlicher freier R -Modul und $M \subset F$ ein Untermodul. Dann ist M endlich erzeugt.*

Es sei F' ein weiterer endlicher freier R -Modul und $f: F' \rightarrow F$ eine R -lineare Abbildung, welche eine Basis von F' auf ein Erzeugendensystem von M abbildet, also mit $\text{Bild } f = M$. Ist dann A eine Matrix mit Koeffizienten aus R , welche f bezüglich geeigneter Basen in F und F' beschreibt, so stimmen die Elementarteiler von A überein mit denjenigen von $M \subset F$.

¹ Die Aussage gilt allgemeiner für endliche freie Moduln über beliebigen Ringen (kommutativ mit 1); vgl. Aufgabe 6.

Aufgaben

Es sei R ein beliebiger Ring (kommutativ mit 1), sofern nichts anderes verlangt ist.

1. Man bestimme die Elementarteiler der folgenden Matrizen:

$$\begin{pmatrix} 2 & 6 & 8 \\ 3 & 1 & 2 \\ 9 & 5 & 4 \end{pmatrix}, \quad \begin{pmatrix} 4 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 15 \end{pmatrix} \in \mathbb{Z}^{3 \times 3}$$

2. Es sei R ein Hauptidealring und $A = (\alpha_{ij})_{i,j} \in R^{m \times n}$ eine nicht-triviale Matrix mit Koeffizienten aus R . Sind dann $\alpha_1, \dots, \alpha_s$ die Elementarteiler von A , so gilt $s > 0$ und $\alpha_1 = \text{ggT}(\alpha_{ij}; i = 1, \dots, m, j = 1, \dots, n)$.
3. Es seien $a_{11}, \dots, a_{1n}$ teilerfremde Elemente eines Hauptidealrings R , d. h. es gelte $\text{ggT}(a_{11}, \dots, a_{1n}) = 1$. Man zeige, es gibt Elemente $a_{ij} \in R$, $i = 2, \dots, n$, $j = 1, \dots, n$, so dass die Matrix $(a_{ij})_{i,j=1,\dots,n}$ in $R^{n \times n}$ invertierbar ist.
4. Es sei $f: M \rightarrow N$ ein Homomorphismus endlicher freier Moduln über einem Hauptidealring R , d. h. M und N mögen jeweils endliche Basen besitzen. Man verwende die Aussage des Elementarteilersatzes und folgere die Existenz von Basen X von M und Y von N , sowie von Null verschiedener Elemente $\alpha_1, \dots, \alpha_s \in R$ mit $\alpha_i | \alpha_{i+1}$ für $1 \leq i < s$, so dass gilt:

$$A_{f,X,Y} = \begin{pmatrix} \alpha_1 & 0 & & 0 & 0 & \dots & 0 \\ 0 & \alpha_2 & & 0 & 0 & \dots & 0 \\ & & \ddots & & & & \\ & & & \ddots & & & \\ 0 & 0 & & & \alpha_s & 0 & \dots & 0 \\ 0 & 0 & & & 0 & 0 & \dots & 0 \\ 0 & 0 & & & 0 & 0 & \dots & 0 \end{pmatrix}$$

Dabei sind die Elemente $\alpha_1, \dots, \alpha_s$ bis auf Assoziiiertheit eindeutig bestimmt.

5. Man bestimme die Länge von $(\mathbb{Z}/15\mathbb{Z})^4$ als $\mathbb{Z}$ -Modul.
6. Es sei M ein endlich erzeugter R -Modul, der zudem frei ist. Man zeige:
- (i) M besitzt eine endliche Basis.
 - (ii) Je zwei Basen von M bestehen aus gleichviel Elementen.
- (Hinweis: Um die Argumentation zu erleichtern, nehme man für Teil (ii) an, dass R ein Integritätsring ist.)
7. Für den Elementarteilersatz (Theorem 4) hatten wir einen Untermodul M eines endlichen freien Moduls F betrachtet. Man zeige, dass die Aussage des Satzes erhalten bleibt, wenn man alternativ F als frei und M als endlich erzeugt voraussetzt.
8. Es sei V ein Vektorraum über einem Körper K . Zu jedem Endomorphismus $f: V \rightarrow V$ kann man auf V die zugehörige Struktur als $K[T]$ -Modul betrachten, welche charakterisiert ist durch $T \cdot v = f(v)$ für Elemente $v \in V$. Man zeige, dass man auf diese Weise eine Bijektion zwischen der Menge $\text{End}_K(V)$ und der Menge der $K[T]$ -Modul-Strukturen auf V erhält, die verträglich sind mit der Struktur von V als K -Vektorraum.

9. (1. *Isomorphiesatz für Moduln*) Es seien N, N' Untermoduln eines R -Moduls M . Man zeige: Die kanonische Abbildung $N \hookrightarrow N + N' \longrightarrow (N + N')/N'$ besitzt $N \cap N'$ als Kern und induziert einen Isomorphismus

$$N/(N \cap N') \xrightarrow{\sim} (N + N')/N'.$$

10. (2. *Isomorphiesatz für Moduln*) Es seien M ein R -Modul und $N \subset N' \subset M$ Untermoduln. Man zeige:

- (i) Die kanonische Abbildung $N' \hookrightarrow M \longrightarrow M/N$ besitzt N als Kern und induziert einen Monomorphismus $N'/N \hookrightarrow M/N$. Folglich lässt sich N'/N mit seinem Bild in M/N identifizieren und somit als Untermodul von M/N auffassen.
- (ii) Die Projektion $M \longrightarrow M/N'$ faktorisiert über M/N , d. h. lässt sich als Komposition $M \xrightarrow{\pi} M/N \xrightarrow{f} M/N'$ schreiben, mit einem Modulhomomorphismus f und der kanonischen Projektion π .
- (iii) f besitzt N'/N als Kern und induziert einen Isomorphismus

$$(M/N)/(N'/N) \xrightarrow{\sim} M/N'.$$

6.4 Endlich erzeugte Moduln über Hauptidealringen

Wir wollen nun einige Folgerungen zur Struktur endlich erzeugter Moduln über Hauptidealringen aus dem Elementarteilersatz ziehen. Im nächsten Abschnitt sollen die gewonnenen Struktursätze dann in Ergebnisse über Normalformen von Endomorphismen von Vektorräumen umgesetzt werden. Als Hilfsmittel benötigen wir noch den so genannten *Chinesischen Restsatz*, den wir als Erstes beweisen.

Satz 1. *Es sei R ein Hauptidealring und*

$$a = \varepsilon p_1^{n_1} \dots p_r^{n_r}$$

eine Primfaktorzerlegung in R mit einer Einheit ε und paarweise nicht-assozierten Primelementen p_i . Ist dann

$$\pi_i: R \longrightarrow R/(p_i^{n_i}), \quad i = 1, \dots, r,$$

jeweils die kanonische Projektion, so ist der Homomorphismus

$$\varphi: R \longrightarrow R/(p_1^{n_1}) \times \dots \times R/(p_r^{n_r}), \quad a \longmapsto (\pi_1(a), \dots, \pi_r(a)),$$

surjektiv und erfüllt $\ker \varphi = (a)$, induziert also einen Isomorphismus

$$R/(a) \xrightarrow{\sim} R/(p_1^{n_1}) \times \dots \times R/(p_r^{n_r}).$$

Dabei ist $R/(p_1^{n_1}) \times \dots \times R/(p_r^{n_r})$ als Ring unter komponentenweiser Addition und Multiplikation zu verstehen.

Beweis. Zunächst zeigen wir, dass φ surjektiv ist. Hierzu genügt es offenbar nachzuprüfen, dass es Elemente $e_1, \dots, e_r \in R$ gibt mit

$$\pi_i(e_j) = \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{sonst.} \end{cases}$$

Da die Elemente $p_j^{n_j}$ und $\prod_{i \neq j} p_i^{n_i}$ für fest gewähltes $j \in \{1, \dots, r\}$ teilerfremd sind, ist das von ihnen erzeugte Hauptideal das Einheitsideal. Folglich existiert für jedes j eine Gleichung des Typs

$$e_j + d_j = 1 \quad \text{mit } e_j \in \left(\prod_{i \neq j} p_i^{n_i}\right), \quad d_j \in (p_j^{n_j}).$$

Es gilt dann $\pi_i(e_j) = 0$ für $i \neq j$ und $\pi_j(e_j) = 1$, wie gewünscht.

Der Kern von φ besteht aus allen Elementen aus R , die durch die Potenzen $p_i^{n_i}$, $i = 1, \dots, r$, teilbar sind, und damit aus den Elementen, die durch a teilbar sind. Somit ergibt sich $\ker \varphi = (a)$, und der behauptete Isomorphismus folgt aus dem Isomorphiesatz 5.1/10. $\square$

In der Situation von Satz 1 lässt sich der Restklassenring $R/(a)$ auch als R -Modul auffassen. Die Aussage des Chinesischen Restsatzes besagt dann, dass $R/(a)$ isomorph zu der konstruierten direkten Summe der R -Moduln $R/(p_i^{n_i})$ ist, also

$$R/(a) \simeq R/(p_1^{n_1}) \oplus \dots \oplus R/(p_r^{n_r}),$$

wobei wir, wie zu Beginn von Abschnitt 6.3 erläutert, den i -ten Summanden $R/(p_i^{n_i})$ mit dem entsprechenden Untermodul

$$0 \times \dots \times 0 \times R/(p_i^{n_i}) \times 0 \times \dots \times 0 \subset R/(p_1^{n_1}) \times \dots \times R/(p_r^{n_r})$$

zu identifizieren haben.

Für einen Modul M über einem *Integritätsring* R definiert man den so genannten *Torsionsuntermodul* T durch

$$T = \{a \in M; \text{ es existiert ein } \alpha \in R - \{0\} \text{ mit } \alpha a = 0\}.$$

Man prüft leicht nach, dass T in der Tat ein Untermodul von M ist, indem man die Nullteilerfreiheit von R benutzt. Es heißt M ein *Torsionsmodul*, wenn M mit seinem Torsionsuntermodul übereinstimmt. Der Torsionsuntermodul eines freien Moduls ist stets trivial, freie Moduln sind daher sozusagen als das “Gegenstück” zu den Torsionsmoduln anzusehen.

