universität wien

DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

Zugänge zur Jordanschen Normalform

Ein Vergleich

verfasst von / submitted by

Pablo Zuder

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Magister der Naturwissenschaften (Mag. rer. nat.)

Wien, 2016 / Vienna, 2016

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 190 406 299
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Lehramtsstudium UF Mathematik UF Psychologie und Philosophie
Betreut von / Supervisor:	Mag. Dr. Günther Hörmann

Mein aufrichtiger Dank gilt in erster Linie Günther, der mich nicht nur fachlich begleitet und gestützt hat, sondern mir auch als Mensch Vorbild wurde.

Vorwort

Der Gegenstand dieser Arbeit ist das Zeigen der Existenz einer sogenannten *Jordan-Basis*. Diese Basis ermöglicht es Matrizen (aufgefasst als Darstellungen linearer Abbildungen) derart zu transformieren, dass die nach der Transformation (Multiplikation mit der Transformationsmatrix bezüglich der Jordan-Basis) entstehende Matrix Normalform, und zwar *Jordansche Normalform* besitzt.

Normalform bezeichnet hierbei eine standardisierte - sozusagen einer Norm entsprechenden - Matrixdarstellung. Eine der herausragenden Eigenschaften der *Jordanschen Normalform* ist nun das Bereitstellen einer bestimmten Kategorisierungsmöglichkeit von linearen Abbildungen. Das heißt: befinden sich zwei Matrizen in *Jordanscher Normalform* wird unmittelbar offensichtlich, ob es sich (bis auf Isomorphie der unterliegenden Vektorräume) um Matrixdarstellungen derselben linearen Abbildung bzw. desselben Endomorphismus (wir betrachten ausschließlich quadratische Matrizen) handelt.

Aus der ursprünglichen Absicht „einer“ Herleitung der Jordanschen Normalform entwickelte sich während der intensiveren Auseinandersetzung mit dieser Thematik ein anderer Gedanke, nämlich die Idee der Betrachtung mehrerer Möglichkeiten zum Existenzbeleg einer Jordan-Basis. Konkret äußert sich das in der Präsentation dreier an sich unabhängiger Darstellungen zur Entwicklung der Jordanschen Normalform. Die Reihenfolge dieser Ansätze ist nach aufsteigendem Abstraktionsgrad gesetzt, was den jeweils nachfolgenden Zugang leichter verdaulich werden lässt und insgesamt zu einem tiefgründigeren Verständnis der *Jordanschen Normalform* beitragen soll.

Die Beweisführungen sind größtenteils vollständig, und ansonsten mit entsprechenden Referenzen versehen. Die Ausführung von Beispielen ist äußerst spärlich, auf die konkrete Bestimmung und auf Vermerke zum anwendungsorientierten Nutzen der *Jordanschen Normalform* (z.B. Lösen von linearen Differentialgleichungen) wird gänzlich verzichtet; ohnehin ist dies Gegenstand gängiger Lehrbücher zur Linearen Algebra oder entsprechender Lehrveranstaltungen. Die Darstellung der *Jordanschen Normalform* für den Spezialfall eines Vektorraumes über $\mathbb{K} = \mathbb{R}$ wird in dieser Arbeit nur am Rande behandelt.

Inhaltsverzeichnis

1. Einleitung

Wie eingangs erwähnt werden in dieser Abhandlung drei eigenständige Zugänge zum Beleg der *Jordanschen Normalform* ausgeführt.

In der ersten Herangehensweise (Kapitel 4), die auf Gerd Fischers *Lineare Algebra* [Fis10] gestützt ist, wird elementar und auch am ausführlichsten gezeigt, dass die *Jordansche Normalform* die in gewissem Sinne bestmögliche Matrixdarstellung von Endomorphismen ist, deren charakteristische Polynome in Linearfaktoren zerfallen. Zuvor werden in Kapitel 3, sozusagen vorbereitend, noch andere markante Matrixdarstellungen von Endomorphismen thematisiert, die aber an sich von unbedeutender Natur sind. Keineswegs von geringer Bedeutung ist die *Allgemeine Normalform*, deren Behandlung dann den Schluss dieser Arbeit bilden wird.
Im letzten Teil von Kapitel 4 wird noch das Klassifikationspotenzial der *Jordanschen Normalform* aufgegriffen.

In Kapitel 5 wird der zweite Ansatz zur Entwicklung der Jordanschen Normalform, der auf Theodor Bröckers *Lineare Algebra und Analytische Geometrie* [Brö03] basiert, Gegenstand unserer Betrachtungen sein. Diese Darstellung ist wie erwähnt abstrakter als die erste und besitzt durch das gesteigerte Abstraktionsmaß eine gewisse Prägnanz. Außerdem können wir hier schon auf Überlegungen des ersten Ansatzes zurückgreifen. Die „grobe“ Hauptraumzerlegung von V erfolgt über das Minimalpolynom, die weitere Zerlegung der Haupträume - wie im ersten Zugang - mittels Betrachtungen für den nilpotenten Anteil des entsprechenden Endomorphismus.

Im dritten, auf Siegfried Bosch' *Lineare Algebra* [Bos14] gründenden Ansatz (Kapitel 7), stoßen wir auf Inhalte, die der Algebra zugeschrieben werden. Letztlich zeigt sich jedoch, wie mit geeigneter Interpretation diese algebraischen Betrachtungen fruchtbar in die Lineare Algebra und in unser Themengebiet überführt werden können und die oben beschriebene Vektorraumzerlegung herzuleiten vermögen.
Gegen Ende dieses dritten Zugangs finden dann noch Überlegungen zur Existenz der *Allgemeinen Normalform* Platz. Ein Zugang nämlich, der weitestgehend von der Bedingung eines in Linearfaktoren zerfallenden Minimalpolynoms (bzw. charakteristischen Polynoms) unabhängig ist und somit allgemeiner abgewickelt werden kann als die beiden vorhergehenden. Die *Jordansche Normalform* kann dann als Spezialfall abgeleitet werden.

Während die ersten beiden Ansätze mittels der Jordan-Basis bis hin zur *Jordanschen Normalform* ausgeführt werden, verzichtet der dritte auf die explizite Konstruktion einer Jordan-Basis und begnügt sich mit dem Belegen der Existenz einer nicht weiter invarianten, direkten Summenzerlegung des dem Endomorphismus zu Grunde liegenden Vektorraumes.

Für entsprechende Stellen werden auch Querverbindungen zwischen den einzelnen Ansätzen generiert.

Anzumerken bleibt noch, dass in dieser Arbeit ausschließlich auf *endlichdimensionalen* Vektorräumen operiert wird.

2. Grundlegende Voraussetzungen

Nachstehend sollen die für diese Abhandlung elementarsten Begrifflichkeiten und wesentlichen Erkenntnisse erörtert werden. Diese sind zwar mehr oder weniger Inhalt einführender Vorlesungen der Linearen Algebra, sollen aber auf Grund ihrer Allgegenwärtigkeit hier nochmals zur Sprache gebracht werden.
Nicht zuletzt auch deshalb um der Leserin oder dem Leser das Wiedererkennen bekannter substanzieller Begriffe in der verwendeten Nomenklatur dieser Arbeit zu ermöglichen.
Andererseits wird auf mit der Behandlung von (endlich-dimensionalen) Vektorräumen über einem Körper $\mathbb{K}$ verbundene Begriffe wie Basis oder Dimension, auf essentielle Kenntnisse über Determinanten sowie auf Eigenschaften von Matrizen oder gar Mengen und (linearen) Abbildungen nicht explizit eingegangen und der Leserin oder dem Leser bei Notwendigkeit als Selbststudium überantwortet. Auch Eigenwerte und Eigenvektoren werden als bekannt vorausgesetzt.

2.1. Matrixdarstellung linearer Abbildungen

Die verwendete Quelle für diesen Abschnitt bildet [Cap10].

Sei $\mathbb{K}$ ein Körper. Wir rufen uns in Erinnerung, dass die Vektorräume der linearen Abbildungen zwischen $\mathbb{K}^n$ und $\mathbb{K}^m$ und der m×n-Matrizen isomorph sind, in Zeichen bedeutet das $L(\mathbb{K}^n, \mathbb{K}^m) \cong M_{m,n}(\mathbb{K})$.

Hierbei wird eine Matrix A als ein n-Tupel von Spaltenvektoren aufgefasst. Diese Spaltenvektoren sind die Bilder der (geordneten) Standardbasis $\mathcal{E}$ mit den Vektoren $e_1, e_2, \ldots, e_n$, d.h. $A = (f(e_1), f(e_2), \ldots, f(e_n)) \in M_{m,n}(\mathbb{K})$, wobei $e_i = (0, \ldots, 1, \ldots, 0) \in \mathbb{K}^n$ mit einer 1 an der i-ten Stelle und sonst nur 0 als Einträge.

Dies bedeutet, dass wir im Sinne eines Isomorphismus nicht mehr zwischen Erkenntnissen linearer Abbildungen und Erkenntnissen von Matrizen unterscheiden müssen, was uns in der Tat bei einigen kommenden Beweisführungen Arbeit ersparen wird.

Koordinatenvektoren. Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$. Für eine (geordnete) Basis $\mathcal{B}$ für V, die aus den Vektoren $v_1, \ldots, v_n$ besteht, gibt es eine eindeutige lineare Abbildung $\varphi = \varphi_{\mathcal{B}} : V \to \mathbb{K}^n$ mit $\varphi(v_i) = e_i$. Diese Abbildung ist ein linearer Isomorphismus.
Sei $v \in V$ ein beliebiger Vektor, dann gibt es eindeutige Elemente $r_1, \ldots, r_n \in \mathbb{K}$,

sodass $v = r_1v_1 + \cdots + r_nv_n$ ist. Wendet man auf diesem Vektor v die Abbildung φ an, ergibt sich

$$\varphi(v) = r_1e_1 + \cdots + r_ne_n = (r_1, \ldots, r_n) \in \mathbb{K}^n.$$

Die Einträge r_i sind also die Koeffizienten der entsprechenden Linearkombination bezüglich der Basis $\mathcal{B}$.
Das Bild $\varphi(v)$ wird als *Koordinatenvektor von v bezüglich der Basis $\mathcal{B}$* bezeichnet. Man schreibt für diesen Vektor $[v]_{\mathcal{B}}$.
Sei nun V' ein weiterer $\mathbb{K}$-Vektorraum mit $\dim(V') = m$ und $\mathcal{B}'$ eine geordnete Basis für V', bestehend aus $v'_1, \ldots, v'_m$. Dann wird mittels des linearen Isomorphismus $\varphi' = \varphi'_{\mathcal{B}'} : V' \to \mathbb{K}^m$ der Koordinatenvektor $[v']_{\mathcal{B}'} \in \mathbb{K}^m$ bezüglich der Basis $\mathcal{B}'$ erzeugt.

$A \in M_{m,n}(\mathbb{K})$ sei die zu der linearen Abbildung $\varphi' \circ \alpha \circ \varphi^{-1} \colon \mathbb{K}^n \to \mathbb{K}^m$ gehörende Matrix, wobei $\alpha \colon V \to V'$ eine beliebige lineare Abbildung darstellt. Diese Matrix wird mit $[\alpha]^{\mathcal{B}}_{\mathcal{B}'}$ bezeichnet und heißt die *Matrixdarstellung von α bezüglich der Basen $\mathcal{B}$ und $\mathcal{B}'$*.

Proposition 2.1.1. Seien V und V' endlichdimensionale $\mathbb{K}$-Vektorräume, $\mathcal{B}$ und $\mathcal{B}'$ geordnete Basen, und $\alpha : V \to V'$ eine lineare Abbildung.
Dann ist die Matrixdarstellung $[\alpha]^{\mathcal{B}}_{\mathcal{B}'}$ charakterisiert durch $[\alpha(v)]_{\mathcal{B}'} = [\alpha]^{\mathcal{B}}_{\mathcal{B}'} \cdot [v]_{\mathcal{B}}$.
Die Spaltenvektoren von $[\alpha]^{\mathcal{B}}_{\mathcal{B}'}$ sind also gerade die Vektoren $[\alpha(v_i)]_{\mathcal{B}'}$, wobei die v_i die Elemente der Basis $\mathcal{B}$ sind.

Beweis. Nach Definition ist $[v]_{\mathcal{B}} = \varphi(v)$, also $v = \varphi^{-1}([v]_{\mathcal{B}})$. Damit ist aber

$$\varphi' \circ \alpha \circ \varphi^{-1}([v]_{\mathcal{B}}) = \varphi'(\alpha(v)) = [\alpha(v)]_{\mathcal{B}'}.$$

Dies bedeutet gerade, dass man $[\alpha(v)]_{\mathcal{B}'}$ erhält, indem man die Matrixdarstellung zu $\varphi' \circ \alpha \circ \varphi^{-1}$, also $[\alpha]^{\mathcal{B}}_{\mathcal{B}'}$, mit $[v]_{\mathcal{B}}$ multipliziert.
Die Spalten von $[\alpha]^{\mathcal{B}}_{\mathcal{B}'}$ sind - wie eingangs erwähnt - durch die Bilder der Standard-Einheitsvektoren e_i $(1 \leq i \leq n)$ gegeben. Da $[v_i]_{\mathcal{B}} = e_i$ für alle $i \in \{1, \ldots, n\}$. gilt, folgt auch der zweite Teil der Proposition. □

Bemerkung. Diese Proposition liefert somit die Anleitung, wie man (üblicherweise) zu der Matrixdarstellung einer linearen Funktion α bezüglich beliebiger Basen $\mathcal{B}$ und $\mathcal{B}'$ gelangt: Man nimmt die Basisvektoren der *alten* Basis $\mathcal{B}$, bildet diese durch α ab und beschreibt die entstehenden Bilder $\alpha(v_i)$ in Termen einer Linearkombination der *neuen* Basis $\mathcal{B}'$. Die dabei eruierten Koeffizienten als (geordnetes) Tupel angereiht ergeben dann genau die jeweiligen Spaltenvektoren $[\alpha(v_i)]_{\mathcal{B}'}$ von $[\alpha]^{\mathcal{B}}_{\mathcal{B}'}$.

Diese Anleitung wird für *Normalform-Darstellungen* von Endomorphismen noch wichtig sein.

Da wir künftig ausschließlich Endomorphismen bezüglich einer einzigen Basis $\mathcal{B}$ betrachten werden, wollen wir die Matrixdarstellung von α bezüglich der Basis $\mathcal{B}$ mit der nicht ganz so technisch wirkenden Bezeichnung $M_{\mathcal{B}}(\alpha)$ versehen (also $M_{\mathcal{B}}(\alpha)$ anstatt $[f]_{\mathcal{B}}^{\mathcal{B}}$).
In diesem Falle schreiben wir auch nur $M_n(\mathbb{K})$ anstatt $M_{n,n}(\mathbb{K})$.
Die Einheitsmatrix $\mathbb{I}_n$ ist gegeben durch

$$\mathbb{I}_n := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix} \in M_n(\mathbb{K}).$$

Mit $\mathrm{End}(V)$ wollen wir den Raum aller Endomorphismen $\alpha\colon V \to V$ bezeichnen. Die Begriffe *lineare Abbildung* und *Endomorphismus* finden synonym Verwendung, sofern für die lineare Abbildung $\alpha\colon V \to W$ eben $V = W$ gilt.

2.2. Dimensionssätze

Die folgenden sogenannten Dimensionssätze werden uns im Laufe dieser Arbeit öfters unterkommen. Sie sind (üblicherweise) Bestandteil einführender Vorlesungen zur Linearen Algebra und ein Beweis dieser sollte bei Bedarf ohne weiteres auffindbar sein. Die folgenden Ausführungen der Dimensionssätze entstammen [Cap10].

Satz 2.2.1 (Dimensionssatz für lineare Abbildungen)**.** Seien V und W endlichdimensionale $\mathbb{K}$–Vektorräume und sei $f : V \to W$ eine lineare Abbildung. Dann ist

$$\dim(\mathrm{Im}(f)) = \dim(V) - \dim(\ker(f)).$$

Der Dimensionssatz für lineare Abbildungen liefert eine bedeutende Beziehung zwischen den Unterräumen $\mathrm{Im}(f) \subset W$, $\ker(f) \subset V$ und dem zu Grunde liegenden Vektorraum V. In seiner Bedeutung in Nichts nachstehend ist der folgende

Satz 2.2.2 (Dimensionssatz für Summen)**.** Für zwei Teilräume W_1 und W_2 eines $\mathbb{K}$–Vektorraumes V sind folgende Bedingungen äquivalent:

(i) $V = W_1 \oplus W_2$

(ii) $\dim(V) = \dim(W_1) + \dim(W_2)$ und $W_1 \cap W_2 = \{0\}$.

(iii) Jedes Element $v \in V$ lässt sich eindeutig in der Form $v = w_1 + w_2$ für $w_1 \in W_1$ und $w_2 \in W_2$ schreiben.

2.3. Charakteristisches Polynom

Literarische Grundlage dieses kurzen Abschnitts bietet [Fis10].

Für eine lineare Abbildung $\alpha : V \to V$ ist $\alpha(v) = \lambda \cdot v$ gleichbedeutend zu $(\alpha - \lambda \cdot \mathrm{id})(v) = 0$ auf Grund der punktweise definierten Addition und Multiplikation (mit einem Skalar) auf $\mathrm{End}(V)$. Das heißt, λ ist genau dann ein Eigenwert von α, wenn es einen Vektor $v \neq 0$ gibt, der letztere Identität erfüllt. Das bedeutet, dass die Abbildung $(\alpha - \lambda \cdot \mathrm{id}_V)$ im Falle der Existenz eines Eigenvektors auf Grund des nicht-trivialen Kerns nicht injektiv und somit nicht bijektiv ist (für Endomorphismen ist Injektivität gleichbedeutend mit Bijektivität) und umgekehrt. Dies wiederum ist äquivalent zu $\det(\alpha - \lambda \cdot \mathrm{id}_V) = 0$.
Somit ist $\mathrm{Eig}(\alpha; \lambda) = \ker(\alpha - \lambda \,\mathrm{id})$, wobei $\mathrm{Eig}(\alpha; \lambda) := \{v \in V : \alpha(v) = \lambda v\}$.