Ist M ein Modul über einem Ring R , so bezeichnet man eine exakte Sequenz von R -linearen Abbildungen

$$R^n \xrightarrow{\psi} R^m \xrightarrow{\varphi} M \longrightarrow 0$$

mit geeigneten endlichen freien R -Moduln R^m und R^n auch als eine *endliche Präsentation* von M . Genauer besteht diese aus einem Epimorphismus

$\varphi: R^m \longrightarrow M$ und einer R -linearen Abbildung $\psi: R^n \longrightarrow R^m$ mit $\text{im } \psi = \ker \varphi$. Eine solche endliche Präsentation existiert stets, wenn M ein endlich erzeugter Modul über einem Hauptidealring R ist. Man wähle nämlich ein endliches Erzeugendensystem $z_1, \dots, z_m$ in M und betrachte die R -lineare Abbildung $\varphi: R^m \longrightarrow M$, welche die kanonische Basis von R^m auf $z_1, \dots, z_m$ abbildet. Dann ist φ ein Epimorphismus, und der Untermodul $\ker \varphi \subset R^m$ ist gemäß 6.3/4 endlich erzeugt. Wir können daher einen weiteren Epimorphismus $\psi: R^n \longrightarrow \ker \varphi \subset R^m$ und folglich eine endliche Präsentation von M finden. Ist in dieser Situation $A \in R^{m \times n}$ eine Matrix, welche die R -lineare Abbildung ψ bezüglich geeigneter Basen in R^m und R^n beschreibt, so bezeichnen wir die Elementarteiler von A auch als die *Elementarteiler der betrachteten endlichen Präsentation* von M . Nach 6.3/8 stimmen diese mit den Elementarteilern von $\ker \varphi \subset R^m$ überein und hängen daher nicht von der Auswahl der Basen in R^m und R^n ab.

Wir wollen nun aus dem Elementarteilersatz 6.3/4 verschiedene Versionen des so genannten Hauptsatzes über endlich erzeugte Moduln über Hauptidealringen herleiten. Wir beginnen mit einer grundlegenden Folgerung aus dem Elementarteilersatz.

Satz 2. *Es sei M ein endlich erzeugter Modul über einem Hauptidealring R und $T \subset M$ sein Torsionsuntermodul. Dann gibt es einen endlichen freien Untermodul $F \subset M$, etwa $F \simeq R^d$, sowie Nichteinheiten $\alpha_1, \dots, \alpha_s \in R - \{0\}$ mit $\alpha_j \mid \alpha_{j+1}$ für $1 \leq j < s$ und*

$$M = F \oplus T, \quad T \simeq \bigoplus_{j=1}^s R/\alpha_j R.$$

Dabei ist d eindeutig bestimmt; man bezeichnet d als den Rang von M . Weiter sind die Elemente $\alpha_1, \dots, \alpha_s$ eindeutig bestimmt bis auf Assoziiertheit. Sie stimmen überein mit den Nichteinheiten unter den Elementarteilern einer jeden endlichen Präsentation von M .

Beweis. Wir gehen von einer endlichen Präsentation von M aus und betrachten den zugehörigen Epimorphismus $\varphi: R^m \longrightarrow M$. Dann folgt $M \simeq R^m / \ker \varphi$ aufgrund des Homomorphiesatzes. Auf die Situation $\ker \varphi \subset R^m$ können wir nun den Elementarteilersatz 6.3/4 anwenden (wobei wir berücksichtigen, dass jede Basis von R^m gemäß 6.3/7 aus genau m Elementen besteht). Es existieren daher Elemente $x_1, \dots, x_m$, die eine Basis von R^m bilden, sowie Elemente $\alpha_1, \dots, \alpha_s \in R - \{0\}$, $s \leq m$, mit $\alpha_j \mid \alpha_{j+1}$ für $1 \leq j < s$, so dass $\alpha_1 x_1, \dots, \alpha_s x_s$ eine Basis von $\ker \varphi$ bilden. Indem wir $\alpha_{s+1} = \dots = \alpha_m = 0$ setzen, können wir den Epimorphismus

$$\varphi': R^m = \bigoplus_{j=1}^m R x_j \longrightarrow \bigoplus_{j=1}^m R x_j / R \alpha_j x_j, \quad (\gamma_1, \dots, \gamma_m) \longmapsto (\overline{\gamma}_1, \dots, \overline{\gamma}_m),$$

betrachten, wobei $\overline{\gamma}_j$ jeweils die Restklasse von γ_j in $R x_j / R \alpha_j x_j$ bezeichne. Nach Konstruktion gilt $\ker \varphi' = \ker \varphi$ und folglich aufgrund des Homomorphiesatzes

$$M \simeq R^m / \ker \varphi = \left(\bigoplus_{j=1}^m Rx_j \right) / \ker \varphi' \simeq \bigoplus_{j=1}^m Rx_j / R\alpha_j x_j \simeq R^{m-s} \oplus \bigoplus_{j=1}^s R / \alpha_j R,$$

also mit $d := m - s$ eine Zerlegung des behaupteten Typs, wenn wir Summanden $R / \alpha_j R = 0$, d. h. mit $\alpha_j \in R^*$ unterdrücken.

In dieser Zerlegung korrespondiert $\bigoplus_{j=1}^s R / \alpha_j R$ zu dem Torsionsuntermodul $T \subset M$ und ist daher eindeutig bestimmt. Indem wir triviale Summanden $R / \alpha_j R$ ignorieren, bzw. annehmen, dass $\alpha_1, \dots, \alpha_s$ keine Einheiten sind, ergibt sich die Eindeutigkeit der α_j mit 6.3/6. Insbesondere sind die α_j unabhängig von der betrachteten Präsentation von M und bestehen gemäß 6.3/8 aus den Nichteinheiten unter den Elementarteilern einer solchen Präsentation.

Um zu sehen, dass auch d eindeutig ist, betrachte man den Epimorphismus

$$M \xrightarrow{\sim} R^d \oplus \bigoplus_{j=1}^s R / \alpha_j R \longrightarrow R^d,$$

der sich aus dem obigen Isomorphismus und der Projektion auf den Summanden R^d zusammensetzt. Der Kern dieser Abbildung ist offenbar gerade der Torsionsuntermodul $T \subset M$, so dass man aufgrund des Homomorphiesatzes einen Isomorphismus $M/T \xrightarrow{\sim} R^d$ erhält. Hieraus folgt die Eindeutigkeit von d mit 6.3/7. $\square$

Speziellere Versionen des Hauptsatzes über endlich erzeugte Moduln über Hauptidealringen lassen sich mittels des Chinesischen Restsatzes aus Satz 2 folgern.

Satz 3. *Es sei M ein endlich erzeugter Modul über einem Hauptidealring R und $T \subset M$ der Torsionsuntermodul. Weiter sei $P \subset R$ ein Vertretersystem der Primelemente von R , und für $p \in P$ bezeichne*

$$M_p = \{x \in M ; p^n x = 0 \text{ für geeignetes } n \in \mathbb{N}\}$$

den so genannten Untermodul der p -Torsion in M . Dann gilt

$$T = \bigoplus_{p \in P} M_p,$$

und es gibt einen endlichen freien Untermodul $F \subset M$, etwa $F \simeq R^d$, mit

$$M = F \oplus T,$$

wobei d eindeutig bestimmt ist und M_p für fast alle $p \in P$ verschwindet. Weiter gibt es zu jedem $p \in P$ natürliche Zahlen $1 \leq n(p, 1) \leq \dots \leq n(p, s_p)$ mit

$$M_p \simeq \bigoplus_{j_p=1 \dots s_p} R / p^{n(p, j_p)} R,$$

wobei die $s_p, n(p, j_p)$ durch die Isomorphie

$$M \simeq F \oplus \bigoplus_{p \in P} \bigoplus_{j_p=1 \dots s_p} R/p^{n(p,j_p)} R$$

eindeutig bestimmt sind und $s_p = 0$ für fast alle p gilt.

Bevor wir zum Beweis kommen, wollen wir diesen Hauptsatz auch noch speziell für endlich erzeugte $\mathbb{Z}$ -Moduln formulieren, als *Hauptsatz über endlich erzeugte abelsche Gruppen*.

Korollar 4. *Es sei G eine endlich erzeugte abelsche Gruppe, P sei die Menge der Primzahlen in $\mathbb{N}$. Dann gestattet G eine Zerlegung in Untergruppen*

$$G = F \oplus \bigoplus_{p \in P} \bigoplus_{j_p=1 \dots s_p} G_{p,j_p},$$

wobei F frei ist, etwa $F \simeq \mathbb{Z}^d$, und G_{p,j_p} zyklisch von p -Potenz-Ordnung, etwa $G_{p,j_p} \simeq \mathbb{Z}/p^{n(p,j_p)}\mathbb{Z}$ mit $1 \leq n(p,1) \leq \dots \leq n(p,s_p)$. Die Zahlen $d, s_p, n(p,j_p)$ sind eindeutig durch G bestimmt, ebenso die Untergruppen $G_p = \bigoplus_{j_p=1 \dots s_p} G_{p,j_p}$, wobei s_p für fast alle $p \in P$ verschwindet.

Wenn G eine endlich erzeugte Torsionsgruppe ist, also ein über $\mathbb{Z}$ endlich erzeugter Torsionsmodul, so besitzt G keinen freien Anteil und besteht daher, wie man insbesondere mit Korollar 4 sieht, nur aus endlich vielen Elementen. Umgekehrt ist jede endliche abelsche Gruppe natürlich eine endlich erzeugte Torsionsgruppe.

Nun zum *Beweis von Satz 3*. Wir beginnen mit der Zerlegung

$$M \simeq R^d \oplus \bigoplus_{j=1}^s R/\alpha_j R$$

aus Satz 2, wobei insbesondere d eindeutig bestimmt ist. Hierbei korrespondiert $\bigoplus_{j=1}^s R/\alpha_j R$ zu dem Torsionsuntermodul $T \subset M$, sowie R^d zu einem freien Modul $F \subset M$, und es gilt $M = F \oplus T$.

Man zerlege die α_j nun in Primfaktoren, etwa $\alpha_j = \varepsilon_j p_1^{n(1,j)} \dots p_r^{n(r,j)}$ mit Einheiten $\varepsilon_j \in R^*$, paarweise verschiedenen Primelementen $p_1, \dots, p_r \in P$, sowie Exponenten $n(i,j) \in \mathbb{N}$, die auch trivial sein dürfen. Dann gilt aufgrund von Satz 1

$$T \simeq \bigoplus_{j=1}^s \bigoplus_{i=1}^r R/p_i^{n(i,j)} R \simeq \bigoplus_{i=1}^r \bigoplus_{j=1}^s R/p_i^{n(i,j)} R.$$

In dieser Zerlegung korrespondiert $\bigoplus_{j=1}^s R/p_i^{n(i,j)} R$ offenbar gerade zu dem Untermodul $M_{p_i} \subset M$ der p_i -Torsion und ist deshalb eindeutig bestimmt; die Restklasse von p_i ist nämlich in jedem Restklassenring der Form $R/p_{i'}^n R$ mit $i' \neq i$ eine Einheit. Somit folgt aus obiger Zerlegung insbesondere $T = \bigoplus_{p \in P} M_p$. Da im Übrigen in der Zerlegung

$$M_{p_i} \simeq \bigoplus_{j=1}^s R/p_i^{n(i,j)} R$$

die Terme zu Exponenten $n(i, j) = 0$ trivial sind und die restlichen Terme gemäß 6.3/6 eindeutig bestimmt sind, ergibt sich insgesamt die Behauptung von Satz 3. $\square$

Aufgaben

Für eine Gruppe G bezeichnet man mit $\text{ord } G$ die Anzahl ihrer Elemente und nennt dies die *Ordnung* von G .

1. Es sei G eine abelsche Gruppe der Ordnung $n < \infty$. Man zeige: Zu jedem Teiler d von n gibt es eine Untergruppe $H \subset G$ der Ordnung d . Andererseits besitzt jede Untergruppe $H \subset G$ eine Ordnung, die ein Teiler von n ist.
2. Zu $n \in \mathbb{N}$ gibt es höchstens endlich viele Klassen isomorpher abelscher Gruppen der Ordnung n . Wie groß ist diese Anzahl für $n = 20$ bzw. $n = 30$?
3. Es seien $a_1, \dots, a_n \in \mathbb{N} - \{0\}$ paarweise teilerfremd und $r_1, \dots, r_n \in \mathbb{N}$ Zahlen mit $0 \leq r_i < a_i$ für $i = 1, \dots, n$. Man zeige: Es existiert eine Zahl $a \in \mathbb{N}$, so dass a bei ganzzahliger Division durch a_i jeweils den Rest r_i lässt.
4. Es sei M ein endlich erzeugter Modul über einem Hauptidealring R . Für den Fall, dass M kein Torsionsmodul ist, zeige man: Es existieren endlich viele freie Untermoduln $F_1, \dots, F_r \subset M$ mit $M = \sum_{i=1}^r F_i$.
5. Über einem Hauptidealring R betrachte man endliche freie Moduln F und F' mit Untermoduln $M \subset F$ und $M' \subset F'$, so dass die Restklassenmoduln F/M und F'/M' jeweils Torsionsmoduln sind. Man zeige: Es existiert genau dann ein Isomorphismus $F/M \xrightarrow{\sim} F'/M'$, wenn die Nichteinheiten unter den Elementarteilern von $M \subset F$ mit denen von $M' \subset F'$ übereinstimmen.
6. Es sei K ein Körper und $G \subset K^*$ eine endliche Untergruppe der multiplikativen Gruppe der Einheiten von K . Man zeige, dass G *zyklisch* ist, d. h. dass es ein $n \in \mathbb{Z}$ mit $G \simeq \mathbb{Z}/n\mathbb{Z}$ gibt. Hinweis: Man betrachte Nullstellen aus K von Polynomen des Typs $T^n - 1 \in K[T]$.

6.5 Allgemeine und Jordansche Normalform für Matrizen

Im Folgenden sei V stets ein Vektorraum über einem Körper K und $f: V \longrightarrow V$ ein Endomorphismus. Wie wir in Abschnitt 6.3 gesehen haben, lässt sich V als Modul über dem Polynomring $K[T]$ auffassen, indem man die Multiplikation der Variablen T mit Elementen von V durch Anwenden von f erklärt. Ist V von endlicher Dimension, so existiert ein eindeutig bestimmtes normiertes Polynom minimalen Grades $p_f \in K[T]$ mit $p_f \cdot v = 0$ für alle $v \in V$; man nennt p_f das *Minimalpolynom* zu f , vgl. 6.2/9. Es ist V dann insbesondere ein $K[T]$ -Torsionsmodul, sogar ein endlich erzeugter, da V als K -Vektorraum

endlich erzeugt ist. Wir können daher das Strukturresultat 6.4/3 auf V als $K[T]$ -Modul anwenden.

Definition 1. Es sei $f: V \longrightarrow V$ ein Endomorphismus eines K -Vektorraums V , sowie $U \subset V$ ein Untervektorraum.

- (i) U heißt f -invariant, wenn $f(U) \subset U$ gilt.
- (ii) U heißt f -zyklisch, wenn es einen Vektor $u \in U$ gibt, so dass die Folge $u, f(u), f^2(u), \dots$ ein Erzeugendensystem von U bildet; insbesondere ist U dann auch f -invariant.
- (iii) U heißt f -unzerlegbar, wenn U ein f -invarianter Untervektorraum $\neq 0$ ist, der sich nicht in eine direkte Summe zweier echter f -invarianter Untervektorräume zerlegen lässt.

Wir wollen die gerade eingeführten Eigenschaften von Unterräumen nun auch modultheoretisch interpretieren. Ein Modul M über einem Ring R wird als *monogen* bezeichnet, wenn er von einem einzigen Element erzeugt wird, d. h. wenn ein $a \in M$ existiert mit $M = Ra$.

Bemerkung 2. Es sei $f: V \longrightarrow V$ ein Endomorphismus eines K -Vektorraums V , wobei man V insbesondere auch als $K[T]$ -Modul bezüglich f betrachte. Für eine Teilmenge $U \subset V$ gilt dann:

- (i) U ist genau dann ein f -invarianter Untervektorraum von V , wenn U ein $K[T]$ -Untermodul von V ist.
- (ii) U ist genau dann ein f -zyklischer Untervektorraum von V , wenn U ein monogener $K[T]$ -Untermodul von V ist.
- (iii) U ist genau dann ein f -unzerlegbarer Untervektorraum von V , wenn U ein $K[T]$ -Untermodul $\neq 0$ von V ist, der sich nicht in eine direkte Summe zweier echter $K[T]$ -Untermoduln zerlegen lässt.

Beweis. Aussage (i) ist unmittelbar klar, denn für Untervektorräume $U \subset V$ bedeutet die Bedingung $f(u) \in U$ für $u \in U$ gerade $T \cdot u \in U$ und ist daher äquivalent zu $q \cdot u \in U$ für $q \in K[T]$, $u \in U$. In ähnlicher Weise erhält man Aussage (ii), denn $U = \sum_{i \in \mathbb{N}} K \cdot f^i(u)$ für ein $u \in U$ ist gleichbedeutend mit $U = \sum_{i \in \mathbb{N}} K \cdot T^i \cdot u$, also mit $U = K[T] \cdot u$. Aussage (iii) schließlich beruht auf der Tatsache, dass für f -invariante Untervektorräume bzw. $K[T]$ -Untermoduln $U', U'' \subset V$ die Summe $U = U' + U''$ genau dann direkt ist, wenn jedes $u \in U$ eine eindeutig bestimmte Darstellung $u = u' + u''$ mit $u' \in U'$, $u'' \in U''$ besitzt. Ob wir bei der skalaren Multiplikation Elemente aus K oder $K[T]$ zulassen, ist dabei ohne Belang. $\square$

Im Weiteren ist es auch wichtig zu wissen, wie man Homomorphismen von $K[T]$ -Moduln mittels Homomorphismen der unterliegenden K -Vektorräume interpretieren kann.

Bemerkung 3. Es seien V und V' zwei K -Vektorräume, welche man bezüglich gegebener Endomorphismen $f: V \rightarrow V$ und $f': V' \rightarrow V'$ als $K[T]$ -Moduln betrachte. Für Abbildungen $\varphi: V \rightarrow V'$ ist dann äquivalent:

- (i) φ ist ein Homomorphismus von $K[T]$ -Moduln.
- (ii) φ ist ein Homomorphismus von K -Vektorräumen, und das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{f} & V, \\ \downarrow \varphi & & \downarrow \varphi \\ V' & \xrightarrow{f'} & V', \end{array} \quad \begin{array}{l} v \mapsto f(v) = T \cdot v, \\ v' \mapsto f'(v') = T \cdot v', \end{array}$$

ist kommutativ.

Eine entsprechende Äquivalenz gilt auch für Mono-, Epi- und Isomorphismen $\varphi: V \rightarrow V'$. Insbesondere stimmen in der Situation eines Isomorphismus φ in (i) bzw. (ii) die Minimalpolynome von f und f' überein. Gleiches gilt für die charakteristischen Polynome von f und f' .

Beweis. Jeder Homomorphismus von $K[T]$ -Moduln $\varphi: V \rightarrow V'$ ist $K[T]$ -linear und damit insbesondere K -linear, also ein Homomorphismus von K -Vektorräumen. Zudem bedeutet die Kommutativität des Diagramms in (ii) gerade die T -Linearität von φ , womit wir

$$\varphi(T \cdot v) = T \cdot \varphi(v) \quad \text{für alle } v \in V$$

meinen. Somit ergibt sich (ii) aus (i) und umgekehrt (i) auch aus (ii), wenn man benutzt, dass aus der T -Linearität die T^n -Linearität von φ für alle $n \in \mathbb{N}$ folgt, und aufgrund der Additivität sogar die $K[T]$ -Linearität von φ . Entsprechend schließt man im Falle von Mono-, Epi- oder Isomorphismen. $\square$

Satz 4. Es sei $f: V \rightarrow V$ ein Endomorphismus eines K -Vektorraums V der Dimension $r < \infty$. Man betrachte V als $K[T]$ -Modul bezüglich f . Dann ist äquivalent:

- (i) V ist f -zyklisch.
- (ii) V ist als $K[T]$ -Modul monogen.
- (iii) Es gibt ein normiertes Polynom $p \in K[T]$, so dass V als $K[T]$ -Modul isomorph zu $K[T]/(p)$ ist.