Der Ausdruck $p_\alpha := \det(\alpha - \lambda \,\mathrm{id}_V)$ bzw. $p_\alpha := \det(M_{\mathcal{B}}(\alpha) - \lambda \mathbb{I}_n)$ wird als *charakteristisches Polynom* von α bezeichnet.

Bemerkung. Die Überlegungen zu Determinanten stimmen auch für $M_n(\mathbb{K}[t])$ (d.h. die Elemente der quadratischen Matrix sind aus dem Raum der Polynome über einem Körper $\mathbb{K}$), sodass die Berechnung von $\det(M_{\mathcal{B}}(\alpha) - \lambda \mathbb{I}_n)$ für λ als Variable keine Probleme ergibt.

2.4. Polynome

Es ist nicht schwer zu zeigen, dass die Eigenwerte zu einer linearen Abbildung α genau die Nullstellen des zugehörigen charakteristischen Polynoms p_α sind. Unter anderem aus diesem Grund wollen wir noch einiges Wissen über Polynome herausarbeiten.
Die hier behandelte Theorie ist hauptsächlich [Bos14] und [Fis13], sowie [Cap11] entnommen.

Satz 2.4.1. Division mit Rest. Sei $\mathbb{K}$ ein Körper und seien $f, g \in \mathbb{K}[t]$ gegeben, wobei $g \neq 0$. Dann gibt es eindeutig bestimmte $q, r \in \mathbb{K}[t]$ mit

$$f = q \cdot g + r \qquad \text{und} \quad \deg(r) < \deg(g)$$

Beweis. Zunächst den Nachweis der *Existenz*: Sei

$$f = a_n \cdot t^n + \cdots + a_0 \text{ und } g = b_m \cdot t^m + \cdots + b_0 \text{ mit } n = deg(f), m = deg(g) \geq 0.$$

Falls $n < m$ ist, erreicht man die gewünschte Darstellung durch Setzen von $q = 0$ und $r = f$, also $f = 0 \cdot g + f$.
Für $n \geq m$ konstruiert man die Polynome

$$q_1 := \frac{a_n}{b_m} \cdot t^{n-m} \quad \text{und} \quad f_1 := f - q_1 \cdot g.$$

Da dann $\deg(q_1 \cdot g) = \deg(q_1) + \deg(g) = n - m + m = n$ gilt (da $\mathbb{K}$ insbesondere ein Integritätsbereich ist; sonst würde „nur“ $\deg(q_1 \cdot g) \leq \deg(q_1) + \deg(g)$ gelten), sowie $(q_1 \cdot g)$ und f den gleichen führenden Koeffizienten a_n besitzen, ergibt sich

$$f = q_1 \cdot g + f_1 \qquad \text{mit} \quad \deg(f_1) < \deg(f).$$

Ist $\deg(f_1) < \deg(g)$ dann erfüllt man mit $q = q_1$ und $r = f_1$ obige Behauptung.
Falls aber $\deg(f_1) \geq \deg(g)$ gilt, kann das gerade beschriebene Vorgehen mit f_1 durchgeführt werden, d.h. wir wählen analog zum vorherigen Schritt ein $q_2 \in \mathbb{K}[t]$ derart, dass

$$f_1 = q_2 \cdot g + f_2 \qquad \text{mit} \quad \deg(f_2) < \deg(f_1)$$

ist. Iterativ weiterführend bricht dieser Prozess nach endlich vielen Schritten ab und wir gelangen letztlich zu einer Zerlegung

$$f_{k-1} = q_k \cdot g + f_k \qquad \text{mit} \quad \deg(f_1) < \deg(g).$$

Somit ergeben sich Polynome $q_1, \ldots, q_k$, sodass

$$f = (q_1 + \cdots + q_k) \cdot g + f_k$$

ist, und erhalten mit $q = (q_1 + \cdots + q_k)$ und $r = f_k$ die gewünschte Darstellung.

Dies ist übrigens spätestens für $k = n - m + 1$ der Fall, da sich der Grad von f_j (für ein beliebiges f_j im gerade beschriebenen Verfahren) in jedem Schritt um mindestens 1 verringert.

Eindeutigkeit: Sei eine Zerlegung

$$f = q \cdot g + r = q' \cdot g + r'$$

gegeben, dann folgt

$$r - r' = (q - q') \cdot g.$$

Da einerseits dem Satz entsprechend $\deg(r), \deg(r') < \deg(g)$ ist, also $\deg(r - r') < \deg(g)$, andererseits aber auch

$$\deg(r - r') = \deg(q - q') + \deg(g)$$

gilt, muss $q - q' = 0$ und dann auch $r - r' = 0$ sein. □

Aus der Division mit Rest - auch als *Euklidischer Algorithmus* bezeichnet - folgt das

Korollar 2.4.2. Sei $p \in \mathbb{K}[t]$ ein Polynom, $p \neq 0$. Dann ist eine Zahl $\lambda \in \mathbb{K}$ genau dann eine Nullstelle von p, wenn das Polynom $(t - \lambda)$ das Polynom p teilt, d.h. wenn ein Polynom $q \in \mathbb{K}[t]$ mit $p = (t - \lambda) \cdot q$ existiert.

Beweis. Die Division mit Rest auf p und $(t - \lambda)$ angewandt ergibt

$$p = q \cdot (t - \lambda) + r \qquad \text{mit} \quad 0 \leq \deg(r) < 1 = \deg((t - \lambda)).$$

Das Restpolynom $r \in \mathbb{K}$ ist somit lediglich ein Skalar. Daraus ergibt sich

$$p(\lambda) = (\lambda - \lambda) + r = r,$$

d.h. wenn λ eine Nullstelle ist, gilt $r = 0$ und umgekehrt. □

Ohne große Mühe ableitbar ist dann auch das

Korollar 2.4.3. Sind $\lambda_1, \ldots, \lambda_m$ paarweise verschieden mit $p(\lambda_j) = 0$ für $j \in \{1, \ldots, m\}$ $(p \in \mathbb{K}[t])$, so existiert folgende Faktorisierung

$$p = (t - \lambda_1) \cdot \cdots \cdot (t - \lambda_m) \cdot q$$

mit $q \in \mathbb{K}[t]$ und $0 \leq \deg(q) < \deg(f) - m$. Dies bedeutet insbesondere, dass die Anzahl m der verschiedenen Nullstellen kleiner oder gleich dem Grad von p sein muss: $m \leq n = \deg(f)$

Gänzlich beantwortet wird die Frage nach der Anzahl der Nullstellen eines Polynoms dann durch den *Fundamentalsatz*, der in erster Linie auch die Existenz von Nullstellen (zumindest im algebraischen Abschluss) sichert.

Satz 2.4.4. Fundamentalsatz der Algebra. Jedes nichtkonstante Polynom mit komplexen Koeffizienten hat eine komplexe Nullstelle.

Ein Beweis für den *Fundamentalsatz der Algebra* ist in [Art98, S. 603] oder in [KM09, S. 285] (allerdings auf Basis der Galois-Theorie) zu finden. □

Aus dem Fundamentalsatz geht hervor, dass jedes komplexe Polynom $p \in \mathbb{C}[t]$ in Linearfaktoren (Polynome ersten Grades) zerfällt und genau $\deg(p)$ viele Nullstellen hat, sofern man die *Vielfachheit* - d.h. wie oft ein Linearfaktor der Form $(t - \lambda)$ multiplikativ in p enthalten ist - der Nullstellen berücksichtigt.

Für den reellen Fall gilt das

Korollar 2.4.5. Sei $p \in \mathbb{R}[t]$ ein reelles Polynom vom Grad $n \geq 1$. Dann gibt es eine Zahl $k \leq n$, sodass $n-k$ gerade ist, k (nicht notwendig verschiedene) Zahlen $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ und $l := \frac{n-k}{2}$ monische Polynome $q_1, \ldots, q_l \in \mathbb{R}[t]$ mit $\deg(q_j) = 2$, sodass

$$p = a_n \cdot (t - \lambda_1) \cdot \ldots \cdot (t - \lambda_k) \cdot q_1 \cdot \ldots \cdot q_l.$$

Beweis. Wir fassen $p \in \mathbb{R}[t]$ als komplexes Polynom auf und obiges Korollar liefert uns somit n (nicht notwendig verschiedene) komplexe Zahlen $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, sodass

$$p = a_n \cdot (t - \lambda_1) \cdot \ldots \cdot (t - \lambda_n)$$

gilt, wobei $a_n \in \mathbb{R}$ als führender Koeffizient von p reell ist.

Ist $\lambda \in \mathbb{C}$ eine Nullstelle von p, dann gilt

$$0 = a_n \cdot \lambda^n + \cdots + a_1\lambda + a_0 \tag{2.1}$$

mit a_i als jeweils entsprechenden reellen Koeffizient des Polynoms. Dann gilt $a_i = \bar{a}_i$, wobei $\bar{a}_i$ die komplexe Konjugation von a_i bezeichne. Die komplexe Konjugation ist mit allen Körperoperationen auf $\mathbb{C}$ verträglich (d.h. die Konjugation der rechten Seite der Gleichung (2.1) wirkt sich auf jeden Summanden und Faktor separat aus). Darum ist

$$0 = \bar{0} = \bar{a}_n \cdot \bar{\lambda}^n + \cdots + \bar{a}_1\bar{\lambda} + \bar{a}_0 = a_n \cdot \bar{\lambda}^n + \cdots + a_1\bar{\lambda} + a_0.$$

Demzufolge ist mit λ automatisch auch $\bar{\lambda}$ eine Nullstelle von p. Andererseits folgt $(t-\lambda) \cdot (t-\bar{\lambda}) = t^2 - (\lambda + \bar{\lambda}) \cdot t + \lambda \cdot \bar{\lambda}$, und $\lambda + \bar{\lambda} = 2\mathrm{Re}(\lambda) \in \mathbb{R}$ und $\lambda \cdot \bar{\lambda} = |\lambda|^2$, also dass dieses Polynom reelle Koeffizienten hat. □

2.5. Potenzen eines Endomorphismus

Der folgende Abschnitt beruht überwiegend auf [Brö03].

Auch zur Vorbereitung der *Jordanschen Normalform* wollen wir uns hier mit Potenzen von Endomorphismen beschäftigen, ohne an dieser Stelle schon ihre Bedeutsamkeit bzw. ihre Tragweite zeigen zu können.

Zur Auffrischung: die Multiplikation einer Matrix $A = M_{\mathcal{B}}(\alpha)$ mit sich selbst entspricht der Komposition (Hintereinanderausführung) des zugehörigen linearen Endomorphismus, d.h. zum Beispiel A^3 entspricht $\alpha \circ \alpha \circ \alpha$. Für diese Komposition schreiben wir einfach α^3.
Eine Kette solcher Kompositionen betrachten wir nun folgend für einen „eingesetzten“ Vektor v.

Ist $v \in \ker(\alpha)$ und $a_0 = 0$, dann gilt:

$$a_r \cdot \alpha^r(v) + \cdots + a_1 \cdot \alpha(v) + a_0 \cdot v = 0, \tag{2.2}$$

ist dabei mindestens ein $a_l \neq 0$ $f\ddot{u}r$ $l \in \{1, \ldots, r\}$ und $v \neq 0$, dann ist (2.2) eine nicht-triviale Linearkombination der Null und die Vektoren $v, \alpha(v), \ldots, \alpha^r(v)$ sind somit linear abhängig. Ist $a_r \neq 0$ (bzw. wir können dies so wählen), so folgt, dass das sogenannte lineare Erzeugnis

$$W = \langle \{v, \alpha(v), \ldots, \alpha^{r-1}(v)\} \rangle = \{\sum_{i=0}^{r-1} a_i \cdot \alpha^i(v) : a_i \in \mathbb{K}\}$$

(mit $\alpha^0 := \mathrm{id}_V$) ein offensichtlich α-invarianter Teilraum von V mit $1 \leq \dim(W) \leq r$ ist.

Aus diesem Resultat können wir sogleich eine nützliche Erkenntnis ableiten. Zuvor benötigen wir aber noch eine

Definition. Man nennt $\alpha \in \mathrm{End}(V)$ **nilpotent**, wenn $\alpha^k = 0$ für ein $k \geq 1$, d.h. wenn a^k jeden Vektor $v \in V$ auf Null abbildet.

Folgerung. Ist $\dim(W) = m$ und φ ein nilpotenter Endomorphismus von W, dann ist $\varphi^m = 0$.

Noch ein Wort zu den Ausdrücken der Form (2.2): diese Linearkombination von Hintereinanderausführungen erinnert sehr stark an ein Polynom

$$p(t) = a_r \cdot t^r + \cdots + a_1 \cdot t + a_0 \in \mathbb{K}[t],$$

in dem t durch α ersetzt wurde.
Diese Kombination aus Kompositionen des linearen Endomorphismus α,

$$p(\alpha) = a_r \cdot \alpha^r + \cdots + a_1 \cdot \alpha + a_0 \cdot \mathrm{id} \in \mathrm{End}(V),$$

ist dann, wenn auch am Anfang vielleicht ungewohnt, selbst wieder eine lineare Abbildung.

Bemerkung. - Es gibt Koeffizienten $a_r, \ldots, a_0$ derart, dass $p(\alpha) = 0$ ist, in diesem Falle ist die Gleichung (2.2) für jedes beliebige $v \in V$ erfüllt.

- Hat man eine Koeffizienten-Kombination mit $p(\alpha) = 0$ gefunden, so gilt auch für alle Vielfachen dieses Polynoms p, etwa $q = r \cdot p$, dass $q(\alpha) = 0$ ist.
- Die Menge $\mathbb{K}[\alpha] = \{p(\alpha) : p \in \mathbb{K}[t]\} \subset \mathrm{End}(V)$ ist ein kommutativer Unterring des (nicht kommutativen) Ringes $\mathrm{End}(V)$, und der Kern von $p \mapsto p(\alpha), \mathbb{K}[t] \to \mathbb{K}[\alpha]$ ist gegeben durch

$$I_\alpha := \{p \in \mathbb{K}[t] : p(\alpha) = 0\} \subset \mathbb{K}[t].$$

- Da I_α ein sogenanntes *Hauptideal* ist, sind alle Polynome $q \in I_\alpha$ von der Gestalt $q = r \cdot m_\alpha$ für ein (eindeutiges) Polynom $m_\alpha \in I_\alpha$ und passendes $r \in \mathbb{K}[t]$. Das Polynom m_α nennt man **Minimalpolynom**.[1]
- Das wohl bekannteste Exemplar mit $p(\alpha) = 0$ ist das *charakteristische Polynom* p_α. Dass das charakteristische Polynom p_α in I_α liegt, ist genau die Aussage des Satzes von Caley-Hamilton.

Satz 2.5.1. Satz von Caley-Hamilton. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum, $\alpha \in \mathrm{End}(V)$ und $p_\alpha \in \mathbb{K}[t]$ sein charakteristisches Polynom. Dann ist

$$p_\alpha(\alpha) = 0 \in \mathrm{End}(V).$$

Natürlich gilt für die zu α gehörende Matrix A:

$$p_\alpha(A) = 0 \in M_n(\mathbb{K}).$$

Ein Beweis für den Satz von Caley-Hamilton ist in [Fis10] ab S. 251 sogar in zwei Varianten präsentiert. □

[1] Diese Definition ist für die folgenden Kapitel genügend, eine saubere Definition sowie Begründung dieser Behauptung folgt dann in 5.1.

Teil I.

Gängige Pfade zur Jordanschen Normalform

3. Besondere Darstellungen linearer Abbildungen

Wie in der Einleitung angesprochen, werden in diesem Kapitel herausragende Matrixdarstellungen von linearen Endomorphismen betrachtet. Herausragend in dem Sinne, dass es - wie auch generell in der gesamten Arbeit - unser Wunsch ist eine spezielle Basis für einen Vektorraum zu finden, sodass die Einträge der Matrix bezüglich eines Basiswechsels mit der neuen Basis ein geordneteres Zahlenbild (im Bezug auf die Darstellung der Matrix mit der ursprünglichen Basis) ergeben.
Diese Überlegungen sind aber weniger auf Grund der resultierenden „hübschen“ Form einer Matrix von Bedeutung, sondern dienen der Vorbereitung zur *Jordanschen Normalform* (hauptsächlich bezüglich des ersten Ansatzes).

3.1. Blockform

In diesem Abschnitt geht es weniger darum, wie man zu einer Matrix in Blockform gelangt, sondern vielmehr darum, was sich von einer Matrix in Blockform ableiten lässt. Den literarisch-fachlichen Hintergrund liefert hierbei [Cap11].

Wenn wir von einer *Blockmatrix* M sprechen meinen wir damit eine Matrix $M = (m_{ij}) \in M_n(\mathbb{K})$, „sodass für ein $k < n$ gilt, dass $m_{ij} = 0$ ist, sofern $i > k$ und $j \leq k$ ist, also in den ersten k Spalten Eintragungen ungleich Null nur in den ersten k Zeilen vorkommen“[1].
Die Matrix M hat also die Form $M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ mit $A \in M_k(\mathbb{K}), B \in M_{k,n-k}(\mathbb{K})$ und $C \in M_{n-k}(\mathbb{K})$. Die Teilmatrizen A und C sind quadratisch, während B rechteckig ist.

[1] [Cap11, Seite 36].

Definition. Sei $\alpha\colon V \to V$ ein Endomorphismus und $W \subset V$ ein Untervektorraum. W heißt α-**invariant**, wenn $\alpha(W) \subset W$.

Proposition 3.1.1. Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$ und $\alpha : V \to V$ ein linearer Endomorphismus. Dann gilt: Es gibt genau dann einen k-dimensionalen α-invarianten Unterraum $W \subset V$, wenn bezüglich einer geeigneten Basis $\mathcal{B}$ von V die Matrixdarstellung von α eine Blockform $M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ mit $A \in M_k(\mathbb{K}), B \in M_{k,n-k}(\mathbb{K})$ und $C \in M_{n-k}(\mathbb{K})$ besitzt. Ist das der Fall, dann ist A eine Matrixdarstellung der Einschränkung $\alpha|_W : W \to W$.
Außerdem ist C eine Matrixdarstellung der induzierten Abbildung $\underline{\alpha} : V/W \to V/W$.