Sind diese Bedingungen erfüllt, so gilt weiter:

- (iv) Das Polynom p aus (iii) stimmt mit dem Minimalpolynom p_f sowie dem charakteristischen Polynom χ_f von f überein; insbesondere gilt $\text{grad } p = r$.

Beweis. Die Bedingungen (i) und (ii) sind aufgrund von Bemerkung 2 äquivalent. Wird weiter V gemäß Bedingung (ii) als $K[T]$ -Modul von dem Element $u \in V$ erzeugt, so betrachte man die $K[T]$ -lineare Abbildung

$$\varphi: K[T] \rightarrow V, \quad h \mapsto h \cdot u = h(f)(u);$$

φ ist surjektiv. Weiter ist $\ker \varphi$ ein $K[T]$ -Untermodul von $K[T]$, also ein Ideal, welches, da $K[T]$ ein Hauptidealring ist, von einem Polynom $p \in K[T]$ erzeugt wird. Der Homomorphiesatz für $K[T]$ -Moduln ergibt dann den in (iii) gewünschten Isomorphismus $K[T]/(p) \simeq V$. Wegen $\dim_K V < \infty$ und $\dim_K K[T] = \infty$ gilt $p \neq 0$, und man darf p als normiertes Polynom annehmen.

Ist andererseits Bedingung (iii) gegeben, so betrachte man die Restklassenabbildung $K[T] \rightarrow K[T]/(p)$, $h \mapsto \bar{h}$. Die Restklasse $\bar{1} \in K[T]/(p)$ erzeugt dann natürlich $K[T]/(p)$ als $K[T]$ -Modul, d. h. es folgt (ii).

Es bleibt nun noch Aussage (iv) zu begründen, wobei wir (i) bzw. (iii) als gegeben annehmen. Wir lesen zunächst aus (iii) ab, dass p das Minimalpolynom von

$$f': K[T]/(p) \rightarrow K[T]/(p), \quad \bar{g} \mapsto T \cdot \bar{g} = \bar{T} \cdot \bar{g},$$

ist; dabei bezeichne $\bar{g}$ für $g \in K[T]$ jeweils die Restklasse in $K[T]/(p)$, entsprechend für $\bar{T}$. Das Minimalpolynom von f' erzeugt nämlich gemäß 6.2/9 bzw. 6.2/10 das Ideal

$$\begin{aligned} & \{h \in K[T] ; h(f')(\bar{g}) = 0 \text{ für alle } \bar{g} \in K[T]/(p)\} \\ &= \{h \in K[T] ; h \cdot \bar{g} = 0 \text{ für alle } \bar{g} \in K[T]/(p)\} \\ &= \{h \in K[T] ; h \cdot \bar{1} = 0\} = \{h \in K[T] ; \bar{h} = 0\} \\ &= (p) \subset K[T]. \end{aligned}$$

Dann stimmt p aber gemäß Bemerkung 3 auch mit dem Minimalpolynom p_f des Endomorphismus $f: V \rightarrow V$ überein.

Dass $p = p_f$ gilt, können wir in etwas ausführlicherer Argumentation auch auf dem Niveau des Endomorphismus f und des Vektorraums V einsehen: Es erzeugt p_f gemäß 6.2/9 bzw. 6.2/10 dasjenige Ideal in $K[T]$, welches aus allen Polynomen h mit $h(f) = 0$ besteht. Also gilt

$$(*) \quad (p_f) = \{h \in K[T] ; h(f)(v) = 0 \text{ für alle } v \in V\},$$

wobei es genügt, wenn man die Bedingung $h(f)(v) = 0$ für v aus einem K -Erzeugendensystem von V testet. Ist daher V als $K[T]$ -Modul monogen, etwa erzeugt von einem Element $u \in V$, so bilden die Elemente $u, f(u), f^2(u), \dots$ ein K -Erzeugendensystem von V , und die Beschreibung $(*)$ ist äquivalent zu

$$(**) \quad (p_f) = \{h \in K[T] ; h(f)(f^n(u)) = 0 \text{ für alle } n \in \mathbb{N}\}.$$

Nun zeigt aber die Rechnung

$$h(f)(f^n(u)) = h \cdot f^n(u) = h \cdot T^n \cdot u = T^n \cdot h \cdot u = T^n \cdot h(f)(u),$$

dass aus $h(f)(u) = 0$ bereits $h(f)(f^n(u)) = 0$ für alle $n \in \mathbb{N}$ folgt. Somit erhält man aus $(**)$

$$(p_f) = \{h \in K[T] ; h(f)(u) = 0\} = \{h \in K[T] ; h \cdot u = 0\} = (p)$$

und damit wegen der Normiertheit von p_f und p bereits $p_f = p$.

Insbesondere stimmt der Grad von p_f aufgrund des nachfolgenden Lemmas 5 mit der K -Vektorraumdimension von $K[T]/(p)$ überein, also mit der Dimension r von V . Da aber das charakteristische Polynom χ_f normiert vom Grad $\dim_K V$ ist und da aufgrund des Satzes von Cayley-Hamilton p_f ein Teiler von χ_f ist, vgl. 6.2/11, gilt bereits $\chi_f = p_f$. Die Gleichung $\chi_f = p$ lässt sich allerdings auch durch explizite Berechnung nachweisen, indem man die Multiplikation mit T auf $K[T]/(p)$ bezüglich der Basis $\bar{T}^0, \dots, \bar{T}^{r-1}$ durch eine Matrix beschreibt und deren charakteristisches Polynom ermittelt. $\square$

Schließlich bleibt noch die K -Vektorraumdimension von Quotienten des Typs $K[T]/(p)$ zu berechnen.

Lemma 5. *Es sei $p \in K[T]$ ein normiertes Polynom. Dann gilt*

$$\dim_K(K[T]/(p)) = \text{grad } p,$$

und die Restklassen $\bar{T}^0, \dots, \bar{T}^{s-1}$ mit $s = \text{grad } p$ bilden eine K -Basis von $K[T]/(p)$.

Beweis. Es genügt zu zeigen, dass für $s = \text{grad } p$ die Restklassen $\bar{T}^0, \dots, \bar{T}^{s-1}$ eine K -Basis von $K[T]/(p)$ bilden. Hierzu betrachte man ein beliebiges Element $\bar{h} \in K[T]/(p)$ und wähle ein Urbild $h \in K[T]$. Division mit Rest ergibt dann eine Zerlegung $h = qp + c$ mit Polynomen $q, c \in K[T]$, wobei c vom Grad $< s$ ist, etwa $c = \sum_{i=0}^{s-1} c_i T^i$ mit Koeffizienten $c_i \in K$. Es folgt

$$\bar{h} = \bar{c} = \sum_{i=0}^{s-1} c_i \bar{T}^i,$$

d. h. $K[T]/(p)$ wird von den Restklassen $\bar{T}^0, \dots, \bar{T}^{s-1}$ als K -Vektorraum erzeugt. Diese Elemente sind aber auch linear unabhängig über K . Hat man nämlich $\sum_{i=0}^{s-1} c_i \bar{T}^i = 0$ mit gewissen Koeffizienten $c_i \in K$, so bedeutet dies $\sum_{i=0}^{s-1} c_i T^i \in (p)$ und damit $c_i = 0$ für alle $i = 0, \dots, s-1$, da alle von Null verschiedenen Elemente in (p) einen Grad $\geq \text{grad } p = s$ haben. $\square$

Wir wollen nun den Struktursatz 6.4/3 umformulieren zu einem Struktursatz für $K[T]$ -Moduln oder, genauer, für die Situation eines endlich-dimensionalen K -Vektorraums V mit einem Endomorphismus $f: V \rightarrow V$.

Theorem 6. *Es sei V ein endlich-dimensionaler K -Vektorraum mit einem Endomorphismus $f: V \rightarrow V$. Dann existieren paarweise verschiedene normierte (und damit nicht-assozierte) Primpolynome $p_1, \dots, p_r \in K[T]$ sowie natürliche Zahlen $1 \leq n(i, 1) \leq \dots \leq n(i, s_i)$ und f -zyklische K -Untervektorräume $V_{ij} \subset V$, $i = 1, \dots, r$, $j = 1, \dots, s_i$, mit*

$$V = \bigoplus_{i=1}^r \bigoplus_{j=1}^{s_i} V_{ij}, \quad V_{ij} \simeq K[T]/(p_i^{n(i,j)}),$$

im Sinne von $K[T]$ -Moduln. Insbesondere ist $p_i^{n(i,j)}$ das Minimal- bzw. charakteristische Polynom von $f|_{V_{ij}}$; vgl. Satz 4.

In der vorstehenden Zerlegung sind die Primpolynome $p_1, \dots, p_r$, die Zahlen $n(i, j)$ und die Vektorräume $V_i := \bigoplus_{j=1}^{s_i} V_{ij}$ für $i = 1, \dots, r$ eindeutig bestimmt, nicht notwendig jedoch die Unterräume V_{ij} selbst; es gilt

$$V_i = \bigcup_{n \in \mathbb{N}} \ker p_i^n(f).$$

Weiter ist $p_f = p_1^{n(1,s_1)} \cdot \dots \cdot p_r^{n(r,s_r)}$ das Minimalpolynom von f , und man hat

$$\dim_K V_{ij} = n(i, j) \cdot \text{grad } p_i,$$

insbesondere also

$$\dim_K V = \sum_{i=1}^r \sum_{j=1}^{s_i} n(i, j) \cdot \text{grad } p_i.$$

Beweis. Die Aussagen zur Existenz und Eindeutigkeit einer Zerlegung von V in f -zyklische Unterräume V_{ij} der behaupteten Art ergeben sich unmittelbar aus 6.4/3 und Satz 4. Dabei ist $V_i = \bigoplus_{j=1}^{s_i} V_{ij}$ in der Sprache der $K[T]$ -Moduln der Untermodul der p_i -Torsion, und der freie Anteil aus 6.4/3 entfällt, da V ein $K[T]$ -Torsionsmodul ist.

Gemäß Satz 4 ist $p_i^{n(i,j)}$ das Minimalpolynom der Einschränkung von f auf V_{ij} . Dann erkennt man leicht, dass das kleinste gemeinsame Vielfache aller dieser Polynome, also

$$p_1^{n(1,s_1)} \cdot \dots \cdot p_r^{n(r,s_r)}$$

das Minimalpolynom von f auf V ist. Die Dimensionsformeln schließlich ergeben sich unter Benutzung von 1.6/6 aus Lemma 5. $\square$

Korollar 7. *Es sei $f: V \rightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V . Dann ist äquivalent:*

- (i) V ist f -unzerlegbar.
- (ii) V ist f -zyklisch, und $p_f = \chi_f$ ist Potenz eines Primpolynoms in $K[T]$.

Insbesondere sind die f -zyklischen K -Untervektorräume $V_{ij} \subset V$ aus Theorem 6 f -unzerlegbar.

Beweis. Man zerlege V gemäß Theorem 6 in eine direkte Summe f -zyklischer Unterräume, deren Minimalpolynom jeweils Potenz eines Primpolynoms aus $K[T]$ ist. Ist V dann f -unzerlegbar, so kann diese Zerlegung nur aus einem Summanden bestehen, und es folgt, dass V f -zyklisch ist.

Ist andererseits V f -zyklisch mit einem Minimalpolynom p_f , welches Potenz eines Primpolynoms in $K[T]$ ist, so gilt $V \simeq K[T]/(p_f)$ im Sinne von $K[T]$ -Moduln gemäß Satz 4, und die Eindeutigkeit der Zerlegung in Theorem 6 zeigt, dass V f -unzerlegbar ist. $\square$

Schließlich wollen wir das Zerlegungstheorem 6 noch umformulieren zu einer Aussage über Normalformen von quadratischen Matrizen. Wir beginnen mit einer trivialen Beobachtung, wobei wir folgende Schreibweise benutzen: Für Matrizen $A_i \in K^{n_i \times n_i}$, $i = 1, \dots, r$, bezeichne $\text{Diag}(A_1, \dots, A_r) \in K^{n \times n}$, $n = \sum_{i=1}^r n_i$, die "Diagonalmatrix" mit den Einträgen $A_1, \dots, A_r$, also

$$\text{Diag}(A_1, \dots, A_r) = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \dots & \\ 0 & & & \dots & A_r \end{pmatrix}$$

Gilt $A_i = (\lambda_i) \in K^{1 \times 1}$ für $i = 1, \dots, r$, so ist $\text{Diag}(A_1, \dots, A_r)$ eine Diagonalmatrix im echten Sinne, und wir schreiben auch $\text{Diag}(\lambda_1, \dots, \lambda_r)$ anstelle von $\text{Diag}(A_1, \dots, A_r)$.

Lemma 8. *Es sei $f: V \rightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V mit einer Zerlegung $V = \bigoplus_{i=1}^r V_i$ in f -invariante Unterräume. Es sei X_i jeweils eine Basis von V_i , sowie X die Basis von V , die sich aus den Basen $X_1, \dots, X_r$ zusammensetzt. Mit $f_i = f|_{V_i}$ gilt dann für die zu f gehörige Matrix sowie das charakteristische bzw. Minimalpolynom:*

- (i) $A_{f,X,X} = \text{Diag}(A_{f_1,X_1,X_1}, \dots, A_{f_r,X_r,X_r})$.
- (ii) $\chi_f = \chi_{f_1} \cdots \chi_{f_r}$.
- (iii) $p_f = \text{kgV}(p_{f_1}, \dots, p_{f_r})$.

Lemma 9. *Es sei $f: V \rightarrow V$ ein Endomorphismus eines K -Vektorraums V mit einer Basis $X = (x_1, \dots, x_n)$. Gilt dann $A_{f,X,X} = \text{Diag}(A_1, \dots, A_r)$ mit Matrizen $A_i \in K^{n_i \times n_i}$ (wobei $n = \sum_{i=1}^r n_i$), so sind für $m_i = \sum_{j=1}^i n_j$ die Unterräume*

$$V_i = \sum_{j=m_{i-1}+1}^{m_i} K \cdot x_j, \quad i = 1, \dots, r,$$

f -invariant, und es gilt $V = \bigoplus_{i=1}^r V_i$.

Beweis zu den Lemmata 8 und 9. Aussage (i) in Lemma 8 sowie Lemma 9 sind unmittelbar klar aufgrund der in 3.1/3 definierten Korrespondenz zwischen linearen Abbildungen und beschreibenden Matrizen. Weiter folgt Aussage (ii) in Lemma 8 aufgrund des Beispiels (2) am Ende von Abschnitt 4.3. Aussage (iii) schließlich ist gültig, da einerseits das kleinste gemeinsame Vielfache aller p_{f_i} den Endomorphismus f annulliert, sowie andererseits p_f ein Vielfaches eines jeden p_{f_i} sein muss. $\square$

Zu einem normierten Polynom $p = T^n + \sum_{i=0}^{n-1} c_i T^i \in K[T]$ vom Grad n nennt man die Matrix

$$A(p) = \begin{pmatrix} 0 & & & & -c_0 \\ 1 & 0 & & & -c_1 \\ & 1 & . & & -c_2 \\ & & . & . & \dots \\ & & & . & \dots \\ & & & & 0 & -c_{n-2} \\ & & & & 1 & -c_{n-1} \end{pmatrix}$$

die Begleitmatrix zu p . Insbesondere gilt $A(p) = (-c_0)$ für $n = 1$. Auch den Sonderfall $n = 0$ wollen wir nicht ausschließen. Es ist $A(p)$ dann die leere Matrix.

Lemma 10. *Es sei $f: V \longrightarrow V$ ein Endomorphismus eines K -Vektorraums der Dimension $n < \infty$ und $p \in K[T]$ ein normiertes Polynom vom Grad n . Dann ist äquivalent:*

- (i) V ist f -zyklisch mit Minimalpolynom $p_f = p$.
- (ii) Es existiert eine Basis X von V , so dass die Matrix $A_{f,X,X}$ die Begleitmatrix zu p ist.

Beweis. Für $V = 0$ ist die Aussage trivial, so dass wir $V \neq 0$ annehmen dürfen. Sei V zunächst f -zyklisch, gelte also $V \simeq K[T]/(p)$ im Sinne von $K[T]$ -Moduln, wobei $p = p_f$ das Minimalpolynom von f ist; vgl. Satz 4. Dann ergeben die Restklassen $\overline{T}^0, \dots, \overline{T}^{n-1}$ gemäß Lemma 5 eine K -Basis von $K[T]/(p)$, bzw. mit anderen Worten, es existiert ein Element $u \in V$ (nämlich dasjenige Element, welches unter dem Isomorphismus $V \simeq K[T]/(p)$ zu der Restklasse $\overline{1} \in K[T]/(p)$ korrespondiert), so dass $f^0(u), f^1(u), \dots, f^{n-1}(u)$ eine K -Basis X von V bilden. Für $p = T^n + c_{n-1}T^{n-1} + \dots + c_0$ folgt dann

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

und man sieht, dass $A_{f,X,X}$ die Begleitmatrix zu p ist.

Sei nun umgekehrt $X = (x_1, \dots, x_n)$ eine Basis von V , so dass $A_{f,X,X}$ die Begleitmatrix zu einem normierten Polynom $p \in K[T]$ vom Grade n ist. Setzen wir dann $u = x_1$, so gilt $x_i = f^{i-1}(u)$ für $i = 1, \dots, n$, und es folgt, dass V f -zyklisch ist. Weiter liest man aus der Matrix $A_{f,X,X}$ die Beziehungen $p(f)(u) = 0$ bzw. $p \cdot u = 0$ im Sinne von V als $K[T]$ -Modul ab. Da man aber $V = K[T] \cdot u$ hat, ergibt sich auch $p \cdot V = 0$, so dass das Minimalpolynom p_f ein Teiler von p ist. Da p_f gemäß Satz 4 mit dem charakteristischen Polynom χ_f übereinstimmt und damit ebenso wie p den Grad n hat, erhält man damit bereits $\chi_f = p_f = p$. $\square$

Nun sind wir in der Lage, aus Theorem 6 die Existenz und Eindeutigkeit der so genannten *allgemeinen Normalform* für Matrizen herzuleiten.

Theorem 11. *Jede Matrix $A \in K^{n \times n}$ ist ähnlich zu einer Matrix der Form $\text{Diag}(A_1, \dots, A_r)$ mit quadratischen Matrizen A_i , wobei A_i jeweils die Begleitmatrix zu einer Potenz q_i eines normierten Primpolynoms aus $K[T]$ ist. Die*

Matrizen $A_1, \dots, A_r$ sind bis auf die Reihenfolge eindeutig durch A bestimmt. Es gilt $p_A = \text{kgV}(q_1, \dots, q_r)$ für das Minimalpolynom von A , sowie $\chi_A = q_1 \cdot \dots \cdot q_r$ für das charakteristische Polynom von A .

Beweis. Interpretieren wir A als Matrix eines Endomorphismus $f: V \rightarrow V$ eines endlich-dimensionalen K -Vektorraums V , so zerfällt V nach Theorem 6 in eine direkte Summe f -zyklischer Unterräume $V_1, \dots, V_r$, etwa $V_i \simeq K[T]/(q_i)$, wobei q_i Potenz eines normierten Primpolynoms ist. Nach Lemma 8 zeigt dies, dass A zu einer Diagonalmatrix $\text{Diag}(A_1, \dots, A_r)$ ähnlich ist, wobei A_i jeweils den Endomorphismus $f_i = f|_{V_i}$ bezüglich einer geeigneten Basis von V_i beschreibt. Aus Satz 4 ergibt sich, dass q_i das Minimalpolynom bzw. das charakteristische Polynom von f_i ist, und aus Lemma 10, dass wir A_i als Begleitmatrix zu q_i annehmen dürfen.

Ist umgekehrt A zu einer Matrix der Form $\text{Diag}(A_1, \dots, A_r)$ ähnlich, wobei A_i jeweils die Begleitmatrix zu einer Potenz q_i eines normierten Primpolynoms ist, so korrespondiert hierzu eine Zerlegung $V \simeq \bigoplus_{i=1}^r K[T]/(q_i)$; vgl. Lemmata 9 und 10 in Verbindung mit Satz 4. Die Eindeutigkeitsaussage in Theorem 6 impliziert dann die Eindeutigkeit der Matrizen $A_1, \dots, A_r$, die ja durch ihre Minimalpolynome bzw. charakteristischen Polynome eindeutig festgelegt sind.