Beweis. Ist $W \subset V$ ein k-dimensionaler α-invarianter Teilraum, dann können wir eine Basis $\mathcal{B}' := \{v_1, \dots, v_k\}$ von W zu einer Basis $\mathcal{B} := \{v_1, \dots, v_n\}$ von V erweitern. Da W α-invariant ist, gilt insbesondere $\alpha(v_j) \in W$ für $j = 1, \dots, k$, es existieren somit für jedes Bild eines Basisvektors $\alpha(v_j)$ $(j = 1, \dots, k)$ Koeffizienten $a_{ij} \in \mathbb{K}$, sodass $\alpha(v_j) = a_{1j}v_1 + \cdots + a_{kj}v_k$ gegeben ist. Mit Proposition 2.1.1 wissen wir, dass die Spalten von $M_{\mathcal{B}}(\alpha)$ durch die Bilder $[\alpha(v_j)]_{\mathcal{B}}$ für $j = 1, \dots, n$ gegeben sind. Das bedeutet die Koeffizienten obiger Linearkombination ergeben die gewünschte Teilmatrix (den Block $A \in M_k(\mathbb{K})$). Dieser Block A entspricht dann $M_{\mathcal{B}'}(\alpha|_W)$.
Den Beweis, dass es sich bei C um die Matrixdarstellung der induzierten Abbildung handelt verschieben wir passenderweise an das Ende des Abschnitts 7.1.

Ist umgekehrt $\mathcal{B} = \{v_1, \dots, v_n\}$ eine Basis von V, sodass $M_{\mathcal{B}}(\alpha)$ die angegebene Blockform hat, dann sei $W := \langle v_1, \dots, v_k \rangle$ der von der ersten k Basisvektoren erzeugte Unterraum. Wegen der Blockform von $M_{\mathcal{B}}(\alpha)$ ist $\alpha(v_j) \in W$ für $j = 1, \dots, k$ und somit ist $\alpha(W) \subset W$, also ist W wie gewünscht α-invariant. □

Proposition 3.1.2. Für eine Matrix M in Blockform, also $M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$, gilt

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det(A) \cdot \det(C).$$

Beweis. Durch Induktion nach der Blockgröße k der Teil-Matrix A: Für k=1 ist $A = a_{11}$, B der Rest der ersten Zeile und C gerade die Matrix die durch Streichung der ersten Zeile und ersten Spalte in der gesamten Matrix M entsteht. Vergegenwärtigt man sich den Entwicklungssatz für Determinanten, dann folgt die Behauptung.
Sei nun $k \geq 2$. A_{ij} bezeichnet die Matrix, die durch Streichung der i-ten Zeile und

j-ten Spalte aus der ursprünglichen Matrix $A = (a_{ij})$ mit $1 \leq i, j \leq k$ entsteht, und B_j diejenige Matrix, die durch Streichung der j-ten Zeile aus der Matrix B entsteht. Entwickeln der Determinante nach der ersten Spalte ergibt dann

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \sum_{j=1}^{k} (-1)^{1+j} \cdot a_{j1} \cdot \underbrace{\det \begin{pmatrix} A_{j1} & B_j \\ 0 & C \end{pmatrix}}_{= \det(A_{j1}) \cdot \det(C)}.$$

Somit ergibt sich $(\sum_{j=1}^{k} (-1)^{1+j} \cdot a_{j1} \cdot \det(A_{j1})) \cdot \det(C) = \det(A) \cdot \det(C)$. □

Korollar 3.1.3. Sei V endlichdimensionaler Vektorraum über einem Körper $\mathbb{K}$, $\alpha : V \to V$ eine lineare Abbildung, $W \in V$ ein α-invarianter Teilraum, $\alpha|_W : W \to W$ die Einschränkung von α und $\underline{\alpha} : V/W \to V/W$ die *induzierte Abbildung*, dann gilt für die charakteristischen Polynome $p_\alpha = p_{\alpha|_W} \cdot p_{\underline{\alpha}}$.

Beweis. Da sich für eine Basis $\mathcal{B}'$ von W und einer Erweiterung dieser zu einer Basis $\mathcal{B}$ von V

$$M_{\mathcal{B}}(\alpha) = \left(\begin{array}{c|c} M_{\mathcal{B}'}(\alpha|_W) & * \\ \hline 0 & C \end{array} \right)$$

ergibt, folgt mit Proposition 3.1.2, dass $p_\alpha = p_{\alpha|_W} \cdot p_C$, da $\det(M_{\mathcal{B}}(\alpha) - t \cdot \mathbb{I}_n) = \det(M_{\mathcal{B}'}(\alpha|_W) - t \cdot \mathbb{I}_{\dim W}) \cdot \det(C - t \cdot \mathbb{I}_{n-\dim W})$ ist. Nach Proposition 3.1.1 ist $p_C = p_{\underline{\alpha}}$. □

Die Tatsache, dass es sich bei $p_{\underline{\alpha}}$ um das charakteristische Polynom der induzierten Abbildung handelt bzw. was überhaupt unter der *induzierten Abbildung* verstanden wird, ist für die nächsten Kapitel irrelevant.
Aus Kapitel 7 wird dann der Zusammenhang zwischen Vektorräumen und Faktorringen (bzw. zwischen einer linearen Abbildung und der *induzierten Abbildung*) hervorgehen.

3.2. Trigonalform

Bezugsquelle dieses Unterpunktes ist [Fis10].

Beschäftigen wir uns zunächst mit der Frage nach den Bedingungen zur Überführung einer Matrix in eine obere Dreiecksmatrix. Ist dies möglich, spricht man von einer *Trigonalisierung* einer Matrix *M*.

Bemerkung. Eine rechte obere Dreiecksmatrix nennt man auch „Darstellung in Schur'scher Normalform."[2]

[2] [Mut06, Seite 200].

Zunächst eine kleine Spezifikation von Korollar 3.1.3, die wir für den kommenden Satz benötigen werden.

Bemerkung. Sei $\alpha\colon V \to V$ linear und $W \subset V$ ein *α-invarianter Unterraum*, so ist $p_{\alpha|_W}$ ein Teiler von p_α.

Beweis. Dies wurde bereits mit Korollar 3.1.3 gezeigt, sofern wir uns nicht am Begriff der *induzierten Abbildung* stoßen, der, wie erwähnt, später erläutert wird. □

Definition. Unter einer **Fahne** (W_r) in einem n-dimensionalen Vektorraum V versteht man eine Kette $\{0\} = W_0 =\subset W_1 \subset \cdots \subset W_n = V$ mit $\dim(W_r) = r$.

Ist $\alpha \in \mathrm{End}(V)$, so heißt eine Fahne α-invariant, wenn

$$\alpha(V_r) \subset V_r \quad \text{für alle } r \in \{0, \ldots, n\}.$$

Satz 3.2.1. Trigonalisierungssatz. Für einen Endomorphismus α eines n-dimensionalen $\mathbb{K}$-Vektorraumes sind folgende Bedingungen äquivalent:

(i) α ist trigonalisierbar, d.h. es gibt eine Basis $\mathcal{B}$ für V, sodass $M_\mathcal{B}(F)$ eine obere Dreiecksmatrix ist.

(ii) Es gibt α-invariante Teilräume $W_0 \subset W_1 \subset \ldots W_{n-1} \subset V$ mit $\dim(W_i) = i$, also eine α-invariante Fahne.

(iii) Das charakteristische Polynom p_α schreibt sich folgenderweise:

$$p_\alpha = \pm(t - \lambda_1) \cdot \ldots \cdot (t - \lambda_n) \qquad \text{mit } \lambda_1, \ldots, \lambda_n \in \mathbb{K},$$

es zerfällt also in Linearfaktoren, d.h. alle Eigenwerte von α liegen in $\mathbb{K}$.

Beweis. (i) $\Leftrightarrow$ (ii): Setze $W_r = \langle (v_1, \ldots, v_r) \rangle$ ($v_1, \ldots, v_r$ linear unabhängig) für $r \in \{1, \ldots, n\}$.
(i) $\Rightarrow$ (iii): $A = (a_{ij}) = M_\mathcal{B}(\alpha)$ ist eine obere Dreiecksmatrix, daher

$$p_\alpha = \det(M_\mathcal{B}(\alpha) - t \cdot \mathbb{I}_n) = (a_{11} - t) \cdot \ldots \cdot (a_{nn} - t).$$

(iii) $\Rightarrow$ (i) mittels Induktion: Für $n = 0, 1$ ist nichts zu zeigen, sei also $n \geq 2$. Man wähle einen Eigenvektor v_1 zu λ_1 und ergänze diesen zu einer Basis $\mathcal{B} = \{v_1, w_2, \ldots, w_n\}$ von V. Dann ist

$$V = V_1 \oplus W \quad \text{mit } V_1 := \langle (v_1) \rangle \text{ und } W := \langle (w_2, \ldots, w_n) \rangle,$$

und

$$M_{\mathcal{B}}(\alpha) = \begin{pmatrix} \lambda_1 & a_{12} & \dots & a_{1n} \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}.$$

Wären die Einträge $a_{12}, \dots, a_{1n}$ allesamt gleich Null, hätten wir eine Basis gefunden aus der ersichtlich wird, dass der (Unter-)Vektorraum W invariant ist ($M_{\mathcal{B}}(\alpha)$ hätte Blockdiagonalform, siehe nächsten Abschnitt, insbesondere Satz 3.3.1). Die Einträge werden im Allgemeinen aber nicht gleich Null sein, da W im Allgemeinen nicht invariant ist.
Wir wollen nun einen Endomorphismus γ finden, der dem Block B entspricht. Mit einem derartigen Endomorphismus $\gamma\colon W \to W$ würde nämlich die Induktionsvoraussetzung „greifen“, da $\dim(W) = n - 1$. Dies ist nun auch nicht so schwierig: Wir definieren die linearen Abbildungen $\beta\colon W \to V_1$, mit $\beta(w_j) = a_{1j}v_1$ und $\gamma\colon W \to W$, mit $\gamma(w_j) = a_{2j}w_2 + \dots + a_{nj}w_n$. Dann ist $\alpha(w_j) = \beta(w_j) + \gamma(w_j)$ für $j \in \{2, \dots, n\}$ und somit $\alpha(w) = \beta(w) + \gamma(w)$ für alle $w \in W$.
Alle Bilder von Vektoren $w \in W$ werden also in ihren Anteil aus V_1 und W zerlegt. Aus Proposition 3.1.2 wissen wir dann, dass

$$p_\alpha = (\lambda_1 - t) \cdot p_\gamma, \quad \text{also } p_\gamma = (\lambda_2 - t) \cdot \ldots \cdot (\lambda_n - t).$$

Laut Induktionsvoraussetzung erhalten wir eine γ-invariante Fahne

$$\{0\} = W_0 \subset W_1 \subset \dots \subset W_{n-1} = W.$$

Wird nun noch $V_r := V_1 + W_{r-1}$ $(1 \leq r \leq n)$ gesetzt, erhalten wir die gewünschte α-invariante Fahne (für V, wobei $V_0 = \{0\}$ und $V_n = V$): jedes $v \in V_r$ kann als $\kappa \cdot v_1 + w$, mit geeignetem $\kappa \in \mathbb{K}$ und geeignetem $w \in W_{r-1}$ geschrieben werden; daher gilt für alle $v \in V_r : \alpha(\kappa v_1 + w) = \underbrace{\lambda_1 \kappa v_1}_{\in V_1} + \underbrace{\beta(w)}_{\in V_1} + \underbrace{\gamma(w)}_{\in W_{r-1}} \in V_r$. □

3.3. Blockdiagonalform

Wir leiten nun den Zusammenhang zwischen direkten Summen α-invarianter Teilräume und Blockdiagonalform - in Anlehnung an die vorigen Abschnitte - her. Der nachfolgende Satz und sein Beweis stützen sich auf [Cap11].

Satz 3.3.1. Sei $V = W_1 \oplus W_2$ eine Zerlegung eines n-dimensionalen $\mathbb{K}$-Vektorraumes in eine direkte Summe von Teilräumen und $\alpha : V \to V$ eine lineare Abbildung. Dann gilt: Eine α-invariante Summenzerlegung $V = W_1 \oplus W_2$ mit $\dim(W_1) = k$ existiert genau dann, wenn eine Basis $\mathcal{B}$ für V existiert, sodass

$$M_{\mathcal{B}}(\alpha) = \begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix} \text{ („Blockdiagonalform“)}$$

mit $A \in M_k(\mathbb{K})$ und $C \in M_{n-k}(\mathbb{K})$ ist.
In diesem Fall ist A eine Matrixdarstellung von $\alpha|_{W_1}$ und C eine Matrixdarstellung von $\alpha|_{W_2}$.

Beweis. Sei zunächst $V = W_1 \oplus W_2$ eine α-invariante Zerlegung mit $\dim(W_1) = k$. Dann erhalten wir für eine gewählte Basis $\mathcal{B}' = \{v_1, \ldots, v_k\}$ für W_1 und $\mathcal{B}'' = \{v_{k+1}, \ldots, v_n\}$ für W_2 eine Basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ für V (Dies liefert im Grunde der Dimensionssatz 2.2.2). Analog zur Überlegung des Beweises von Proposition 3.1.1 folgt, dass die Vektoren $[\alpha(v_i)]_{\mathcal{B}}$ für $i = 1, \ldots, k$ nur in den ersten k Zeilen Eintragungen ungleich Null und für $i = k+1, \ldots, n$ erst ab der $(k+1)$-ten Zeile Eintragungen ungleich Null besitzen. Es ist somit $A = M_{\mathcal{B}'}(\alpha|_{W_1})$ und $B = M_{\mathcal{B}''}(\alpha|_{W_2})$.
Ist umgekehrt $\mathcal{B} = \{v_1, \ldots, v_n\}$ eine Basis von V, sodass $M_{\mathcal{B}}(\alpha)$ die angegebene Blockform hat. Dann verfahren wir wieder analog zum Beweis von Proposition 3.1.1: wir setzen $W_1 = \langle v_1, \ldots, v_k \rangle$ und $W_2 = \langle v_{k+1}, \ldots, v_n \rangle$. Aus der Blockform von $M_{\mathcal{B}}(\alpha)$ lesen wir ab, dass W_1 und W_2 α-invariant sind. Auf Grund der Wahl von W_1 und W_2 kann natürlich jedes $v \in V$ in der Form von $v = w_1 + w_2$ dargestellt werden, also ist $V = W_1 + W_2$. Da $\dim(V) = \dim(W_1) + \dim(W_2)$ gilt, folgt aus dem Dimensionssatz 2.2.2, dass $W_1 \cap W_2 = 0$ und damit $V = W_1 \oplus W_2$. □

Mit Proposition 3.1.2 folgt dann auch sofort, dass für das charakteristische Polynom p_α die Identität

$$p_\alpha = p_{\alpha|_{W_1}} \cdot p_{\alpha|_{W_2}}$$

gilt.

3.4. Diagonalform

Besitzt eine Matrix A Einträge ungleich Null nur auf der Hauptdiagonale, so heißt es diese befindet sich in *Diagonalform*. Da die *Jordansche Normalform* vor allem dann gewinnbringend ist[3], wenn die Bedingungen der Diagonalisierbarkeit einer Matrix B, d.h. das Überführen von B in eine Diagonalmatrix, nicht gegeben sind, wollen wir uns hier mit eben diesen Bedingungen auseinandersetzen.

Fachliche Quelle hierbei stellt [Fis10] dar.

Der Vektorraum aller Eigenvektoren eines Endomorphismus α bezüglich des Eigenwerts λ, auch *Eigenraum* genannt, wird mit $\mathrm{Eig}(\alpha;\lambda)$ bezeichnet.

Lemma 3.4.1. Ist λ ein Eigenwert von α, so gilt $1 \leq \dim(\mathrm{Eig}(\alpha;\lambda)) \leq \mu(p_\alpha;\lambda)$, wobei $\mu(p_\alpha;\lambda)$ die algebraische Vielfachheit von λ in p_α beschreibt.

Beweis. Sei $v_1,\ldots,v_s$ eine Basis von $\mathrm{Eig}(\alpha;\lambda)$. Es gibt immer mindestens einen (dann sowieso linear unabhängigen) Eigenvektor v_1 zu λ, also $s \geq 1$.
Ergänzt man die Basis von $\mathrm{Eig}(\alpha;\lambda)$ zu einer Basis

$$\mathcal{B} = \{v_1,\ldots,v_s,v_{s+1},\ldots,v_n\}$$

von V, dann ist

$$M_{\mathcal{B}}(\alpha) = \left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & * \\ 0 & & \lambda & \\ \hline & 0 & & C \end{array}\right) \in M_n(\mathbb{K}).$$

Aus Proposition 3.1.2 folgt, dass $\det(M_{\mathcal{B}}(\alpha) - t\cdot\mathbb{I}_n) = p_\alpha = (\lambda - t)^s \cdot p_C$ ist. Es ist somit offensichtlich, dass λ auf jeden Fall s-fache Nullstelle des charakteristischen Polynoms p_α ist, also $\dim(\mathrm{Eig}(\alpha;\lambda)) = s \leq \mu(p_\alpha;\lambda)$. □

Satz 3.4.2. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und α ein linearer Endomorphismus auf V. Dann sind folgende Bedingungen äquivalent:

(i) α ist diagonalisierbar.

[3] Ist eine Matrix diagonalisierbar, so entspricht die Jordansche Normalform der Diagonalform.

(ii) a) Das charakteristische Polynom p_α zerfällt in Linearfaktoren (d.h. alle Eigenwerte liegen in $\mathbb{K}$) und
b) $\dim(\mathrm{Eig}(\alpha;\lambda)) = \mu(p_\alpha;\lambda)$ für alle Eigenwerte λ von α.