Die Aussagen über das Minimalpolynom p_A sowie das charakteristische Polynom χ_A wurden bereits in Lemma 8 hergeleitet. $\square$

Als Folgerung können wir nochmals den Satz von Cayley-Hamilton ablesen, sogar in einer etwas verbesserten Version:

Korollar 12 (Cayley-Hamilton). *Es sei $f: V \rightarrow V$ ein Endomorphismus eines K -Vektorraums der Dimension $n < \infty$ mit Minimalpolynom p_f und charakteristischem Polynom χ_f . Dann ist p_f ein Teiler von χ_f , und jeder Primfaktor, der in der Primfaktorzerlegung von χ_f vorkommt, kommt auch in der Primfaktorzerlegung von p_f vor, im Allgemeinen allerdings mit geringerer Vielfachheit.*

Eine Diagonalmatrix $A = \text{Diag}(\lambda_1, \dots, \lambda_n) \in K^{n \times n}$ ist bereits von allgemeiner Normalform, da die Matrix $A_i = (\lambda_i) \in K^{1 \times 1}$ als Begleitmatrix des Polynoms $T - \lambda_i \in K[T]$ aufgefasst werden kann. Aus Theorem 11 ergibt sich daher folgendes Diagonalisierbarkeitskriterium:

Korollar 13. *Folgende Aussagen sind für einen Endomorphismus f eines endlich-dimensionalen K -Vektorraums (bzw. eine Matrix $A \in K^{n \times n}$) äquivalent:*

- (i) *f (bzw. A) ist diagonalisierbar.*
- (ii) *Das Minimalpolynom p_f (bzw. p_A) zerfällt in ein Produkt $\prod_{i=1}^r (T - \lambda_i)$ mit paarweise verschiedenen Nullstellen $\lambda_1, \dots, \lambda_r \in K$.*

Es soll als Nächstes die so genannte *Jordansche Normalform* für Matrizen hergeleitet werden. Diese Normalform kann nur in den Fällen konstruiert werden, in denen das charakteristische bzw. Minimalpolynom vollständig in lineare

Faktoren zerfällt, also beispielsweise dann, wenn der Körper K algebraisch abgeschlossen ist. Für $\lambda \in K$ bezeichnen wir die Matrix

$$J(\lambda, n) = \begin{pmatrix} \lambda & & & & 0 \\ 1 & \lambda & & & \\ & 1 & \lambda & & \\ & & \cdot & \cdot & \\ & & & \cdot & \cdot \\ 0 & & & & 1 & \lambda \end{pmatrix} \in K^{n \times n}$$

als Jordankästchen der Länge n zum Eigenwert λ .

Lemma 14. Für $A \in K^{n \times n}$ und $\lambda \in K$ ist äquivalent:

- (i) A ist ähnlich zu $J(\lambda, n)$.
- (ii) A ist ähnlich zur Begleitmatrix $A(q)$ des Polynoms $q = (T - \lambda)^n \in K[T]$.

Beweis. Man interpretiere A als beschreibende Matrix eines Endomorphismus $f: V \rightarrow V$, wobei V ein n -dimensionaler K -Vektorraum sei. Ist dann A ähnlich zu $J(\lambda, n)$, so existiert eine K -Basis $x_1, \dots, x_n$ von V mit

$$(*) \quad f(x_i) = \lambda x_i + x_{i+1}, \quad i = 1, \dots, n,$$

wenn wir $x_{n+1} = 0$ setzen. Mit Induktion ergibt sich hieraus

$$x_1, \dots, x_i \in \langle f^0(x_1), \dots, f^{i-1}(x_1) \rangle, \quad i = 1, \dots, n;$$

denn für $i = 1$ ist dies trivial, und unter Benutzung der Induktionsvoraussetzung zum Index i schließt man

$$\begin{aligned} x_{i+1} &= f(x_i) - \lambda x_i \\ &\in \langle f^1(x_1), \dots, f^i(x_1) \rangle + \langle f^0(x_1), \dots, f^{i-1}(x_1) \rangle \\ &= \langle f^0(x_1), \dots, f^i(x_1) \rangle. \end{aligned}$$

Insbesondere gilt daher

$$V = \langle f^0(x_1), \dots, f^{n-1}(x_1) \rangle,$$

d. h. V ist f -zyklisch und wird als $K[T]$ -Modul von x_1 erzeugt.

Man betrachte nun den surjektiven $K[T]$ -Modulhomomorphismus

$$\varphi: K[T] \rightarrow V, \quad h \mapsto h \cdot x_1 = h(f)(x_1).$$

Aus $(*)$ ergibt sich

$$(f - \lambda \text{id})(x_i) = x_{i+1}, \quad i = 1, \dots, n,$$

also

$$(f - \lambda \text{id})^i(x_1) = \begin{cases} x_{i+1} \neq 0 & \text{für } i = 0, \dots, n-1 \\ 0 & \text{für } i = n \end{cases}.$$

Der Kern von φ enthält daher das Polynom $(T - \lambda)^n$, nicht aber $(T - \lambda)^{n-1}$. Als Hauptideal wird $\ker \varphi$ von einem normierten Polynom $q \in K[T]$ erzeugt. Dieses teilt $(T - \lambda)^n$, nicht aber $(T - \lambda)^{n-1}$, und stimmt folglich mit $(T - \lambda)^n$ überein. Aufgrund des Isomorphiesatzes für Moduln induziert φ dann einen Isomorphismus $K[T]/((T - \lambda)^n) \xrightarrow{\sim} V$ im Sinne von $K[T]$ -Moduln, und man erkennt V gemäß Satz 4 als f -zyklisch mit Minimalpolynom $p_f = (T - \lambda)^n$. Schließlich folgt mit Lemma 10, dass A zur Begleitmatrix des Minimalpolynoms $(T - \lambda)^n$ ähnlich ist. Die Implikation von (i) nach (ii) ist daher bewiesen.

Zum Nachweis der umgekehrten Implikation sei A nun ähnlich zur Begleitmatrix des Polynoms $q = (T - \lambda)^n \in K[T]$. Dann ist V nach Lemma 10 f -zyklisch mit Minimalpolynom q zu f , und es gilt $V \simeq K[T]/(q)$ im Sinne von $K[T]$ -Moduln. Wir wollen zeigen, dass die Restklassen

$$(\overline{T} - \lambda)^0, (\overline{T} - \lambda)^1, \dots, (\overline{T} - \lambda)^{n-1}$$

eine K -Basis von $K[T]/(q)$ bilden. Gemäß Lemma 5 ist bereits bekannt, dass dies für die Potenzen $\overline{T}^0, \dots, \overline{T}^{n-1}$ gilt. Es genügt daher, wenn wir induktiv

$$\overline{T}^0, \dots, \overline{T}^{i-1} \in \langle (\overline{T} - \lambda)^0, \dots, (\overline{T} - \lambda)^{i-1} \rangle, \quad i = 1, \dots, n,$$

zeigen. Für $i = 1$ ist dies trivial, und unter Verwendung der Induktionsvoraussetzung zum Index i können wir wie folgt schließen:

$$\begin{aligned} \overline{T}^i &= (\overline{T} - \lambda) \cdot \overline{T}^{i-1} + \lambda \overline{T}^{i-1} \\ &\in \langle (\overline{T} - \lambda)^1, \dots, (\overline{T} - \lambda)^i \rangle + \langle (\overline{T} - \lambda)^0, \dots, (\overline{T} - \lambda)^{i-1} \rangle \\ &= \langle (\overline{T} - \lambda)^0, \dots, (\overline{T} - \lambda)^i \rangle \end{aligned}$$

Folglich bilden die Potenzen $(\overline{T} - \lambda)^{i-1}$ für $i = 1, \dots, n$ eine K -Basis von $K[T]/(q)$.

Wir betrachten nun die Gleichungen

$$\overline{T} \cdot (\overline{T} - \lambda)^{i-1} = \lambda \cdot (\overline{T} - \lambda)^{i-1} + (\overline{T} - \lambda)^i, \quad i = 1, \dots, n,$$

wobei $(\overline{T} - \lambda)^n = 0$ gilt. Sie besagen, dass der im Sinne von Bemerkung 3 zu f korrespondierende Endomorphismus

$$K[T]/(q) \longrightarrow K[T]/(q), \quad \overline{h} \longmapsto \overline{T} \cdot \overline{h},$$

bezüglich obiger Basis durch das Jordankästchen $J(\lambda, n)$ beschrieben wird. Dann gibt es aber auch in V eine Basis, bezüglich der f durch $J(\lambda, n)$ beschrieben wird, und die Implikation von (ii) nach (i) ist bewiesen. $\square$

Nun lassen sich Existenz und Eindeutigkeit der Jordanschen Normalform leicht aus Theorem 11 folgern.

Theorem 15. *Es sei $A \in K^{n \times n}$ eine Matrix, deren Minimal- bzw. charakteristisches Polynom vollständig in Linearfaktoren zerfällt. Dann ist A ähnlich zu einer so genannten Jordan-Matrix*

$$\text{Diag}(J(\lambda_1, n_1), \dots, J(\lambda_r, n_r)),$$

deren “Einträge” Jordankästchen sind. Die Elemente λ_i, n_i sind, abgesehen von der Reihenfolge, eindeutig durch A bestimmt, wobei die λ_i (unter eventueller Mehrfachaufzählung) gerade die Eigenwerte von A durchlaufen. Wählt man einen n -dimensionalen K -Vektorraum V und realisiert A als Matrix eines Endomorphismus $f: V \rightarrow V$, so ist

$$V \simeq \bigoplus_{i=1}^r K[T]/((T - \lambda_i)^{n_i})$$

gerade die Zerlegung aus 6.4/3 bzw. Theorem 6.