(iii) Sind $\lambda_1,\dots,\lambda_k$ die paarweise verschiedenen Eigenwerte von α, so ist

$$V = \mathrm{Eig}(\alpha;\lambda_1) \oplus \cdots \oplus \mathrm{Eig}(\alpha;\lambda_k).$$

Beweis. (i) $\Rightarrow$ (ii): Ist α diagonalisierbar, so existiert laut Definition eine Basis $\mathcal{B}$ für V aus Eigenvektoren. Diese wird entsprechend den verschiedenen Eigenwerten $\lambda_1,\dots,\lambda_k$ geordnet, wir betrachten dann für $1 \le i \le k$ eine Basis

$$\{v_1^{(i)},\dots v_{s_i}^{(i)}\} \quad \text{von } \mathrm{Eig}(\alpha;\lambda_i).$$

Setzen wir $r_i := \mu(p_\alpha;\lambda_i)$, so gilt

$$s_1 + \cdots + s_k = n \quad , \quad r_1 + \cdots + r_k = n \quad \text{und} \quad s_i \le r_i.$$

Dies erreichen wir nur mit $s_i = r_i$ für alle $i = 1,\dots,k$, also folgt (ii)b).
Weiters liefert die Bedingung (i) eine Basis $\mathcal{B}$ für die das charakteristische Polynom durch

$$p_\alpha = \det \begin{pmatrix} \lambda_1 - t & & 0 \\ & \ddots & \\ 0 & & \lambda_n - t \end{pmatrix} = (\lambda_1 - t) \cdot \ldots \cdot (\lambda_n - t)$$

gegeben ist und somit einen Beleg für (ii) a).
(ii) $\Rightarrow$ (iii): Wir wissen um die n linear unabhängigen Eigenvektoren, die eine Basis $\mathcal{B}$ von V liefern, und, dass es keinen Vektor $v \neq 0$ mit $v \in \mathrm{Eig}(\alpha;\lambda_i) \cap Eig(\alpha;\lambda_j)$ für $i \neq j$ gibt. Also muss

$$V = \mathrm{Eig}(\alpha;\lambda_1) \oplus \cdots \oplus \mathrm{Eig}(\alpha;\lambda_k)$$

sein.
(iii) $\Rightarrow$ (i): Für jedes $i \in \{1,\dots,k\}$ sei $\mathcal{B}_i = \{v_1^{(i)},\dots,v_{s_i}^{(i)}\}$ eine Basis von $\mathrm{Eig}(\alpha;\lambda_i)$. Dann ist

$$\mathcal{B} := \{v_1^{(1)},\dots,v_{s_1}^{(1)},\dots,v_1^{(k)},\dots,v_{s_k}^{(k)}\}$$

eine Basis von V bestehend aus Eigenvektoren, α also diagonalisierbar.
Insbesondere gilt $s_i = r_i$ für alle i und die Matrixdarstellung von α bezüglich $\mathcal{B}$ ist durch

$$M_{\mathcal{B}}(\alpha) = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_1 & & 0 & \\ & & & \ddots & & \\ & 0 & & & \lambda_k & \\ & & & & & \ddots \\ & & & & & & \lambda_k \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\end{matrix}}\right\} r_1\text{-mal} \\ \vdots \\ \left.\vphantom{\begin{matrix}a\\a\end{matrix}}\right\} r_k\text{-mal} \end{matrix}$$

gegeben. □

Wir haben somit die Bedingungen für die wohl ideale Matrixdarstellung einer entsprechenden linearen Abbildung hergeleitet.

3.5. Ausflug ins Reelle

Auch hier ist [Fis10] die fachliche Bezugsquelle.

Natürlich ist es auch interessant, wie eine *Normalform* über einem reellen Vektorraum V, bei dem die meisten Sätze auf Grund der algebraischen Unabgeschlossenheit von $\mathbb{R}$ nur bedingt greifen, aussehen könnte. Einen Schritt dorthin wagt der

Satz 3.5.1. Ist α ein Endomorphismus eines reellen Vektorraumes V, so gibt es eine Basis $\mathcal{B}$ derart, dass

$$M_{\mathcal{B}}(\alpha) = \begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & * & \\ & & \lambda_r & & \\ & & & \boxed{B_1} & \\ & 0 & & & \ddots \\ & & & & & \boxed{B_m} \end{pmatrix}$$

mit $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ und

$$B_i = \begin{pmatrix} 0 & -c_i \\ 1 & -b_i \end{pmatrix}$$

mit $b_i, c_i \in \mathbb{R}$ und $b_i^2 - 4c_i < 0$ für $i = 1, \ldots, m$ ist.

Zuvor noch das

Lemma 3.5.2. Jeder Endomorphismus α eines reellen Vektorraumes V mit $\dim(V) \geq 1$ hat einen invarianten Unterraum W mit $1 \leq \dim(W) \leq 2$.

Beweis. Das reelle charakteristische Polynom kann nach 2.4.5 wie folgt zerlegt werden:

$$p_\alpha = \pm(t-\lambda_1)\cdot\ldots\cdot(t-\lambda_r)\cdot q_1(t)\cdot\ldots\cdot q_m(t) \quad (\lambda_1,\ldots,\lambda_r \in \mathbb{R})$$

mit $\deg(q_i) = 2$ und nicht-reellen Nullstellen.
Ist $r \geq 1$, so gibt es einen Eigenvektor v_1 zu λ_1, der dann einen α-invarianten Unterraum $W = \langle v_1\rangle := \{\mu \cdot v_1 : \mu \in \mathbb{R}\}$ liefert.

Im Falle $r = 0$ wählen wir einen beliebigen Vektor $v \in V$, $v \neq 0$ und wissen nach dem Satz von Caley-Hamilton 2.5.1, dass

$$p_\alpha(\alpha)(v) = (q_1(\alpha)\circ\ldots\circ q_m(\alpha))(v) = 0$$

ist. Das bedeutet, dass eben jeder Vektor $v \in V$ im Laufe dieser Komposition auf das Nullelement abgebildet wird. Es gibt somit eine klare „Grenze“ $i \in \{1,\ldots,m\}$, sodass

$$(q_{i+1}(\alpha)\circ\ldots\circ q_m(\alpha))(v) \neq 0 \quad \text{aber} \quad (q_i(\alpha)\circ\ldots\circ q_m(\alpha))(v) = 0$$

gilt. Ist $i \lneqq m$ setzen wir $w := (q_{i+1}(\alpha)\circ\ldots\circ q_m(\alpha))(v)$, ansonsten $w := v$. Weiters setzen wir

$$W := \langle w, \alpha(w)\rangle.$$

Da nach Korollar 2.4.5 das Polynom q_i monisch ist, etwa $q_i(t) = t^2 + b_i\cdot t + c_i$ und der Wahl von w, ist

$$q_i(\alpha)(w) = \alpha(\alpha(w)) + b_i\cdot\alpha(w) + c_i\cdot w = 0, \text{ also } \alpha(\alpha(w)) = -b_i\cdot\alpha(w) - c_i\cdot w.$$

Damit gilt für ein Element $\tilde{w} \in W$

$$\alpha(\tilde{w}) = \alpha(b_i\cdot w + c_i\cdot\alpha(w)) = b_i\alpha(w) + c_i\underbrace{\alpha(\alpha(w))}_{\in W} \in W,$$

und somit die α-Invarianz von W. Weiters gilt

$$M_{\mathcal{B}}(\alpha|_W) = \begin{pmatrix} 0 & -c_i \\ 1 & -b_i \end{pmatrix}$$

und somit $p_{\alpha|_W} = q_i(t)$.

Wir müssen noch zeigen, dass w und $\alpha(w)$ linear unabhängig sind und damit $\dim(W) = 2$ gewährleisten. Die Gleichung $a \cdot w + b \cdot \alpha(w) = 0$ muss aber trivial sein, also $a = b = 0$, da ansonsten $\alpha(w) = -\frac{a}{b} \cdot w$ ein Widerspruch zur, in unserer Fallunterscheidung getroffenen, Annahme, dass p_α keine reellen Nullstellen und α somit keinen Eigenwert besitzt[4], wäre. □

Mit diesem Werkzeug können wir nun obigen Satz belegen.

Beweis Satz 3.5.1. Der $\mathbb{K}$-Vektorraum V wird unter zwei Etappen in α-invariante Summen zerlegt.
Zunächst spalten wir von $V := W_1$ schrittweise die zu $\lambda_1, \ldots, \lambda_r$ gehörigen Eigenräume ab und fassen diese jeweils zu V_i zusammen, W_{i+1} ist dabei jeweils ein Komplementärraum, somit ergibt sich

$$V = V_0 \oplus W_1 = V_1 \oplus W_2 = V_2 \oplus W_3 = \cdots = V_r \oplus W_{r+1}.$$

Nun können wir das vorige Lemma benutzen um ausgehend von W_{r+1} zweidimensionale α-invariante Unterräume abzubauen. Etwas genauer ausgedrückt: wir haben alle Eigenräume abgeschöpft und zu V_r vereinigt, somit haben wir mit Bezug auf den Beweis von 3.5.2 den Fall $r = 0$ bezüglich des Unterraumes W_{r+1}. Diesem Unterraum W_{r+1} können wir nach Lemma 3.5.2 einen weiteren zweidimensionalen Unterraum abschöpfen und zu V_r hinzufügen. Das Lemma sichert für jeden Unterraum W_{r+j} $(j = 3, 5, \ldots, 2m + 1)$ die Existenz eines zweidimensionalen Unterraums bis $W_{r+2m+1} = \{0\}$, wobei W_{r+j} jeweils den „restlichen“ Teilraum bezeichne, nachdem ein zweidimensionaler Unterraum abgeschöpft wurde. Diese Vorgehensweise also weiterführend erhalten wir die Zerlegung

$$\begin{aligned} V &= V_0 \oplus W_1 = V_1 \oplus W_2 = V_2 \oplus W_3 = \cdots = V_r \oplus W_{r+1} \\ &= V_{r+2} \oplus W_{r+3} = V_{r+4} \oplus W_{r+5} = \cdots = V_{r+2m} \oplus W_{r+2m+1}, \end{aligned}$$

die $\alpha(V_k) \subset V_k$ erfüllt; nach Konstruktion gilt $\dim(V_k) = k$ und

$$\{0\} = V_0 \subset V_1 \subset \cdots \subset V_r \subset V_{r+2} \subset \cdots \subset V_{r+2m} = V.$$

Wir sollten uns aber noch kurz Gedanken über die Abbildungen machen, die uns jeweils mit Lemma 3.5.2 die zweidimensionalen Teilräume gesichert haben. Sie sind gegeben durch die eindeutige Zerlegung von

$$\begin{aligned} \alpha|_{W_i} : W_i \to V = V_{i-1} \oplus W_i \quad \text{in} \quad \alpha|_{W_i} : \gamma_i + \alpha_i \\ \text{mit } \gamma_i : W_i \to V_{i-1} \quad \text{und} \quad \alpha_i : W_i \to W_i. \end{aligned}$$

Diese Zerlegung erinnert an den Beweis des Trigonalisierungssatzes 3.2.1. □

[4]Nicht-reelle Nullstellen von p_α werden nicht berücksichtigt, da wir über dem Skalarenkörper $\mathbb{R}$ operieren.

4. Die Jordansche Normalform - elementar

Wir sind nun soweit, uns einer Matrixdarstellung zu widmen, für den Fall, dass das charakteristische Polynom des zugehörigen Endomorphismus α in Linearfaktoren zerfällt, diese zugehörige Matrix aber trotzdem nicht diagonalisierbar ist (Bedingungen siehe 3.4.2). In diesem Fall besteht immer noch die Möglichkeit des Trigonalisierens. Unbefriedigend dabei ist die, sofern wir nicht konkret trigonalisieren, Unwägbarkeit über die Einträge oberhalb der Hauptdiagonale. Für einen Endomorphismus α, dessen charakteristisches Polynom p_α in Linearfaktoren zerfällt, aber keine hinreichende Anzahl linear unabhängiger Vektoren liefert, kann jedoch durch „Basteln“ einer geschickten Basis eine bessere Matrixdarstellung als die Trigonalform aus Abschnitt 3.2 gefunden werden. Aus dieser zugeschnittenen Basis resultiert die bestmögliche Matrixdarstellung für obigen Fall. Das Ergebnis wird als *Jordansche Normalform* bezeichnet.
Der in diesem Kapitel dargelegte Weg ermöglicht es uns die Jordansche Normalform elementarer als in den kommenden Ansätzen herzuleiten; so wird zum Beispiel das sogenannte Minimalpolynom nur grob definiert und für die eigentlichen Beweise der Hauptsätze nicht verwendet (bzw. hätte auf die Verwendung des Minimalpolynoms ganz verzichtet werden können). Auch äußerst spärlich wird die Teilbarkeitstheorie im Polynomring in Anspruch genommen, die hier angewendeten Erkenntnisse sind vielleicht schon aus der Schule bekannt.
Der Titel soll übrigens keinesfalls abwertend erscheinen, sondern lediglich auf den erfreulichen Umstand hinweisen, dass wir dieser Präsentation ohne weitere Kenntnis von tieferer Mathematik folgen können.

In Ansatz Nummer Eins wird zunächst gezeigt, dass die Eigenräume, um die entsprechende Dimension erweitert, dann Haupträume genannt, eine invariante direkte Summenzerlegung von V bilden, wobei V den Vektorraum über $\mathbb{K}$ für die lineare Abbildung $\alpha\colon V \to V$ bezeichne. Das Lemma von Fitting und dessen Erkenntnisse über nilpotente Abbildungen (die lineare Abbildung α besitzt nämlich jeweils auf einen Hauptraum eingeschränkt einen nilpotenten Anteil φ) wird uns dann einiges an Beweisarbeit vorweg nehmen und den eigentlichen Beweis der Hauptraumzerlegung recht kurz darstellbar machen.
Ein zweiter Hauptsatz zeigt dann - zumindest implizit - dass die Haupträume bis zu einem klar definierten Grad weiters invariant zerlegbar sind.

In Matrizen ausgedrückt bedeuten diese zwei Hauptsätze folgendes: der erste Hauptsatz liefert invariante Räume (die Haupträume) mittels deren (nach Wahl einer Basis für diese und der anschließenden Transformation) eine Matrixdarstellung (des Endomorphismus α) in Blockform gegeben ist. Mit dem Beweis des

zweiten Hauptsatzes wird die Konstruktion einer speziellen Basis gezeigt. Die Struktur dieser Basis demonstriert dann, warum es keine bessere α-invariante Zerlegung des jeweiligen Haupraumes geben kann und ermöglicht somit die handlichste Gestaltung eines Blockes.
Die Basen sämtlicher Haupträume vereinigt, ergeben die sogenannte Jordan-Basis, aus der dann wiederum die *Jordansche Normalform* der linearen Abbildung α resultiert.

Das gesamte Kapitel ist - mit Ausnahme der Schlussbemerkung, die auf [Jän08] gründet - in enger Anlehnung an [Fis10] entstanden.

4.1. Hauptraumzerlegung

Nach 3.4.2 wissen wir, dass ein Vektorraum V über $\mathbb{K}$ genau dann als direkte Summe der α-invarianten Eigenräume $\mathrm{Eig}(\alpha;\lambda_i)$ dargestellt werden kann, wenn

$$\dim(\mathrm{Eig}(\alpha;\lambda_i)) = \mu(p_\alpha;\lambda_i) = r_i \text{ ist.}$$

Liefern die Eigenwerte λ_i allerdings nicht genügend linear unabhängige Eigenvektoren ist es möglich, den Eigenraum zu einem neuen Teilraum mit hinreichend großer Dimension zu erweitern.

Für einen Eigenwert λ der Vielfachheit $r \geq 1$ setzt man

$$\ker(\alpha - \lambda \cdot \mathrm{id})^r =: \mathrm{Hau}(\alpha;\lambda)$$

und bezeichnet diesen Kern als den *Hauptraum* (oder *verallgemeinerten Eigenraum*) von α zum Eigenwert λ.

Offensichtlich ist

$$\mathrm{Eig}(\alpha;\lambda) = \ker(\alpha - \lambda \cdot \mathrm{id}) \subset \ker(\alpha - \lambda \cdot \mathrm{id})^r = \mathrm{Hau}(\alpha;\lambda).$$

Es gilt auch allgemein für beliebige $\varphi \in \mathrm{End}(V)$:

$$\{0\} \subset \ker(\varphi) \subset \ker(\varphi^2) \subset \cdots \subset \ker(\varphi^l),$$

$$V \supset \mathrm{Im}(\varphi) \supset \mathrm{Im}(\varphi^2) \supset \cdots \supset \mathrm{Im}(\varphi^l).$$

Vorsicht: Es ist zwar laut Dimensionssatz 2.2.1 $\dim(\ker(\varphi^j) + \dim(\mathrm{Im}(\varphi^j)) = \dim(V)$ für alle $j = 1, \ldots, l$, aber im Allgemeinen ist $\ker(\varphi^j) \cap \mathrm{Im}(\varphi^j) \neq \{0\}$ und es ist somit durch diese Unterräume keine direkte Summenzerlegung von V gegeben.

Bevor wir mittels der Hauptraumzerlegung dem gerade erwähnten Problem der Summenzerlegung entgegen treten wollen, brauchen wir noch das

Lemma 4.1.1 (Lemma von Fitting)**.** Für $\varphi \in \mathrm{End}(V)$ wird

$$d := \min\{l \in \mathbb{N} : \ker(\varphi^l) = \ker(\varphi^{l+1})\} \text{ und } r := \mu(p_\varphi; 0),$$

wobei $\varphi^0 := \mathrm{id}$, betrachtet. Dann gilt:

(i) $d = \min\{l : \mathrm{Im}(\varphi^l) = \mathrm{Im}(\varphi^{l+1})\}$

(ii) $\ker(\varphi^d) = \ker(\varphi^{d+i})$, $\mathrm{Im}(\varphi^d) = \mathrm{Im}(\varphi^{d+i})$ für alle $i \in \mathbb{N}$

(iii) Die Räume $U := \ker(\varphi^d)$ und $W := \mathrm{Im}(\varphi^d)$ sind φ-invariant.

(iv) $(\varphi|_U)^d = 0$ und $\varphi|_W : W \to W$ ist ein Isomorphismus.

(v) Für das Minimalpolynom von $\varphi|_U$ gilt $m_{\varphi|_U} = t^d$.

(vi) $V = U \oplus W$, $\dim(U) = r \geq d$, $\dim(W) = n - r$.

Insbesondere gibt es eine Basis $\mathcal{B}$ von V, sodass

$$M_{\mathcal{B}}(\varphi) = \begin{pmatrix} N & 0 \\ 0 & C \end{pmatrix} \quad mit\ N^d = 0.$$

Beweis. Man erhält die folgenden Äquivalenzen:

$$\begin{aligned} \ker(\varphi^l) = \ker(\varphi^{l+1}) &\overset{\ker(\varphi^l) \subset \ker(\varphi^{l+1})}{\Leftrightarrow} \dim(\ker(\varphi^l)) = \dim(\ker(\varphi^{l+1}) \\ &\Leftrightarrow \quad \dim(\mathrm{Im}(\varphi^l)) = \dim(\mathrm{Im}(\varphi^{l+1}) \\ &\underset{\mathrm{Im}(\varphi^{l+1}) \subset \mathrm{Im}(\varphi^l)}{\Leftrightarrow} \mathrm{Im}(\varphi^l) = \mathrm{Im}(\varphi^{l+1}). \end{aligned}$$

Damit wäre (i) geklärt. Die Gleichheit $\ker(\varphi^l) = \ker(\varphi^{l+1})$ ist somit auch äquivalent dazu, dass $\varphi|_{\mathrm{Im}(\varphi^l)} : \mathrm{Im}(\varphi^l) \to \mathrm{Im}(\varphi^{l+1})$ ein Isomorphismus ist, d.h. $\varphi|_W : W \to W$ ist ein Isomorphismus (der zweite Teil von (iv) folgt).
Damit ist auch (ii) klar.
Zu (iii) gibt es auch nicht allzu viel zu sagen: Nachdem (ii) gilt (also sich die Räume $\mathrm{Im}(\varphi^i)$ und $\ker(\varphi^i)$ ab $i \geq d$ nicht mehr ändern), ist - die Leserin oder der Leser verzeihe mir den Ausdruck - trivialerweise die φ-Invarianz, also (iii), gegeben.

Die Gleichheit $(\varphi|_U)^d = 0$ ist klar, da $v \in \ker(\varphi^d)$ per Definition $\varphi^d(v) = 0$ erfüllt. (v) wird aus 4.2.1 folgen, da $\varphi|_U$ nilpotent mit Nilpotenzindex d ist, d.h. $(\varphi|_U)^d = 0$, aber $(\varphi|_U)^{d-1} \neq 0$, da sonst $\ker(\varphi^d) = U \subset \ker(\varphi^{d-1})$ gelten würde, im Widerspruch zu $\ker(\varphi^{i-1}) \subsetneq \ker(\varphi^i)$ für $i \leq d$.
(iv): Da nach 2.2.1 (Dimensionssatz für lineare Abbildungen) $\dim(V) = \dim(U) + \dim(W)$ gilt, zeigen wir zunächst, dass $U \cap W$ trivial ist, und somit nach 2.2.2 (Dimensionssatz für direkte Summen) $V = U \oplus W$ gilt. Sei also $v \in U \cap W$, dann ist $\varphi^d(v) = 0$ und es existiert ein $w \in V$ mit $\varphi^d(w) = v$. Folglich gilt dann, dass

$$w \in \ker(\varphi^{2d}) = \ker(\varphi^d)$$

ist, wegen 2) und $\varphi^{2d}(w) = 0$. Das bedeutet dann gerade, dass $v = \varphi^d(w) = 0$ ist. Aus Satz 3.3.1 geht hervor, dass eine invariante Summenzerlegung äquivalent zu einer Blockdiagonalform ist, N entspricht hierbei genau $\varphi|_U$, also ist $N^d = 0$.
Bleibt noch $\dim(U) = r \geq d$ zu zeigen. Da $\ker(\varphi^{i-1}) \subsetneq \ker(\varphi^i)$ für $i \leq d$ gilt, vergrößert sich auch die jeweilige Dimension dieser Teilräume, d.h. $\dim(U) \geq d$. Wegen $r = \mu(p_\varphi; 0)$ und $V = U \oplus W$ gilt

$$t^r \cdot q = p_\varphi = p_{\varphi|_U} \cdot p_{\varphi|_W} \quad \text{mit } q(0) \neq 0.$$

Aus der Annahme $p_{\varphi|_W}(0) := (a_k t^k + \cdots + a_0)(0) = 0$ ließe sich, im Widerspruch zur Bijektivität der Einschränkung $\varphi|_W$ stehend, $\det(\varphi|_W) = a_0 = 0$ folgern.
Damit ist $p_{\varphi|_W}(0) \neq 0$, also $p_{\varphi|_U} = t^r$ und $\dim(U) = r$. □

Dies war nun etwas mühselig. Umso eleganter können wir mit den Ergebnissen aus diesem Lemma ausgestattet den ersten Hauptsatz dieses Kapitels angehen.

Satz 4.1.2. Hauptraumzerlegung. Sei $\alpha : V \to V$ eine lineare Abbildung, wobei V ein $\mathbb{K}$-Vektorraum ist, und

$$p_\alpha = \pm(t - \lambda_1)^{r_1} \cdot \ldots \cdot (t - \lambda_k)^{r_k}$$

mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$. Es sei $V_i := \mathrm{Hau}(\alpha; \lambda_i) \subset V$ für jedes λ_i der Hauptraum. Dann gilt:

(i) $\alpha(V_i) \subset V_i$ und $\dim(V_i) = r_i$ für $i = 1, \ldots, k$.

(ii) $V = V_1 \oplus \cdots \oplus V_k$.

(iii) α hat eine Zerlegung $\alpha = \alpha_D + \alpha_N$ (α_D diagonalisierbar, α_N nilpotent).

Dieser Satz offenbart in Matrizen gedacht sogleich seine Bedeutung für die Entwicklung einer *Normalform*:

Korollar 4.1.3. Sei $A \in M_n(\mathbb{K})$, sodass $p_\alpha = \pm(t-\lambda_1)^{r_1} \cdot \ldots \cdot (t-\lambda_k)^{r_k}$. Dann gibt es eine invertierbare Matrix $S \in GL_n(\mathbb{K})$ derart, dass

$$SAS^{-1} = \begin{pmatrix} \boxed{\lambda_1 \mathbb{I}_{r_1} + N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_k \mathbb{I}_{r_k} + N_k} \end{pmatrix} =: \tilde{A},$$

wobei für $i = 1, \ldots, k$

$$\lambda_i \mathbb{I}_{r_i} + N_i = \begin{pmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix} \in M_{r_i}(\mathbb{K}) \text{ ist.}$$

Insbesondere gilt $\tilde{A} = D + N$, wobei D eine Diagonalmatrix und N nilpotent ist, und weiters

$$D \cdot N = N \cdot D.$$

□

Beweis des Satzes zur Hauptraumzerlegung 4.1.2. Mittels Induktion über die Zahl $k \geq 1$ der verschiedenen Eigenwerte von α: Wir setzen $\varphi := \alpha - \lambda_1 \cdot \mathrm{id}$.
Es gilt

$$p_\varphi(t - \lambda_1) = p_\alpha(t), \text{ also } \quad \mu(p_\varphi; 0) = \mu(p_\alpha; \lambda_1) = r_1.$$

Zur Erinnerung: $\ker(\alpha - \lambda \cdot \mathrm{id})^m =: \mathrm{Hau}(\alpha; \lambda)$ für ein geeignetes $m \in \mathbb{N}$. Demnach folgt mit Lemma 4.1.1, dass $V = \mathrm{Hau}(\alpha; \lambda_1) \oplus W$ und $r_1 = \dim(\mathrm{Hau}(\alpha; \lambda_1))$ ist. Weiters gilt nach diesem Lemma, dass die Summanden $\mathrm{Hau}(\alpha; \lambda_1)$ und W φ-invariant sind. Da $\alpha = \varphi + \lambda_1 \cdot \mathrm{id}$ ist, gilt auch $\alpha(\mathrm{Hau}(\alpha; \lambda_1)) \subset \mathrm{Hau}(\alpha; \lambda_1)$ und $\alpha(W) \subset W$, also die α-Invarianz dieser Summanden.
Diese α-invarianten Teilräume „liefern“ eine Matrixdarstellung von α in Blockdiagonalform (siehe 3.3), dann gilt

$$p_{\alpha|_W} = \pm(t-\lambda_2)^{r_2} \cdot \ldots \cdot (t-\lambda_k)^{r_k},$$

und somit folgt nach Induktionsvoraussetzung (i) und (ii).
Die gewünschte Existenz aus (iii) erhält man mit (bezüglich der Bezeichnungen von Korollar 4.1.3)

$$D := \begin{pmatrix} \boxed{\lambda_1 \mathbb{I}_{r_1}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_k \mathbb{I}_{r_k}} \end{pmatrix} \quad \text{und} \quad N := \begin{pmatrix} \boxed{N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{N_k} \end{pmatrix}.$$

Leicht überprüfbar ist dann auch

$$D \cdot N = \begin{pmatrix} \boxed{\lambda_1 N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_k N_k} \end{pmatrix} = N \cdot D.$$

□

Es wurde hiermit die Zerlegbarkeit obiger Abbildung α gezeigt, dabei gilt demnach die Überführbarkeit des diagonalisierbaren Anteils α_D in eine ideale Matrixdarstellung - hier also Diagonalform - als gesichert.
Die Sicherung einer ebenfalls „ansehnlichen“ Matrixdarstellung für den nilpotenten Anteil α_N wird Gegenstand des nachfolgenden Abschnitts sein.

4.2. Nilpotente Endomorphismen

Ist für ein lineare Abbildung $\varphi : V \to V$ und einer geeigneten Potenz $d \in \mathbb{N}$ $\varphi^d(v) = 0$ für jedes $v \in V$, so heißt φ **nilpotent** mit *Nilpotenzindex* d, sofern $d = \min\{l : \varphi^l = 0\}$ gilt. Derartige Abbildungen sind vor allem auch betrachtenswert geworden seit wir aus der *Hauptraumzerlegung* 4.1.2 um die Existenz des nilpotenten Anteils eines beliebigen linearen Endomorphismus α (dessen charakteristisches Polynom p_α in Linearfaktoren zerfällt) wissen.

Satz 4.2.1. Ist $\varphi \in \mathrm{End}(V)$ und $n = \dim(V)$, so sind folgende Aussagen äquivalent:

(i) φ ist nilpotent.

(ii) $\varphi^d = 0$ für ein d mit $1 \leq d \leq n$.

(iii) Das charakteristische Polynom p_φ ist gegeben durch $p_\varphi = \pm t^n$.

(iv) Es existiert eine Basis $\mathcal{B}$ von V, sodass

$$M_{\mathcal{B}}(\varphi) = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix} \text{ gilt.}$$

Beweis. (i) $\Rightarrow$ (ii) auf Grund der Folgerung aus 2.5.
(ii) $\Rightarrow$ (iii) : Das Minimalpolynom m_φ muss von der Form $m_\varphi = t^d$ mit $1 \leq d \leq n$

sein, weil aus (ii) folgt, dass $t^k \in I_\varphi$ für ein $1 \leq k \leq n$ ist und aus der Bemerkung über den Zusammenhang des Minimalpolynoms m_α und dem Ideal I_α aus 2.5 geht $t^k = r \cdot m_\varphi$ hervor. Da m_φ und p_φ dieselben Primfaktoren besitzen[1], folgt mit $\deg(p_\varphi) = n$, dass $p_\varphi = \pm t^n$ ist.
(iii) $\Rightarrow$ (iv): Nach (iii) zerfällt das charakteristische Polynom p_φ in Linearfaktoren, somit liefert der Trigonalisierungssatz 3.2.1 die Existenz einer Basis $\mathcal{B}$, sodass $M_\mathcal{B}(\varphi)$ obere Dreiecksform hat. Die Diagonale von $M_\mathcal{B}(\varphi)$ kann keine Einträge ungleich 0 besitzen wegen der Form des charakteristischen Polynoms $p_\varphi = \pm t^n$ (bezüglich eines Basiswechsel ist p_φ ja gerade invariant). So entsteht die gewünschte Form.
(iv) $\Rightarrow$ (i): Dies kann auch anhand der Struktur von $M_\mathcal{B}(\varphi)$ (mit einer Basis $\mathcal{B}$ wie in (iv) gegeben) überlegt werden:

$$(M_\mathcal{B}(\varphi))^k = \underbrace{M_\mathcal{B}(\varphi) \cdot \ldots \cdot M_\mathcal{B}(\varphi)}_{k \; mal}$$

ist spätestens für $k = n$ gleich Null, da für jede Multiplikation mindestens eine weitere Nebendiagonale verschwindet. □

Der nun folgende zweite Hauptsatz demonstriert die Konstruktion einer Basis mittels derer ein nilpotenter Endomorphismus in eine ideale Matrixform gebracht werden kann.
Der Einfachheit halber definieren wir dazu die *Jordanmatrix*

$$J_k := \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \in M_k(\mathbb{K}),$$

„für die $J_k^k = 0$ gilt, und k die minimale Potenz mit dieser Eigenschaft ist.“[2]

Satz 4.2.2. Sei φ ein nilpotenter Endomorphismus eines $\mathbb{K}$-Vektorraumes V und $d := \min\{l : \varphi^l = 0\}$. Dann gibt es eindeutig bestimmte Zahlen $s_1, \ldots, s_d \in \mathbb{N}$ mit

$$d \cdot s_d + (d-1) \cdot s_{d-1} + \cdots + s_1 = \dim(V)$$

[1] Das haben wir nicht bewiesen, einen Beleg für den Fall $\mathbb{K} = \mathbb{C}$ liefert [Fis10] auf Seite 257.
[2] [Fis10, Seite 264].

und eine Basis $\mathcal{B}$ von V, sodass

$$M_{\mathcal{B}}(\varphi) = \begin{pmatrix} J_d & & & & & & & & \\ & \ddots & & & & & & & \\ & & J_d & & & & & & \\ & & & J_{d-1} & & & 0 & & \\ & & & & \ddots & & & & \\ & & & & & J_{d-1} & & & \\ & & 0 & & & & \ddots & & \\ & & & & & & & J_1 & \\ & & & & & & & & \ddots \\ & & & & & & & & & J_1 \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} s_d\text{-mal} \\ \left.\vphantom{\begin{matrix}a\\a\end{matrix}}\right\} s_{d-1}\text{-mal} \\ \vdots \\ \left.\vphantom{\begin{matrix}a\\a\end{matrix}}\right\} s_1\text{-mal} \end{matrix}$$

ist.

Beweis. Sei $U_l := \ker(\varphi^l)$, dann gilt

$$\{0\} = U_0 \subset U_1 \subset \cdots \subset U_{d-1} \subset U_d = V.$$

Wir beginnen bei $U_d = V$ und unterteilen diesen Raum in alle Vektoren, die genau d-mal abgebildet Null ergeben (W_d) und in diejenigen für die das nicht gilt (U_{d-1}); demgemäß erreicht man $U_{d-1} \oplus W_d = V$.
Der Unterraum U_{d-1} wird wiederum in Vektoren unterteilt, die nach exakt $(d-1)$-mal Abbilden durch φ verschwinden, d.h. $\varphi^{d-1}(v) = 0$, aber $\varphi^{d-2}(v) \neq 0$ und in die anderen, es ensteht $U_{d-1} = U_{d-2} \oplus W_{d-1}$.

Wir bauen nach diesem Muster iterativ die Räume U_{d-i} für $1 \leq i \leq d$ ab und erhalten folgende Zerlegung:

$$\begin{array}{ccccccccc} U_d & & & & & & & & \\ \downarrow & & & & & & & & \\ U_{d-1} & \oplus & W_d & & & & & & \\ \downarrow & & \downarrow & & & & & & \\ U_{d-2} & \oplus & W_{d-1} & \oplus & W_d & & & & \\ \downarrow & & \downarrow & & \downarrow & & & & \\ \vdots & & \vdots & & \vdots & & & & \\ U_1 & \oplus & W_2 & \oplus & W_3 & \oplus \cdots \oplus & W_d & & \\ \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\ U_0 & \oplus & W_1 & \oplus & W_2 & \oplus \cdots \oplus & W_{d-1} & \oplus & W_d\,. \end{array}$$

Jede Zeile ist dabei eine Zerlegung von V, dabei ist $U_0 = \{0\}$, also

$$V = W_1 \oplus \cdots \oplus W_d.$$

Nun kann man schrittweise durch ergänzen von Basen aller W_l zu einer Basis von V gelangen:

$$\begin{array}{llll} w_1^{(d)}, \ldots\ldots, w_{s_d}^{(d)} & & & \\ \varphi(w_1^{(d)}), \ldots\ldots, \varphi(w_{s_d}^{(d)}), & & w_1^{(d-1)}, \ldots\ldots\ldots, w_{s_{d-1}}^{(d-1)} & \\ \vdots & \vdots & \vdots & \vdots \\ \varphi^{d-1}(w_1^{(d)}), \ldots, \varphi^{d-1}(w_{s_d}^{(d)}), & \varphi^{d-2}(w_1^{(d-1)}), \ldots, & \varphi^{d-2}(w_{s_{d-1}}^{(d-1)}), \ldots, & w_1^{(1)}, \ldots, w_{s_1}^{(1)} \end{array}.$$

Dabei ist die erste Zeile eine Basis von W_d, die zweite von W_{d-1}, und schlussendlich bildet die letzte Zeile eine Basis von

$$W_1 = U_1 = \ker(\varphi).$$

Nun ordnen wir die Basis folgenderweise: In jeder Spalte läuft man von unten nach oben und liest die Spalten von links nach rechts. Die größte Beweisarbeit ist hiermit erbracht, wir „ergattern“ nämlich durch diese Anordnung die gewünschte Form von $M_{\mathcal{B}}(\varphi)$.

Die Zahlen $s_1, \ldots, s_d$ sind genau die Dimensionen bzw. die Erweiterung auf eine Dimension der jeweiligen Unterräume W_j für $j = 1, \ldots, d$, also rein durch φ bestimmt.

Was noch nicht ausreichend behandelt wurde ist die Frage, ob φ angewandt auf die Vektoren $w_1^{(l)}, \ldots\ldots, w_{s_l}^{(l)}$ wirklich eine Menge linear unabhängiger Vektoren für W_{l-1} liefert. Dies ergibt sich aber direkt aus der Tatsache, dass φ injektiv ist. Da $\varphi(W_l) \subset W_{l-1}$, also $\dim(W_l) \leq \dim(W_{l-1})$ ist, gilt außerdem, dass $s_{l-1} \not< s_l$ für $l = 1, \ldots, d$. □

Die eben konstruierte Basis stellt, wie wir gleich sehen werden, den Schlüssel zur *Jordanschen Normalform* dar.