Beweis. Man benutze Lemma 14, um von der allgemeinen Normalform aus Theorem 11 von A zur Jordanschen Normalform zu gelangen bzw. umgekehrt von der Jordanschen Normalform zur allgemeinen Normalform. $\square$

Wir sagen, eine Matrix $A \in K^{n \times n}$ sei *trigonalisierbar*, wenn sie ähnlich zu einer Matrix der Form $B = (\beta_{ij}) \in K^{n \times n}$ ist mit $\beta_{ij} = 0$ für $i < j$; man nennt B eine (untere) *Dreiecksmatrix*. Aus der Existenz der Jordanschen Normalform kann man insbesondere ein Kriterium für Trigonalisierbarkeit ableiten.

Korollar 16. *Eine Matrix $A \in K^{n \times n}$ ist genau dann trigonalisierbar, wenn das Minimal- bzw. charakteristische Polynom von A vollständig in Linearfaktoren zerfällt.*

Beweis. Sei A trigonalisierbar, also ähnlich zu einer Matrix $B = (\beta_{ij}) \in K^{n \times n}$ mit $\beta_{ij} = 0$ für $i < j$. Dann gilt $\chi_A = \chi_B = \prod_{i=1}^n (T - \beta_{ii})$; insbesondere zerfällt χ_A und damit nach Korollar 12 auch das Minimalpolynom p_A vollständig in Linearfaktoren. Ist umgekehrt letztere Bedingung gegeben, so zeigt die Existenz der Jordanschen Normalform, dass A trigonalisierbar ist. $\square$

Wir wollen als Nächstes ein erstes (recht grobes) praktisches Verfahren zur expliziten Berechnung der Jordanschen Normalform einer Matrix angeben. Man betrachte also eine Matrix $A \in K^{n \times n}$, deren charakteristisches Polynom vollständig in Linearfaktoren zerfällt, etwa

$$\chi_A = (T - \lambda_1)^{n_1} \cdot \dots \cdot (T - \lambda_r)^{n_r};$$

im Unterschied zu Theorem 15 setzen wir hierbei voraus, dass die Eigenwerte $\lambda_1, \dots, \lambda_r$ paarweise verschieden sind. Die Jordansche Normalform J zu A ist dann eine Matrix, auf deren Diagonalen Jordankästchen des Typs $J(\lambda, s)$ stehen.

Für $i = 1, \dots, r$ und $j = 1, \dots, n_i$ bezeichne $k_{i,j}$ diejenige Anzahl, mit der das Jordankästchen $J(\lambda_i, j)$ in J vorkommt. Es gilt

$$k_{i,1} + 2k_{i,2} + \dots + n_i k_{i,n_i} = n_i,$$

da die Vielfachheit des Eigenwertes λ_i gerade n_i ist, das Element λ_i also genau n_i -mal auf der Diagonalen von J vorkommt.

Da die Matrizen A und J zueinander ähnlich sind, sind auch die Matrizen $A - \lambda E$ und $J - \lambda E$ sowie die Potenzen $(A - \lambda E)^\ell$ und $(J - \lambda E)^\ell$ mit $\ell \in \mathbb{N}$ zueinander ähnlich; dabei sei $\lambda \in K$ und $E \in K^{n \times n}$ die Einheitsmatrix. Insbesondere folgt dann mit 3.4/6

$$\operatorname{rg}(A - \lambda E)^\ell = \operatorname{rg}(J - \lambda E)^\ell \quad \text{für } \ell \in \mathbb{N},$$

und es ist $(J - \lambda E)^\ell$ wiederum eine "Diagonalmatrix", gebildet aus Kästchen, nämlich aus den ℓ -ten Potenzen der Jordankästchen von $J - \lambda E$. Nun berechnet sich für ein Jordankästchen des Typs $J(\lambda, m)$ der Rang einer ℓ -ten Potenz zu

$$\operatorname{rg} J(\lambda, m)^\ell = \begin{cases} \max\{0, m - \ell\} & \text{für } \lambda = 0 \\ m & \text{für } \lambda \neq 0 \end{cases}.$$

Daher bestehen folgende Gleichungen:

$$\begin{aligned} \operatorname{rg}(A - \lambda_i E)^{n_i} &= \operatorname{rg}(J - \lambda_i E)^{n_i} = n - n_i \\ \operatorname{rg}(A - \lambda_i E)^{n_i-1} &= \operatorname{rg}(J - \lambda_i E)^{n_i-1} = n - n_i + k_{i,n_i} \\ \operatorname{rg}(A - \lambda_i E)^{n_i-2} &= \operatorname{rg}(J - \lambda_i E)^{n_i-2} = n - n_i + 2k_{i,n_i} + k_{i,n_i-1} \\ &\vdots \\ \operatorname{rg}(A - \lambda_i E)^1 &= \operatorname{rg}(J - \lambda_i E)^1 = n - n_i + (n_i - 1)k_{i,n_i} + \dots + k_{i,2} \end{aligned}$$

Ermittelt man also die Ränge der Potenzen von $A - \lambda_i E$, so lassen sich die Zahlen $k_{i,j}$, $j = n_i, n_i - 1, \dots, 2$, der Reihe nach berechnen. Weiter gilt aufgrund obiger Gleichung

$$k_{i,1} = n_i - 2k_{i,2} - \dots - n_i k_{i,n_i},$$

so dass man auch $k_{i,1}$ berechnen kann. Sind aber alle Zahlen $k_{i,j}$ bekannt, so hat man insgesamt die Jordansche Normalform J von A bestimmt.

Als Beispiel wollen wir die Jordansche Normalform der Matrix

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ -1 & 1 & 0 & 3 \end{pmatrix} \in \mathbb{R}^{4 \times 4}$$

bestimmen. Es gilt $\chi_A = (T - 2)^4$, also $r = 1$ und $n_1 = 4$ in obiger Notation. Wir haben

$$\operatorname{rg}(A - 2E) = \operatorname{rg} \begin{pmatrix} -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 1 \\ -1 & 1 & 0 & 1 \end{pmatrix} = 1,$$

sowie $\operatorname{rg}(A - 2E)^2 = \operatorname{rg}(0) = 0$ und damit auch $\operatorname{rg}(A - 2E)^s = 0$ für $s \geq 2$.
Daher bestehen die Gleichungen

$$\begin{aligned} 0 &= \operatorname{rg}(A - 2E)^4 = n - n_1 = 0, \\ 0 &= \operatorname{rg}(A - 2E)^3 = k_{1,4}, \\ 0 &= \operatorname{rg}(A - 2E)^2 = 2k_{1,4} + k_{1,3}, \\ 1 &= \operatorname{rg}(A - 2E)^1 = 3k_{1,4} + 2k_{1,3} + k_{1,2}, \end{aligned}$$

und dies ergibt $k_{1,4} = k_{1,3} = 0$, $k_{1,2} = 1$, sowie

$$k_{1,1} = 4 - 2k_{1,2} = 2.$$

Folglich ist $J = \operatorname{Diag}(J(2, 1), J(2, 1), J(2, 2))$, also

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

die Jordansche Normalform zu A .

Wir behandeln abschließend noch ein weiteres viel effektiveres Verfahren, mit dessen Hilfe man neben der Jordanschen auch die allgemeine Normalform von Matrizen ermitteln kann. Das Vorgehen stellt sich wie folgt dar: Ausgehend von einer Matrix $A \in K^{n \times n}$ interpretieren wir diese als Endomorphismus f des n -dimensionalen K -Vektorraums $V = K^n$, so dass wir V unter f als $K[T]$ -Modul auffassen können. Sodann konstruieren wir zu A eine kanonische endliche Präsentation von V als $K[T]$ -Modul und bestimmen die zugehörigen Elementarteiler. Wir gelangen auf diese Weise zu der in 6.4/2 angegebenen Zerlegung von V in monogene $K[T]$ -Untermodule, aus der sich alle weiteren Zerlegungen und insbesondere auch die Normalformen von A ergeben.

Lemma 17. *Für eine Matrix $A \in K^{n \times n}$ betrachte man K^n als $K[T]$ -Modul unter dem durch $x \mapsto A \cdot x$ gegebenen Endomorphismus. Es sei*

$$\varphi: K[T]^n \longrightarrow K^n, \quad (p_1, \dots, p_n) \longmapsto \sum_{i=1}^n p_i(A) \cdot e_i,$$

diejenige $K[T]$ -lineare Abbildung, die die kanonische $K[T]$ -Basis von $K[T]^n$ auf die kanonische K -Basis $e_1, \dots, e_n$ von K^n abbildet; φ ist surjektiv. Weiter sei

$$\psi: K[T]^n \longrightarrow K[T]^n, \quad P \longmapsto (TE - A) \cdot P,$$

die durch die Matrix $TE - A \in K[T]^{n \times n}$ gegebene $K[T]$ -lineare Abbildung; $E \in K^{n \times n}$ sei die Einheitsmatrix. Dann ist die Sequenz

$$K[T]^n \xrightarrow{\psi} K[T]^n \xrightarrow{\varphi} K^n \longrightarrow 0$$

eine endliche Präsentation von K^n als $K[T]$ -Modul.