4.3. Jordansche Normalform

Die bisher erbrachten Anstrengungen münden nun in diesen Abschnitt und wir können nach dem Formulieren des Satzes über die *Jordansche Normalform* den Beweis desselbigen äußerst knapp gestalten.

Satz 4.3.1. Sei V ein $\mathbb{K}$-Vektorraum und $\alpha : V \to V$ eine lineare Abbildung derart, dass das charakteristische Polynom in Linearfaktoren zerfällt, also

$$p_\alpha = \pm(t-\lambda_1)^{r_1} \cdot \dots \cdot (t-\lambda_k)^{r_k}$$

mit paarweise verschiedenen $\lambda_1, \dots, \lambda_k \in \mathbb{K}$. Dann gibt es eine Basis $\mathcal{B}$ von V, sodass

$$M_{\mathcal{B}}(\alpha) = \begin{pmatrix} \boxed{\lambda_1 \mathbb{I}_{r_1} + N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_k \mathbb{I}_{r_k} + N_k} \end{pmatrix} \text{ gilt,}$$

wobei N_i für $i = 1, \dots, k$ die Gestalt einer Jordanmatrix aus Abschnitt 4.2 hat.

Man beachte, dass diese Darstellung täuschen kann: die Zahlen $r_i \in \mathbb{N}$ $(1 \leq i \leq k)$ sind im Allgemeinen verschieden und somit besitzen die Teilmatrizen $\lambda_i \mathbb{I}_{r_i} + N_i$ auch unterschiedliche Größen.

Eine Teilmatrix $\lambda_i \mathbb{I}_{r_i} + N_i$ ist dann durch

$$\lambda_i \mathbb{I}_{r_i} + N_i = \begin{pmatrix} \lambda_i & 1 & & & & & & & & & \\ & \ddots & \ddots & & & & & & & & \\ & & \ddots & 1 & & & & & & & \\ & & & \lambda_i & 0 & & & & & & \\ & & & & \lambda_i & 1 & & & & & \\ & & & & & \ddots & \ddots & & & & \\ & & & & & & \ddots & 1 & & & \\ & & & & & & & \lambda_i & 0 & & \\ & & & & & & & & \ddots & \ddots & \\ & & & & & & & & & \ddots & 0 \\ & & & & & & & & & & \lambda_i \end{pmatrix} \text{ gegeben.}$$

Die nicht beschriebenen Einträge dieser *Jordanmatrix* zum Eigenwert λ_i sind Nullen.

Beweis. Die Beweisarbeit für diesen langen Satz ist schon geleistet. Für $i = 1, \ldots, k$ setzen wir

$$V_i := \mathrm{Hau}(\alpha; \lambda_i) \qquad \text{und} \qquad \varphi_i := (\alpha - \lambda_i \cdot \mathrm{id})|_{V_i}.$$

Mit Satz 4.1.2 wissen wir um die Zerlegbarkeit von α.
Die entsprechende Anordnung der Basiselemente aus Satz 4.2.2 bezüglich des nilpotenten Anteils φ_i liefert dann eine Basis $\mathcal{B}_i$ von V_i und $\mathcal{B}_1, \ldots, \mathcal{B}_k$ dann insgesamt eine Basis $\mathcal{B}$, sodass $M_{\mathcal{B}}(\alpha)$ die gewünschte Form - nämlich die *Jordansche Normalform* - besitzt. □

In $\lambda_i \mathbb{I}_{r_i} + N_i$ wird ein Kästchen der Gestalt

$$\begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$$

mit d Einträgen λ_i und $d-1$ Einsen als *Jordanblock* der Länge d zu λ_1 bezeichnet. Wir können uns anhand der in Satz 4.2.2 illustrierten Anordnung von Basiselementen überlegen, dass ein Jordanblock genau einer dieser Spalten (von unten nach oben gelesen) entspricht. Die ersten $s_{d_i}^{(i)}$ Jordanblöcke in $\lambda_i \mathbb{I}_{r_i} + N_i$ sind somit die größten mit Länge $d_i = \min\{l : N_i^l = 0\}$, d.h. d_i ist der Nilpotenzindex von φ_i.[3] Die letzten $s_1^{(i)}$ Jordanblöcke in $\lambda_i \mathbb{I}_{r_i} + N_i$ bestehen dann nur noch jeweils aus einer Null (Länge ist Eins). Mit einer Anpassung der Bezeichnung aus Satz 4.2.2 an φ_i, nämlich $\dim(W_d) := s_{d_i}^{(i)}$ usw., ergibt sich

$$d_i \cdot s_{d_i}^{(i)} + (d_i - 1) \cdot s_{d_i - 1}^{(i)} + \cdots + s_1^{(i)} = r_i,$$

außerdem gilt klarerweise

$$r_1 + \cdots + r_k = n.$$

Wir haben somit den Zusammenhang zwischen den Elementen $r_i, d_i, s_j^{(i)} \in \mathbb{N}$ $(1 \leq j \leq d_i)$ generiert. Diese Elemente sowie der Eigenwert $\lambda_i \in \mathbb{K}$ sind unabhängig von der Wahl einer Basis für V und somit allein durch den Endomorphismus $\alpha : V \to V$ bestimmt. Man nennt diese daher **Invarianten** von α.
Das Studium linearer Abbildungen $\alpha : V \to V$ kann hiermit als abgerundet

[3] Man könnte hier noch zeigen, dass d_i der Vielfachheit der Nullstelle λ_i im Minimalpolynom m_α entspricht.

betrachtet werden: Diese Invarianten sind im Grundlegenden alles was uns ein Endomorphismus an Informationen liefert. „Umgekehrt bedeutet dies, daß es so viele Möglichkeiten gibt, wesentlich verschiedene Endomorphismen zu konstruieren, wie man 'Sätze' von solchen Invarianten zusammenstellen kann.“[4]

Vermöge der in diesem Kapitel formulierten Erkenntnisse können hier noch die Erläuterungen aus der Einleitung bezüglich der **Klassifikation von Matrizen** präzisiert werden.

Definition. Zwei Matrizen $A, B \in M_n(\mathbb{K})$ heißen *ähnlich* wenn es eine invertierbare Matrix $P \in M_n(\mathbb{K})$ gibt, sodass $B = PAP^{-1}$ gilt.

Diese *Ähnlichkeit* stellt eine Äquivalenzrelation dar und somit ein Werkzeug für die Klassifikation von Matrizen (vgl. [Jän08, S. 233]). Anhand der gerade angeeigneten Konzepte ist es naheliegend Matrizen derselben *Jordanschen Normalform* einer gemeinsamen Klasse zuzuordnen. Dies ist klar, denn wir interpretieren mit obiger Definition Matrizen die *ähnlich* zu einer Jordanmatrix J sind als verschiedene Matrixdarstellungen derselben linearen Abbildung.
Die Jordansche Normalform eines Endomorphismus wiederum kann als durch die angesprochenen Invarianten bestimmt aufgefasst werden. Wir haben somit mit der *Jordanschen Normalform* und der Bestimmung der *Invarianten* eine Klassifizierungsmöglichkeit für Matrizen gewonnen.

[4] [Fis10, Seite 269].

5. Die Jordansche Normalform - medium

Wie bereits angekündigt liefert dieser Ansatz im Vergleich zu Kapitel 4 eine abstraktere Version der Herleitung der Jordanschen Normalform. Dieser Ansatz stellt hier das Bindeglied zwischen Ansatz Eins und Ansatz Drei dar, weil der Bezug zu beiden Ansätzen noch recht deutlich ist.
Der elementarere Zugang aus Kapitel 4 bietet hierbei durchaus eine Hilfestellung zum schnelleren Verständnis. Der nun folgende Ansatz zeigt allgemeiner den Zusammenhang zwischen der Zerlegbarkeit des charakteristischen Polynoms p_α bezüglich einer linearen Abbildung $\alpha : V \to V$ und einer direkten Summenzerlegung von V (Vektorraum über $\mathbb{K}$).

Es ist recht bündig erläutert was in diesem Ansatz passiert: nach Bedingung eines in Linearfaktoren zerfallenden Minimalpolynoms greift der *Erste Zerlegungssatz* und liefert die erste direkte Summenzerlegung von V. Durch die Konstruktion eines offensichtlich α-invarianten, zyklischen Teilraums können wir mit dem *Zerlegungssatz* die gerade erwähnten Summanden mittels des *halbeinfachen Anteils* von α weiter zerlegen. Diese Zerlegung schöpft dann alle invarianten Unterräume aus, sodass wie in Kapitel 4 mit der Wahl einzelner Basen und dem Zusammenlegen dieser eine Jordan-Basis entsteht und die *Jordansche Normalform* folgt.

Ausgenommen von Satz 5.1.1 ([Art98]) und Satz 5.1.2 ([Cap11]) liefert das theoretische Fundament dieses Kapitels ausschließlich [Brö03].

5.1. Grundbegriffe

Zunächst werden wir zeigen, dass jedes Ideal (siehe 6.2) des Polynomringes $\mathbb{K}[t]$ (über einem Körper $\mathbb{K}$) von einem einzigen Element $p \in I \subset \mathbb{K}[t]$ erzeugt wird, d.h. jedes Element $f \in I$ ist ein „Vielfaches“ von p, besitzt also die Form $f = p \cdot q$ mit $q \in \mathbb{K}[t]$.
Ein derartiges Ideal heißt Hauptideal (siehe 6.4), diese Bezeichnung wird uns dann vor allem ab Kapitel 6 öfters begegnen.

Vorweg noch: ein Polynom wird **monisch** genannt, wenn sein führender Koeffizient (also der Koeffizient der höchsten Potenz) gleich Eins ist.

Satz 5.1.1 (Hauptidealsatz)**.** Zu jedem Ideal $I \subset \mathbb{K}[t]$ gibt es ein erzeugendes Polynom $p \in I \subset \mathbb{K}[t]$ (monisch und eindeutig), d.h. $I = p \cdot \mathbb{K}[t] := \{p \cdot q : q \in \mathbb{K}[t]\}$.

Beweis. Sei $I \subset \mathbb{K}[t]$ ein Ideal. Ist I das Nullideal sind wir fertig, da dieses ein Hauptideal ist. Sei also $I \neq \{0\}$. „Im Fall von $\mathbb{Z}$ findet man ein erzeugendes Element einer Untergruppe, die ungleich Null ist, indem man das kleinste positive Element dieser Untergruppe wählt. Der Ersatz in $\mathbb{K}[t]$ besteht darin, in I ein Polynom $p \neq 0$ kleinsten Grades zu wählen.“[1] Wir zeigen nun in eben gehörter Anlehnung an $\mathbb{Z}$, dass für dieses minimale Polynom $p \cdot \mathbb{K}[t] = I$ gilt. Nach Definition eines Ideals (siehe 6.2) gilt $p \cdot q \subset I$ für jedes $q \in \mathbb{K}$, also $p \cdot \mathbb{K}[t] \subset I$. Umgekehrt ergibt die Division mit Rest (2.4.1) für ein beliebiges Element $q \in I \subset \mathbb{K}[t]$ und p, dass $q = p \cdot g + r$ ist, wobei $r < p$ ist. Da $r = q - p \cdot g \in I$ ist, wäre $r \neq 0$ ein Widerspruch zur Minimalität von p. Somit haben wir $I \subset p \cdot \mathbb{K}[t]$ gezeigt.

Zur Eindeutigkeit: Seien p_1, p_2 monische Polynome mit $p_1 \cdot \mathbb{K}[t] = p_2 \cdot \mathbb{K}[t]$. Dann existieren Polynome $q_1, q_2 \in \mathbb{K}[t]$, sodass $p_1 = p_2 \cdot q_2$ und $p_2 = p_1 \cdot q_1$ gilt. Das bedeutet p_1 und p_2 haben gleichen Grad und $q_1, q_2 \in \mathbb{K}$ sind Skalare (also konstante Polynome). Da die Polynome p_1, p_2 monisch sind, folgt $q_1 = q_2 = 1$ und somit $p_1 = p_2$. □

Sei V ein $\mathbb{K}$-Vektorraum und $\alpha : V \to V$ linear. Wir haben in Abschnitt 2.5 bereits die Menge $I_\alpha := \{p \in \mathbb{K}[t] : p(\alpha) = 0\} \subset \mathbb{K}[t]$, also alle Polynome ($\in \mathbb{K}[t]$) die zur Nullabbildung ($V \to V$) werden, wenn wir die Variable $t \in \mathbb{K}$ durch α ersetzen, kennen gelernt. Die Menge I_α ist der Kern der Abbildung $\psi : \mathbb{K}[t] \to \mathrm{End}(V)$ mit $p \mapsto p(\alpha)$.

Zur Vorbereitung des ersten Zerlegungssatzes und der Spezifizierung des Minimalpolynoms benötigen wir folgenden

Satz 5.1.2. Sei V ein $\mathbb{K}$–Vektorraum und $\alpha : V \to V$ eine lineare Abbildung. Dann gilt:

(i) Für $p, q \in \mathbb{K}[t]$ und $\lambda \in \mathbb{K}$ gilt $(p + \lambda \cdot q)(\alpha) = p(\alpha) + \lambda \cdot q(\alpha)$ und $(p \cdot q)(\alpha) = p(\alpha) \circ q(\alpha)$.

(ii) Die Menge I_α ist ein Ideal in $\mathbb{K}[t]$, d.h $I_\alpha \subset \mathbb{K}[t]$ ist ein Teilraum und für $p \in I_\alpha$ und $q \in \mathbb{K}[t]$ ist $p \cdot q \in I_\alpha$.

Beweis. Zu (i): Das Maximum der Grade von p und q wird mit N bezeichnet, dann geben wir die Polynome p, q in der Darstellung $p = \sum_{i=0}^{N} a_i t^i$ und $q = \sum_{i=0}^{N} b_i t^i$ an, und somit $p(\alpha) = \sum_{i=0}^{N} a_i \alpha^i$ und $q(\alpha) = \sum_{i=0}^{N} b_i \alpha^i$. Damit ergibt sich offensichtlich $p(\alpha) + \lambda \cdot q(\alpha) = \sum_{i=0}^{N} (a_i + \lambda \cdot b_i)\alpha^i = (p + \lambda \cdot q)(\alpha)$.

[1] [Art98, Seite 410].

Für den zweiten Teil erinnern wir uns, dass nach Definition ist $t^i \cdot t^j = t^{i+j}$ in $\mathbb{K}$ und $\alpha^i \circ \alpha^j = \alpha^{i+j}$ in $\mathrm{End}(V)$ gilt. Man kann sich nun vor Augen führen, dass es keinen Unterschied macht, ob die Polynome p, q zuerst multiplizieren (mit Verwendung des Distributivgesetzes) und dann α in dieses Produkt einsetzten oder die Komposition $p(\alpha) \circ q(\alpha)$ (unter Verwendung der Linearität) auswerten, das bedeutet genau $(p \cdot q)(\alpha) = p(\alpha) \circ q(\alpha)$.
Zu (ii): Für $p, q \in I_\alpha$ und $\lambda \in \mathbb{K}$ gilt nach Teil (i): $(p+\lambda \cdot q)(\alpha) = p(\alpha)+\cdot\lambda\cdot q(\alpha) = 0$. Damit ist $(p + \lambda \cdot q) \in I_\alpha$, und somit $I_\alpha \subset \mathbb{K}[t]$ ein Unterraum. Für $p \in I_\alpha$ und $q \in \mathbb{K}$ beliebig gilt wieder nach Teil (i): $(p \cdot q)(\alpha) = p(\alpha) \circ q(\alpha) = 0 \circ q(\alpha) = 0$. Damit ist $p \cdot q \in I_\alpha$. □

Definition. Das normierte Polynom kleinsten Grades $m \in I_\alpha \subset \mathbb{K}[t]$ wird als **Minimalpolynom** bezeichnet.

Da mit Satz 5.1.2 der Teilraum $I_\alpha \subset \mathbb{K}[t]$ ein Hauptideal ist (nach Satz 5.1.1 ist $\mathbb{K}[t]$ ein Hauptideal), gilt $I_\alpha = m \cdot \mathbb{K}[t]$, die Menge I_α ist also das Erzeugnis des Minimalpolynoms m.

Unmittelbar ergibt sich folgende

Bemerkung. Ist m das Minimalpolynom von α und $g \in I_\alpha$, also $g(\alpha) = 0$, so ist m ein Teiler von g.

Im Folgenden bezeichnet α stets einen Endomorphismus auf einem Vektorraum V über einem algebraisch abgeschlossenen Körper $\mathbb{K}$ mit $\dim(V) = n$.

Definition. Ein Unterraum $U \subset V$ heißt α-**Modul**[2], wenn gilt $\alpha U \subset U$, also einfach wenn U α-invariant (siehe 3.1) ist.

Definition. Ein α-Modul heißt **zyklisch**, wenn

$$U = L_\alpha(v) := \{\sum_{j=0}^{k} \lambda_j \alpha^j(v) \ : \ \lambda_j \in \mathbb{K},\, k \in \mathbb{N}_0\} \text{ für ein } v \in V.$$

Dieses sogenannte α-**Erzeugnis** von v ist der kleinste α-Modul, welcher v enthält.

Zu Beginn gleich ein Satz mit weit reichenden Folgen, der uns nämlich mit einem zerfallenden Minimalpolynom unmittelbar die Hauptraumzerlegung wie in Satz 4.1.2 liefern wird.

[2]Man könnte in diesem Kapitel *Modul* jedes Mal mit α-invarianter Unterraum ersetzen, die Bezeichnung *Modul* ist aber in Ausblick auf Kapitel 7 von großer Bedeutung und soll darum propädeutisch schon zum Einsatz kommen.

Satz 5.1.3. Erster Zerlegungssatz. Sei $f \in \mathbb{K}[t]$, $f = g \cdot h$, und g, h seien teilerfremd, d.h. ihr größter gemeinsamer Teiler sei 1. Sei α ein Endomorphismus von V.
Dann gilt:

(i) $\ker(f(\alpha)) = \ker(g(\alpha)) \oplus \ker(h(\alpha))$

(ii) Die beiden Projektionen $\ker(f(\alpha)) \to \ker(g(\alpha))$ und $\ker(f(\alpha)) \to \ker(h(\alpha))$ auf die Summanden sind durch Endomorphismen $(q \cdot h)(\alpha)$ und $(p \cdot g)(\alpha)$ mit $p, q \in \mathbb{K}[t]$ gegeben.

Beweis. Sei V ein endlichdimensionaler Vektorraum über einem Körper $\mathbb{K}$ und $\alpha \colon V \to V$ linear. Der Übersichtshalber ändern wir hier die Nomenklatur: statt $f(\alpha)$ schreiben wir f_α usw. bzw. den an einem Vektor ausgewerteten Endomorphismus $f(\alpha)(v)$ bezeichnen wir dann mit $f_\alpha v$.
Der Satz wird in drei Schritten bewiesen: wir zeigen zuerst, dass sich jeder Vektor $v \in \ker(f_\alpha)$ in $\mathrm{Im}(q_\alpha \cdot h_\alpha) + \mathrm{Im}(p_\alpha \cdot g_\alpha)$ befindet. Danach, dass $\mathrm{Im}(q_\alpha \cdot h_\alpha) \subset \ker(g_\alpha)$ und $\mathrm{Im}(p_\alpha \cdot g_\alpha) \subset \ker(h_\alpha)$. Und zuletzt, dass falls sich ein Element v in $\ker(g(\alpha)) \cap \ker(h(\alpha))$ befindet, es sich nur um den Nullvektor $v = 0$ handeln kann.
Dieser Beweis wirkt vielleicht - auch auf Grund der behäbigen Nomenklatur - etwas kompliziert; Teil (i) von Satz 5.1.2 und die darin erklärte Multiplikation von Endomorphismen der obigen Form sollte dabei für Abhilfe sorgen.

Sei also $v \in \ker(f_\alpha)$.
Der Euklidische Algorithmus 2.4.1 liefert Polynome $p, q \in \mathbb{K}$, sodass $p \cdot g + q \cdot h = 1$ und daraus folgt
$v = \mathrm{id}(v) = (p_\alpha \cdot g_\alpha + q_\alpha \cdot h_\alpha)v = (p_\alpha \cdot g_\alpha)v + (q_\alpha \cdot h_\alpha)v$, was den ersten Schritt beweist.
Dass die Bilder von $(q_\alpha \cdot h_\alpha)$ im Kern von g_α liegen folgt mit
$g_\alpha \cdot (q_\alpha \cdot h_\alpha v) = q_\alpha \cdot (h_\alpha \cdot g_\alpha v) = q_\alpha \cdot (f_\alpha v) = 0$. Wir haben dabei bedacht, dass die Komposition aus Teil (i) von Satz 5.1.2 kommutativ ist und die Klammern eigentlich nicht relevant sind (Assoziiertheit). Analog gilt $h_\alpha \cdot (p_\alpha \cdot g_\alpha v) = 0$.
Für den letzten Schritt sei $w \in V$ mit $g_\alpha w = h_\alpha w = 0$, dann ist $w = (p_\alpha \cdot g_\alpha)w + (q_\alpha \cdot h_\alpha)w = 0$ und der Satz somit bewiesen. □

Bemerkung. In Kapitel 7 wird der Zusammenhang zwischen teilerfremden Faktoren eines Polynoms und invarianten Teilräume noch verdeutlicht.

5.2. Jordansche Normalform

Wie eingangs erwähnt basiert die folgende Vorgehensweise auf [Brö03].

Wir werden den zur linearen Abbildung $\alpha\colon V \to V$ gehörenden $\mathbb{K}$-Vektorraum V in eine direkte Summe zyklischer α-Moduln zerlegen, die nicht weiter α-invariant zerlegbar sind.

Anhand des in Linearfaktoren zerfallenden Minimalpolynoms m von α, d.h.

$$m(t) = (t - \lambda_1)^{r_1} \cdot \ldots \cdot (t - \lambda_k)^{r_k},$$

können wir den Vektorraum V in α-invariante Summen zerlegen:

$$V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_k), \qquad V(\lambda_j) = \ker(t - \lambda_j)^{r_j}.$$

Dies folgt aus dem Ersten Zerlegungssatz 5.1.3 und entspricht dem Satz zur Hauptraumzerlegung 4.1.2.

Analog zum Abschnitt 4.2 sind wir nun an einer weiteren Zerlegung der jeweiligen Haupträume $V(\lambda_j)$ interessiert. Natürlich versuchen wir die Zerlegungen so weit zu treiben, bis das Maß an α-invarianten Unterräumen (oder Summanden der Zerlegung) erschöpft ist.
Eine für uns bereits aus Kapitel 4 bekannte Eigenschaft eines nilpotenten Endomorphismus $\varphi : V \to V$ ist, dass dieser eine im Allgemeinen nicht-triviale φ-invariante Zerlegung des Vektorraumes V induziert.

Dies liefert auch das folgende Lemma. Zuvor aber noch eine benötigte

Definition. Sei W ein $\mathbb{K}$-Vektorraum und $\varphi\colon W \to W$ linear. Die **Periode** $k \in \mathbb{N}$ eines Vektors $w \in W$ bezeichne diejenige Zahl für die $\varphi^k(w) = 0$, aber $\varphi^{k-1}(w) \neq 0$ gilt.

Falls φ nilpotent mit Nilpotenzindex $d \in \mathbb{N}$ ist, gilt im Falle $k = d$, dass w offensichtlich *maximale Periode* besitzt.

Lemma 5.2.1. Zerlegungslemma für nilpotente Endomorphismen.
Sei φ ein nilpotenter Endomorphismus eines m-dimensionalen Raumes W, und sei $w \in W$ von maximaler Periode k, dann existiert ein φ-invarianter Unterraum U von W, sodass

$$W = L_\varphi(w) \oplus U.$$

Beweis. Wir wählen einen φ-invarianten Unterraum $U \subset W$ von maximal möglicher Dimension, sodass $L_\varphi(w) \cap U = \{0\}$, d.h. die Summe $L_\varphi(w) + U$ ist direkt. Wir behaupten nun, dass jedes Element aus W bereits ein Element dieser direkten Summe sein muss, also $L_\varphi(w) \oplus U = W$.

Angenommen es gäbe ein $v \in W$, das nicht in $L_\varphi(w) \oplus U$ liegt. Dann existiert ein minimales $j \in \mathbb{N}$ mit $\varphi^j(v) \in L_\varphi(w) \oplus U$, denn spätestens $\varphi^m(v) = 0 \in U$ (resultiert aus der Folgerung in Abschnitt 2.5 und $m = \dim(W)$). Einfachheitshalber setzen wir $\varphi^{j-1}(v) = s$, dann gilt also

$$s \notin L_\varphi(w) \oplus U \quad \text{und} \quad \varphi(s) \in L_\varphi(w) \oplus U.$$

Jedes Element $x \in L_\varphi(w)$ lässt sich offensichtlich in der Form

$$x = \tilde{\mu} \cdot w + \varphi(\tilde{w}), \quad \tilde{w} \in L_\varphi(w),\ \tilde{\mu} \in \mathbb{K}$$

schreiben. Für $\varphi(s)$ bedeutet das, es existieren $\mu \in \mathbb{K}$, $w' \in L_\varphi(w)$, sodass

$$\varphi(s) = \underbrace{\mu \cdot w + \varphi(w')}_{\in L_\varphi(w)} + \underbrace{u}_{\in U}. \tag{5.1}$$

Das Element $\varphi(s)$ hat nach Voraussetzung eine Periode $\leq k-1$, da die Summen aus obiger Gleichung (5.1) φ-invariant sind, ist insbesondere auch die Periode von $\mu \cdot w + \varphi(w')$ echt kleiner k. Nach $k-1$ maligen Anwenden von φ auf $(\mu \cdot w + \varphi(w'))$ erhält man

$$0 = \varphi^{k-1}(\mu \cdot w + \varphi(w')) = \mu \cdot \varphi^{k-1}(w) + \underbrace{\varphi^k(w')}_{=0}$$

und leitet, da laut Voraussetzung $\varphi^{k-1}(w) \neq 0$ ist, daraus ab, dass $\mu = 0$ sein muss.
Demnach gilt

$$\varphi(s) = \varphi(w') + u \quad w' \in L_\varphi(w),\ u \in U.$$

Setzen wir nun $s' = s - w'$ und erkennen $s' \notin L_\varphi(w) \oplus U$, denn wäre $s' = s - w' \in L_\varphi(w) \oplus U$, dann auch $s \in L_\varphi(w) \oplus U$, also ein Widerspruch zu unserer Annahme. Somit ist

$$s' \notin L_\varphi(w) \oplus U \text{ mit } \varphi(s') = u \in U.$$

Dies bedeutet gerade, dass wir einen Vektor s' gefunden haben, der weder in U noch in $L_\varphi(w)$ liegt, und wegen $\varphi(s') = u \in U$ ist das zyklische Erzeugnis $L_\varphi(s')$ invariant.
Es existiert somit ein φ-invarianter Unterraum $U + L_\varphi(s')$ der auf Grund von $L_\varphi(s') \not\subset U$ „größer“ als U ist und somit der Maximalität von U widerspricht. $\square$

Dieses Lemma benötigen wir nun sofort für den

Satz 5.2.2. Zerlegungssatz für nilpotente Endomorphismen.
Sei φ ein nilpotenter Endomorphismus eines endlich-dimensionalen Vektorraumes W. Dann existieren Elemente $w_1, \ldots, w_k \in W$, sodass

$$W = \bigoplus_{j=1}^{k} L_\varphi(w_j).$$

Ist $r_j = \dim(L_\varphi(w_j))$ und ist die Anordnung so gewählt, dass $0 < r_1 \leq r_2 \leq \cdots \leq r_k$, so ist das Tupel $(r_1, r_2, \ldots, r_k)$ durch φ eindeutig bestimmt.

Beweis. Die gewünschte Zerlegung ergibt sich aus Lemma 5.2.1, entweder durch Induktion oder man überlege sich konzepthaft, dass man unmittelbar mit Lemma 5.2.1 einem beliebigen (endlich-dimensionalem) Vektorraum (über $\mathbb{K}$) iterativ φ-invariante Teilräume „abschöpfen“ kann.
Die Anzahl der Summanden ist $k = \dim(\ker(\varphi))$, dies ist anhand des Beweises von 4.1.2 klar. Wie man dort sieht, besitzen die dort konstruierten Basisvektoren von $\ker(\varphi)$ jeweils maximale Periode. Im Beweis von Satz 4.1.2 ist auch ersichtlich, dass $(r_1, r_2, \ldots, r_k)$ ausschließlich von φ abhängen. □

Im Vergleich zur Darstellung einer Basis $\mathcal{B}_i$ für V_i aus 4.2 werden hier die einzelnen Spalten (also das von unten nach oben gehen) zu einem α-Modul zusammengefasst. Diese Moduln angeordnet liefern dann das gleiche Ergebnis wie in 4.2 und letztlich in 4.3. Das werden wir gleich sehen.

Zusammenfassend erhält man:

Satz 5.2.3. Jordanzerlegung. Sei α ein Endomorphismus eines endlich-dimensionalen Vektorraumes V über einem Körper $\mathbb{K}$, dessen Minimalpolynom in Linearfaktoren zerfällt. Dann gibt es eine α-invariante Zerlegung wie folgt:

$$V = \bigoplus_{j=1}^{k} V(\lambda_j), \quad V(\lambda_j) = \bigoplus_{r=1}^{s_j} L_{\varphi_j}(w_{jr}), \quad \varphi_j = \alpha - \lambda_j \cdot \mathrm{id}\,.$$

Die Zerlegung in die verallgemeinerten Eigenräume $V(\lambda_j)$ und die Perioden der w_{jr} sind bis auf die Reihenfolge eindeutig durch α bestimmt.

Beweis. Dieser Beweis ist lediglich eine Zusammenfassung der bisherigen Ergebnisse. Die (eindeutige) Zerlegung $V(\lambda_j)$ ist durch Satz 5.1.3 gewährleistet, die aus der Zerlegung der einzelnen Haupträume $V(\lambda_j)$ in $\bigoplus_{r=1}^{s_j} L_{\varphi j}(w_{jr})$ resultierenden w_{jr} sind nach obigem Satz 5.2.2 eindeutig bestimmt. □

Dies führt auch sogleich zu der von uns beabsichtigten Matrixdarstellung von α.

Die Basen der zyklischen Summanden der Form $L_\varphi(w)$ werden zu einer Basis $\mathcal{B}_{w_{ji}}$ (für ein $1 \leq i \leq s_j$), einer sogenannten **Jordanbasis**, von V zusammengefügt. Für diese Basis $\mathcal{B}_{w_{ji}}$ ergibt sich analog zu Abschnitt 4.3 ein Block der Form

$$\begin{array}{ccccc} 0 & 1 & 0 & \dots & 0 \\ \vdots & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & 0 & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & 0 \end{array}$$

und demnach ist $\alpha = \lambda_j \cdot \mathrm{id} + \varphi_j$ durch den Jordanblock

$$\begin{array}{ccccc} \lambda_j & 1 & 0 & \dots & 0 \\ 0 & \lambda_j & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & \ddots & \lambda_j & 1 \\ 0 & \dots & \dots & 0 & \lambda_j \end{array}$$

gegeben.

Wir können diesen Abschnitt mit weniger Sorgfalt wie noch in Abschnitt 4.3 zu Ende bringen: es ist aus letzterem bereits bekannt, dass die gerade konstruierte Jordanbasis - über *Jordankästchen* und *Jordanmatrizen* - zur *Jordanschen Normalform* führt.

Teil II.

Darf's ein bisschen mehr Algebra sein?

Wir wollen noch eine dritte Herangehensweise zur Entwicklung der *Jordanschen Normalform* betrachten. Mein Interesse gilt bei diesem Ansatz auch stark der Einbindung und Verknüpfung der nun folgenden Theorie mit dem bereits bekannten Wissen aus Kapitel 4 und Kapitel 5, d.h. wie lässt sich also die Methode der Zerlegung eines Vektorraumes (bezüglich eines linearen Endomorphismus) aus Kapitel 7 mit den bereits bekannten Zerlegungen der vorigen Kapitel vereinbaren.

Insofern man Lineare Algebra von Algebra trennen will muss sich die Liebhaberin oder der Liebhaber der *reinen* Linearen Algebra in diesem Teil der Arbeit in Geduld üben, da die vorgestellte Theorie ihren Zusammenhang mit Vektorräumen oder Matrizen über weite Strecken verborgen hält.

Kapitel 6 liefert das notwendige Handwerkszeug aus der Algebra, welches dann in Kapitel 7 zur Vektorraumzerlegung entsprechend Anwendung finden wird. Mittels der - auf den Homomorphiesatz für Ringe gründenden - Zerlegungssätze wird die Existenz einer Zerlegung eines Vektorraumes (bezüglich eines zu betrachtenden Endomorphismus) hergeleitet. Die Erkenntnisse aus diesem Kapitel vermögen uns hinsichtlich der *Jordanschen Normalform* natürlich nichts neues zu offenbaren, liefern jedoch einen fruchtbaren Beitrag für eine *allgemeinere Normalform*.

6. Algebraische Voraussetzungen

Um nachfolgend präzise arbeiten zu können werden in diesem Kapitel die bedeutendsten Definitionen und deren Folgerungen dargelegt. Die Unterschiede zu bereits bekannten algebraischen Strukturen sind dabei oft subtilerer Natur.

6.1. Ringe

Der folgende Abschnitt beruht überwiegend auf [Art98].

Wir werden hier kurz Ringe und Ringhomomorphismen definieren. Ringe können als erste *Verallgemeinerungsstufe* von Körpern gesehen werden und es ist oft von weitreichender Folge, anstatt auf Körpern, mit Ringen zu operieren.

Definition. Ein Ring R ist eine Menge mit zwei Verknüpfungen $+$ und $\cdot$, die man Addition bzw. Multiplikation nennt und die folgenden Axiome erfüllt:

(i) Mit der Addition bildet R eine abelsche Gruppe. Das zugehörige neutrale Element wird mit 0 bezeichnet.

(ii) Die Multiplikation ist assoziativ und hat ein neutrales Element, das mit 1 bezeichnet wird.

(iii) *Distributivgesetze*: Für alle $a, b, c \in R$ gelten:

$$(a + b) \cdot c = a \cdot c + b \cdot c \text{ und } c \cdot (a + b) = c \cdot a + c \cdot b$$

Man kann in Ringen im Allgemeinen also nicht mehr auf die Existenz von inversen Elemente zurückgreifen.

Definition. Seien R und R' Ringe. Ein **Homomorphismus** $\varphi : R \to R'$ ist eine Abbildung (zwischen zwei Ringen) mit folgenden Eigenschaften:

$$\varphi(a + b) = \varphi(a) + \varphi(b), \quad \varphi(a \cdot b) = \varphi(a) \cdot \varphi(b) \quad \text{und} \quad \varphi(1_R) = 1_{R'}$$

für alle $a, b \in R$. Ein **Isomorphismus** von Ringen ist ein bijektiver Homomorphismus.

6.2. Ideale

Als fachliche Grundlage für diesen Abschnitt dient [KM09].

Mit den sogenannten Idealen sind wir bereits in Zusammenhang mit dem Minimalpolynom in Kontakt gekommen. Das Attribut *ideal* erhalten sie deshalb, weil sie sich - im Gegensatz zu bloßen Unterräumen - auf der gesamten Struktur als abgeschlossen erweisen. Hier ist nun die geeignete Stelle dies sauber zu definieren.

Definition. Eine Untergruppe I von R bezüglich $+$ heißt ein **Ideal** von R, wenn folgendes gilt:

(i) $r \cdot a \in I$ für alle $a \in I$ und $r \in R$ kurz: $RI \subseteq I$

(ii) $a \cdot r \in I$ für alle $a \in I$ und $r \in R$ kurz: $IR \subseteq I$.

Das bedeutet die Menge I ist bezüglich der Multiplikation mit Elementen aus ganz R abgeschlossen.

Beispiele.

(i) $\{0\}$ (*Nullideal*) und R (*Einsideal*) sind die sogenannten **trivialen** Ideale des Ringes R.