Beweis. Bezeichnet e_i in $K[T]^n$, wie auch in K^n , den i -ten Einheitsvektor, so gilt für $i = 1, \dots, n$

$$\psi(e_i) = (TE - A)e_i = Te_i - Ae_i, \quad \varphi(Te_i) = Ae_i, \quad \varphi(Ae_i) = Ae_i,$$

und damit

$$\varphi \circ \psi(e_i) = \varphi(Te_i - Ae_i) = Ae_i - Ae_i = 0.$$

Es folgt $\varphi \circ \psi = 0$, also im $\psi \subset \ker \varphi$. Aufgrund des Homomorphiesatzes zerlegt sich φ dann in eine Komposition

$$\varphi: K[T]^n \xrightarrow{\pi} K[T]^n / \text{im } \psi \xrightarrow{\bar{\varphi}} K^n$$

surjektiver Abbildungen, und wir behaupten, dass $\bar{\varphi}$ sogar bijektiv ist. Da $K[T]^n$ als K -Vektorraum von den Elementen $T^\nu e_i$ mit $i = 1, \dots, n$ und $\nu \in \mathbb{N}$ erzeugt wird, erzeugen entsprechend die Bilder $\pi(T^\nu e_i)$ den Quotienten $K[T]^n / \text{im } \psi$ als K -Vektorraum. Nun gilt aber für alle $\nu \in \mathbb{N}$

$$T^{\nu+1}e_j - T^\nu Ae_j = \psi(T^\nu e_j) \in \text{im } \psi, \quad j = 1, \dots, n,$$

und deshalb

$$T^{\nu+1}e_j \in T^\nu \sum_{i=1}^n Ke_i + \text{im } \psi, \quad j = 1, \dots, n.$$

Per Induktion folgt hieraus für $\nu \in \mathbb{N}$

$$T^\nu e_j \in \sum_{i=1}^n Ke_i + \text{im } \psi, \quad j = 1, \dots, n,$$

und damit

$$K[T]^n = \sum_{i=1}^n Ke_i + \text{im } \psi.$$

Dies bedeutet, dass $K[T]^n / \text{im } \psi$ als K -Vektorraum von n Elementen erzeugt wird und damit über K eine Dimension $\leq n$ besitzt. Aufgrund der Dimensionsformel 2.1/10 ist $\bar{\varphi}$ dann notwendigerweise injektiv und damit bijektiv, wie behauptet. Dann gilt aber $\text{im } \psi = \ker \pi = \ker \varphi$, und es folgt, dass die angegebene Sequenz eine endliche Präsentation von K^n als $K[T]$ -Modul darstellt. $\square$

Theorem 18. *Es sei $f: V \rightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V . Sei $A = A_{f,X,X} \in K^{n \times n}$ die Matrix, welche f bezüglich einer gegebenen Basis X von V beschreibt, und seien $\alpha_1, \dots, \alpha_s \in K[T]$ mit $\alpha_j | \alpha_{j+1}$ diejenigen Elementarteiler der Matrix $TE - A \in K[T]^{n \times n}$, die nicht invertierbar sind; wir nehmen die α_j als normierte Polynome an. Weiter betrachten wir die Primfaktorzerlegungen $\alpha_j = p_1^{n(1,j)} \dots p_r^{n(r,j)}$, $j = 1, \dots, s$, der α_j mit paarweise verschiedenen normierten Primpolynomen $p_1, \dots, p_r \in K[T]$ und Exponenten $n(i, j) \geq 0$. Dann gilt:*

(i) Fasst man V als $K[T]$ -Modul unter f auf, so folgt

$$V \simeq \bigoplus_{j=1}^s K[T]/\alpha_j K[T] \simeq \bigoplus_{i=1}^r \bigoplus_{j=1}^s K[T]/p_i^{n(i,j)} K[T],$$

und die letztere Zerlegung gibt Anlass zu einer Zerlegung von V in f -zyklische Untervektorräume gemäß Theorem 6.

(ii) Die Kästchen der allgemeinen Normalform von A sind die Begleitmatrizen zu den Primpotenzen $p_i^{n(i,j)}$ mit $n(i,j) > 0$, $i = 1, \dots, r$, $j = 1, \dots, s$. Falls alle p_i linear sind, erhält man hieraus die Jordansche Normalform mittels Lemma 14.

(iii) Charakteristisches Polynom χ_f und Minimalpolynom p_f von f berechnen sich zu

$$\chi_f = \alpha_1 \dots \alpha_s, \quad p_f = \alpha_s.$$

Beweis. Wir können $V = K^n$ und f als den durch $x \mapsto Ax$ gegebenen Endomorphismus von V annehmen. Unter Benutzung der in Lemma 17 bereitgestellten endlichen Präsentation von V erhalten wir dann mit 6.4/2 die erste Zerlegung aus (i) und mittels des Chinesischen Restsatzes 6.4/1 auch die zweite, wobei der freie Anteil in 6.4/2 entfällt, da V ein $K[T]$ -Torsionsmodul ist. Die restlichen Aussagen ergeben sich mittels Satz 4 sowie mit den Lemmata 8 und 10, wie im Beweis zu Theorem 11. $\square$

Als Beispiel betrachten wir nochmals die Matrix

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ -1 & 1 & 0 & 3 \end{pmatrix} \in \mathbb{R}^{4 \times 4}$$

und bestimmen zunächst die Elementarteiler der Matrix $TE - A \in \mathbb{R}^{4 \times 4}$, indem wir das Verfahren aus dem Beweis zu 6.3/5 anwenden:

$$\begin{aligned} & \begin{pmatrix} T-1 & -1 & 0 & -1 \\ 0 & T-2 & 0 & 0 \\ 1 & -1 & T-2 & -1 \\ 1 & -1 & 0 & T-3 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -1 & 0 & T-3 \\ 0 & T-2 & 0 & 0 \\ 1 & -1 & T-2 & -1 \\ T-1 & -1 & 0 & -1 \end{pmatrix} \\ & \mapsto \begin{pmatrix} 1 & -1 & 0 & T-3 \\ 0 & T-2 & 0 & 0 \\ 0 & 0 & T-2 & -(T-2) \\ 0 & T-2 & 0 & -(T-2)^2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & T-2 & 0 & 0 \\ 0 & 0 & T-2 & 0 \\ 0 & 0 & 0 & (T-2)^2 \end{pmatrix} \end{aligned}$$

Damit erhalten wir $T-2, T-2, (T-2)^2$ als die nicht-invertierbaren Elementarteiler von $TE - A$ und gemäß Theorem 18 dann $\chi_A = (T-2)^4$ als charakteristisches Polynom, sowie $p_A = (T-2)^2$ als Minimalpolynom zu A . Da die

Elementarteiler bereits Primpotenzen sind, entfällt die Anwendung des Chinesischen Restsatzes und allgemeine bzw. Jordansche Normalform von A ergeben sich zu

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & -4 \\ 0 & 0 & 1 & 4 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$$

Aufgaben

Im Folgenden sei K ein Körper.

1. Es sei $A \in \mathbb{R}^{7 \times 7}$ eine Matrix mit charakteristischem Polynom

$$\chi_A = (T^2 + 1)^2(T - 2)(T^2 - 1) \in \mathbb{R}[T].$$

Man untersuche, welche Gestalt die allgemeine Normalform von A haben kann.

2. Man zeige, dass zwei Matrizen $A, B \in \mathbb{R}^{3 \times 3}$ genau dann ähnlich sind, wenn ihre Minimalpolynome sowie ihre charakteristischen Polynome übereinstimmen.
3. Für die Matrizen

$$\begin{pmatrix} 2 & 2 & 0 & -3 \\ 1 & 1 & 0 & -1 \\ 1 & 2 & -1 & -1 \\ 1 & 2 & 0 & -2 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 1 & -2 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & -2 \\ -1 & 1 & 1 & -1 \end{pmatrix}, \quad \begin{pmatrix} 3 & 0 & 1 & -2 \\ 4 & 1 & 1 & -3 \\ 0 & 0 & -1 & 0 \\ 4 & 0 & 2 & -3 \end{pmatrix} \in \mathbb{R}^{4 \times 4}$$

berechne man jeweils charakteristisches und Minimalpolynom sowie, falls existent, die Jordansche Normalform.

4. Es seien $A, B \in K^{n \times n}$ Matrizen mit den Minimalpolynomen $p_A, p_B \in K[T]$ und mit Primfaktorzerlegung $p_A = p_1^{n_1} \dots p_r^{n_r}$, wobei $p_1, \dots, p_r \in K[T]$ paarweise verschiedene normierte Primpolynome seien. Man zeige, dass A und B genau dann ähnlich sind, wenn $p_A = p_B$ sowie $\operatorname{rg} p_i^{k_i}(A) = \operatorname{rg} p_i^{k_i}(B)$ für $i = 1, \dots, r$ und $1 \leq k_i \leq n_i$ gilt.
5. Man betrachte normierte Primpolynome $p_1, \dots, p_n \in K[T]$, die paarweise verschieden seien, sowie natürliche Zahlen $1 \leq r_i \leq s_i$, $i = 1, \dots, n$. Gibt es einen Endomorphismus $f: V \rightarrow V$ eines endlich-dimensionalen K -Vektorraums V mit Minimalpolynom $p_f = p_1^{r_1} \dots p_n^{r_n}$ und mit charakteristischem Polynom $\chi_f = p_1^{s_1} \dots p_n^{s_n}$?
6. Es sei V ein endlich-dimensionaler K -Vektorraum mit einem Endomorphismus $f: V \rightarrow V$ und $U \subset V$ ein f -invarianter Unterraum. Man zeige:
 - (i) f induziert einen Endomorphismus $\bar{f}: V/U \rightarrow V/U$.
 - (ii) Es gilt $p_{\bar{f}} | p_f$ für die Minimalpolynome von $\bar{f}$ und f .
 - (iii) Es gilt $\chi_f = \chi_{f|U} \cdot \chi_{\bar{f}}$ für die charakteristischen Polynome von f , von $f|_U$ (der Einschränkung von f auf U) und von $\bar{f}$.
7. Es sei V ein endlich-dimensionaler K -Vektorraum mit einem Endomorphismus $f: V \rightarrow V$. Das Minimalpolynom p_f sei Potenz eines Primpolynoms $p \in K[T]$, etwa $p_f = p^r$ mit $r > 0$. Man zeige:
 - (i) Es existiert ein Vektor $u \in V$ mit $p^{r-1}(f)(u) \neq 0$.

- (ii) Ist $u \in V$ wie in (i), und ist $U \subset V$ der von u erzeugte f -zyklische Untervektorraum, so existiert ein f -invarianter Untervektorraum $U' \subset V$ mit $V = U \oplus U'$.
8. Es sei V ein endlich-dimensionaler f -zyklischer K -Vektorraum unter einem Endomorphismus $f: V \rightarrow V$. Man zeige, dass jeder f -invariante Unterraum $U \subset V$ wiederum f -zyklisch ist.
9. Es sei V ein endlich-dimensionaler f -zyklischer K -Vektorraum unter einem Endomorphismus $f: V \rightarrow V$, und es sei $u \in V$ ein erzeugendes Element. Man zeige:
- (i) Es gibt auf V eine eindeutig bestimmte Ringstruktur, deren Addition mit der Vektorraumaddition auf V übereinstimmt und deren Multiplikation $(af^i(u)) \cdot (bf^j(u)) = abf^{i+j}(u)$ für $a, b \in K$ und $i, j \in \mathbb{N}$ erfüllt.
 - (ii) Unter dieser Ringstruktur ist V genau dann ein Körper, wenn das Minimalpolynom zu f prim ist.