(ii) Die Menge $n\mathbb{Z}$ für jedes $n \in \mathbb{N}_0$ ist ein Ideal im Ring $(\mathbb{Z}, +, \cdot)$.[1]

Definition. Sei R ein Ring und $A \subseteq R$. Das sogenannte **Erzeugnis** $\langle A \rangle$ bezeichnet den kleinsten Unterraum, der A enthält.

Definition. Das Ideal I von R heißt **endlich erzeugt**, wenn es Ringelemente $a_1, \ldots, a_n \in R$ gibt, sodass $I = \langle \{a_1, \ldots, a_n\} \rangle$ gilt. Wird ein Ideal I von einem einzigen Element a erzeugt, d.h. $I = \langle \{a\} \rangle$, dann nennt man I ein **Hauptideal**.

Es lässt sich leicht zeigen, dass $\langle \{a\} \rangle = Ra := \{r \cdot a : r \in R\}$ gilt, also dass das Hauptideal $\langle \{a\} \rangle$ durch alle „Vielfachen“ von a gegeben ist.

Beispiele.

(i) Hat der Ring R ein Einselement 1, dann gilt $\langle \{1\} \rangle = R$.

(ii) Jedes Ideal von $\mathbb{Z}$ und $\mathbb{Z}_n$ ($n \in \mathbb{N}_0$) ist ein Hauptideal.

[1] Es haben sogar alle Ideale von $\mathbb{Z}$ die Form $n\mathbb{Z}$ mit $n \in \mathbb{N}_0$.

6.3. Restklassenringe

Die hier ausgeführte Theorie gründet auf [Fis13].

Wir werden bald sehen, dass gewisse Restklassenringe ein isomorphes Gegenstück zu entsprechenden Moduln bilden (Isomorphiesatz: 7.1.2). Der Bedeutsamkeit von Restklassenringe soll an dieser Stelle noch ausführlicher Rechnung getragen werden.

Sei I ein Ideal eines Ringes R. Wir betrachten die Menge

$$R/I := \{x + I \colon x \in R\}$$

und definieren auf dieser die Operationen

$$\text{Addition } (+) \colon (x + I) + (y + I) := (x + y) + I$$

und

$$\text{Multiplikation } (\cdot) \colon (x + I) \cdot (y + I) := (x \cdot y) + I.$$

Diese Menge wird mit der gerade definierten Addition zu einer abelschen Gruppe - dann als Faktorgruppe bezeichnet - und mit der Multiplikation zu einem Ring. Außerdem gibt es eine herausragende Abbildung bezüglich R und R/I. Genau dies besagt folgender

Satz 6.3.1. Sei R ein Ring, $I \subset R$ ein Ideal und $\pi : R \to R/I$ mit $x \mapsto x + I$. Dann ist R/I ein Ring und π ist surjektiv, ein Ringepimorphismus mit $\ker(\pi) = I$ Ist R kommutativ, so auch R/I; hat R ein Einselement 1, so ist $1 + I$ ein bzw. das Einselement von R/I.

Beweis. Da wir auf Äquivalenzklassen operieren, müssen wir zunächst die Wohldefiniertheit der obigen Verknüpfungen nachweisen:
Wir wählen beliebige Restklassen aus R/I für die $(x + I) = (x' + I)$ und $(y + I) = (y' + I)$ sei. Dann ist $(x - x') \in I$ und $(y - y') \in I$ (siehe nächsten Absatz). Demzufolge gilt

$$xy - x'y' = (x - x')y' + x(y - y') \in I,$$

d.h. $(xy + I) = (x'y' + I)$ bzw. $(x + I)(y + I) = (x' + I)(y' + I)$. Das Produkt $(x + I)(y + I)$ hängt somit nicht von der Wahl der Repräsentanten ab.
Der Nachweis für + kann ähnlich geführt werden.

R/I ist eine abelsche Gruppe: Assoziativität und Kommutativität bezüglich $+$ übertragen sich von R auf R/I. Das Nullelement $\bar{0}$ in R/I ist gegeben durch $\bar{0} = 0 + I = I$. Dies bedeutet auch gleich $\ker(\pi) = I$.
Die Existenz eines additiv inversen Elementes ist gesichert: für jedes $\bar{x} \in R/I$ gilt die Identität $\bar{x} + (-\bar{x}) = \bar{0}$.
R/I ist ein Ring: Assoziativität bezüglich der Multiplikation und Distributivität übertragen sich wieder von R auf R/I.

Dementsprechend ist π ein Ringhomomorphismus, denn mit

$$\pi(x) + \pi(y) = (x + I) + (y + I) = (x + y) + I = \pi(x + y)$$

und

$$\pi(x) \cdot \pi(y) = (x + I) \cdot (y + I) = (x \cdot y) + I = \pi(x \cdot y)$$

sind auch die Homomorphie-Axiome aus Abschnitt 6.1 erfüllt.
Die Surjektivität von π ist per Definition von R/I klar. □

Diese „Art“ von Ringen wird uns im kommenden Kapitel permanent begleiten.
Man bezeichnet den Ring R/I auch als **Faktorring** oder **Restklassenring** von R modulo I.
Die Elemente des Faktorringes R/I werden auch als Nebenklassen oder Äquivalenzklassen bezeichnet (zwei Elemente $x, y \in R$ liegen nämlich genau dann in derselben Nebenklasse bezüglich I wenn $(x-y) \in I$ gilt; durch $\sim_I$ mit $x \sim_I y := (x-y) \in I$ wird eine Äquivalenzrelation auf R induziert).

Die Abbildung $\pi : R \to R/I, \quad x \mapsto x + I$ wird als der *kanonische Ringhomomorphismus* bezeichnet.

6.4. Hauptidealringe und Euklidische Ringe

Die aus diesem Abschnitt resultierende Erkenntnis wird sein, dass $\mathbb{K}[t]$ *faktoriell* ist. Hierfür herangezogene Quellen sind [Art98] und [KM09], sowie [Bos14].

Definition. Ein **Integritätsbereich** ist ein vom Nullring verschiedener kommutativer Ring R mit 1, der keine Nullteiler besitzt, d.h. in einem ein derartigen Ring ist $1 \neq 0$, und aus $ab = 0$ $(a, b \in R)$ folgt $a = 0$ oder $b = 0$.

In einem Integritätsbereich R gilt die Kürzungsregel:

$$\text{Ist } ab = ac \text{ mit } a \neq 0, \text{ so ist } b = c. \tag{6.1}$$

Denn aus $ab = ac$ mit $a \neq 0$ folgt $a(b - c) = 0$, also $b = c$. □

Definition. Ein Ringelement $r \in R$ für das ein $r' \in R$ mit $r \cdot r' = r' \cdot r = 1$ existiert wird als **Einheit** bezeichnet.

Bemerkung. Der Begriff der *Teilbarkeit* in Integritätsbereichen wird wie folgt aufgefasst: ein Ringelement a *teilt* ein anderes Ringelement b (abgekürzt $a|b$), wenn ein $q \in R$ mit $b = aq$ existiert.
Das Element a ist ein *echter* Teiler von b, wenn in der Zerlegung $b = aq$ weder a noch q eine Einheit ist.

Definitionen. Sei $a \neq 0$ eine Nichteinheit von R.

(i) Das Element a heißt **irreduzibel** (unzerlegbar), wenn es keine echten Teiler besitzt: Aus $a = bq$ folgt also $a \in R^\times$ oder $b \in R^\times$ ist ($R^\times$ bezeichnet die Menge aller Einheiten in R).

(ii) Das Element a heißt **Primelement**, wenn aus $a|(b \cdot q)$ $(b, q \in R)$ sofort $a|b$ oder $a|q$ folgt.

Definition. Zwei Elemente a und a' heißen **assoziiert**, wenn jeweils ein Element das andere Element teilt, dies ist äquivalent zu $a' = \epsilon a$ mit einer Einheit $\epsilon \in R^\times$.

Lemma 6.4.1. Sei R ein Integritätsring.

(i) Jedes Primelement $p \in R$ ist unzerlegbar.

(ii) Für $a, b \in R$ gilt: $a|b \Leftrightarrow \langle b \rangle \subseteq \langle a \rangle$.[2]

Beweis. Zu (i): Es sei $p \in R$ ein Primelement mit $p = ab$ mit $a, b \in R$. Dann ist o.B.d.A. $p|a$ und somit $a = c \cdot p$ mit entsprechendem $b \in R$. Nun ist $p = a \cdot b = c \cdot p \cdot b$ und sofern R ein *Integritätsbereich* ist, folgt für $p \neq 0$ nach (6.1), dass $c \cdot b = 1$ gilt, also $b \in R^\times$ eine Einheit und p somit irreduzibel ist.
Zu (ii): $a|b \Leftrightarrow b \in Ra \Leftrightarrow \{b\} = Rb \subseteq Ra = \{a\}$. □

Die Aussage (ii) aus 6.4.1 verknüpft die Teilbarkeit von Elementen in Integritätsbereichen mit den entsprechenden Idealen (Vgl. [KM09] S. 179). Dieser Zusammenhang wird noch wichtig sein.

[2] Für ein beliebiges Element r ist $\langle r \rangle$ abkürzend für $\langle \{r\} \rangle$.

Definition. Ist jedes Ideal $I \subset R$ ein Hauptideal, dann nennt man den Integritätsring R einen **Hauptidealring**.

Lemma 6.4.2. Für ein Element $p \in R$ in einem Hauptidealring R ist folgendes äquivalent:

(i) p ist irreduzibel.

(ii) p ist ein Primelement.

Beweis. (ii) $\Rightarrow$ (i) wurde bereits durch Lemma 6.4.1 gezeigt.
(i) $\Rightarrow$ (ii): Sei also $p \in R$ irreduzibel mit $p|ab$ $(a, b \in R)$ aber nicht $p|a$. Die Summe

$$\langle a\rangle + \langle p\rangle := \{ra + sp : r, s \in R\}$$

ist selbst wieder ein Ideal, da diese eine Summe von Idealen ist. Laut Voraussetzung ist R ein Hauptideal, es gibt demnach ein Element $d \in R$ mit $\langle a\rangle + \langle p\rangle = \langle d\rangle$. Somit gilt $d|a$ und $d|p$ (siehe Lemma 6.4.1). Da p irreduzibel ist folgt $p = cd$, sodass c oder d eine Einheit ist. Angenommen c ist eine Einheit, dann gilt $d = c^{-1}p$ und aus $d|a$ bzw. $p|a$ würde ein Widerspruch zur obigen Annahme entstehen. Somit ist also d eine Einheit, d.h. es gilt $\langle a\rangle + \langle p\rangle = R$ und weiters $\langle ab\rangle + \langle pb\rangle = \langle b\rangle$ (nach Multiplikation mit b). Es existieren also geeignete $r, s \in R$ mit $rab + spb = b$ und weil p das Produkt ab teilt, folgt auch $p|b$. □

Satz 6.4.3. In einem Hauptidealring R lässt sich jedes Element $a \in R \setminus \{0\}$, welches keine Einheit ist, als endliches Produkt von Primelementen schreiben.

Beweis. Nach Lemma 6.4.2 ist jedes irreduzible Element aus R bereits ein Primelement, somit genügt uns eine Zerlegung in irreduzible Elemente. Sei $a \neq 0$ eine Nichteinheit von R. Angenommen a ließe sich nicht als ein endliches Produkt irreduzibler Elemente schreiben, dann lässt sich folglich $a = a_1 \cdot a_1'$ mit Nichteinheiten $a_1 \cdot a_1' \in R$ darstellen. Da a keine endliche Faktorisierung in irreduzible Elemente besitzt muss dies auch o.B.d.A. für a_1 gelten. Dieses Element kann wiederum in $a_1 = a_2 \cdot a_2'$ mit Nichteinheiten zerlegt werden. Iterativ liefert dieses Verfahren eine Folge von Elementen

$$a = a_0, a_1, a_2, \cdots \in R,$$

sodass a_{i+1} jeweils ein echter Teiler von a_i (also nicht assoziiert) ist. Dies generiert weiters eine aufsteigende Folge von Idealen

$$\langle a\rangle \subsetneq \langle a_1\rangle \subsetneq \langle a_2\rangle \dots \ .$$

Die Vereinigung $\mathfrak{b} = \bigcup_{i=0}^{\infty} \langle a_i\rangle$ ergibt wieder ein Ideal und muss laut Voraussetzung durch ein Element b erzeugt werden, d.h. $\mathfrak{b} = \langle b\rangle$. Dies führt sofort zu einem Widerspruch: $\mathfrak{b}$ müsste nach Konstruktion das „letzte“ Ideal in obiger Folge sein, was allerdings laut Annahme nicht möglich ist.

□

Aus dem bereits Gesagten und der letzten Beweisführung ergibt sich schon fast der folgende

Satz 6.4.4. Sei R ein Integritätsbereich. Dann sind die folgenden Bedingungen äquivalent:

(i) Jede Nichteinheit $a \neq 0$ von R besitzt eine endliche Zerlegung $a = a_1 \cdot \ldots \cdot a_m$, wobei $a_1, \ldots a_m$ irreduzible Faktoren von R sind.

(ii) Der Ring R enthält keine unendlich aufsteigende Hauptidealkette

$$\langle b_1 \rangle \subsetneq \langle b_2 \rangle \subsetneq \langle b_3 \rangle \subsetneq \ldots \ ,$$

d.h. diese aufsteigende Kette wird *stationär*.

(iii) Jedes unzerlegbare Element von R ist ein Primelement.

Dieser Satz ist zwar aus algebraischer Sicht für faktorielle Ringe bedeutend, für uns aber nur noch bedingt von Interesse. Aus diesem Grund verzichten wir auch auf einen vollständigen Beweis und belegen nur die Äquivalenz von (i) und (ii).

Beweis. Die Äquivalenz von (i) und (ii) bekommen wir aus der Beweisführung von Satz 6.4.3, mit dem Zusatz, dass eine unendlich aufsteigende Kette von Hauptidealen umgekehrt auch eine unendliche Zerlegung eines Elementes a liefert.
Eine vollständige Ausführung (allerdings nicht hier anknüpfend) findet sich in [KM09, S. 182]. □

Definition. Ein Integritätsbereich, in dem jede Nichteinheit $a \neq 0$ ein Produkt (bis auf Assoziation eindeutig) unzerlegbarer Elemente ist, wird als **faktorieller Ring** bezeichnet.

Das folgende Lemma verdeutlicht den Zusatz *bis auf Assoziation eindeutig*.

Lemma 6.4.5. In einem Integritätsring R habe man die Gleichung

$$p_1 \cdot \ldots \cdot p_r = q_1 \cdot \ldots \cdot q_s$$

für Primelemente $p_1, \ldots, p_r \in R$ und irreduzible Elemente $q_1, \ldots, q_s \in R$. Dann gilt $r = s$ und nach Umnummerierung der $q_1, \ldots, q_s$ existieren Einheiten $\epsilon_1, \ldots, \epsilon_s \in R^\times$ mit $q_i = \epsilon_i \cdot p_i$ für $i = 1, \ldots, r$, d.h. p_i ist jeweils assoziiert zu q_i.

Beweis. Aus $p_1 \cdot \ldots \cdot p_r = q_1 \cdot \ldots \cdot q_s$ folgt insbesondere, dass p_1 das Produkt $q_1 \cdot \ldots \cdot q_s$ teilt. Da p_1 ein Primelement ist, teilt es sogar einen Faktor dieses Produktes und durch entsprechende Umnummerierung erhält man $p_1 | q_1$. Da q_1 irreduzibel ist,

muss ϵ_1 in $q_1 = \epsilon_1 \cdot p_1$ eine Einheit sein. In einem Integritätsring ergibt sich mit der Kürzungsregel somit

$$p_2 \cdot \ldots \cdot p_r = \epsilon_1 \cdot q_2 \cdot \ldots \cdot q_s.$$

So fortfahrend gelangt man zu der Gleichung

$$1 = q_{r+1} \cdot \ldots \cdot q_s.$$

Diese Gleichung schildert $q_{r+1}, \ldots, q_s$ als Einheiten aus, was laut Voraussetzung aber nicht möglich ist, folglich gilt $r = s$. □

Bemerkung. Aus schulischer Sicht erscheint die Eindeutigkeit einer Primfaktorzerlegung auf Grund der Assoziation wohl eingeschränkt. Wie man sich anhand von Polynomen vorstellen kann muss eine Zerlegung in Primelemente für allgemeine Ringe eben nicht immer so „schön“ verlaufen wie es im Ring der ganzen Zahlen $\mathbb{Z}$ der Fall ist. In diesem gibt es die Assoziations-Einschränkung im eigentlichen Sinne nicht, da hier nur $\pm 1 \in \mathbb{Z}^\times$ sind (also ± 1 sind die einzigen Einheiten in $\mathbb{Z}$).

Aus Satz 6.4.3 ergibt sich nun unmittelbar folgende Erkenntnis:

Satz 6.4.6. Jeder Hauptidealring ist faktoriell.

Definition. Ein Integritätsring R heißt **euklidischer Ring**, wenn es eine Abbildung $\delta : R \setminus \{0\} \to \mathbb{N}$ mit folgender Eigenschaft gibt: Zu beliebigen $a, b \in R \setminus \{0\}$ existieren $q, r \in R$ mit

$$a = q \cdot b + r \quad \text{und} \quad r = 0 \text{ oder } \delta(r) < \delta(b).$$

Bemerkung. Die Abbildung δ wird auch als *Gradabbildung* bezeichnet.

Satz 6.4.7. Jeder euklidische Ring ist ein Hauptidealring.

Beweis. Sei R euklidisch und $I \subset R$ ein Ideal. Da $\{0\}$ ein Hauptideal ist kann man $I \neq \{0\}$ annehmen. Wir wählen ein Element $a \in I \setminus \{0\}$ mit minimalen Grad $k = \delta(a)$, d.h. für jedes $c \in I$ gilt $\delta(a) = k \leq \delta(c)$. Wir zeigen nun, dass ein solches Element bereits das Ideal I erzeugt.
Angenommen es gäbe ein $b \in I \setminus \langle a \rangle$. Die Division mit Rest für a und b ergibt

$$b = q \cdot a + r \text{ mit } \delta(r) < \delta(a) = k.$$

Da $b \notin \langle a \rangle$ muss $r \neq 0$ gelten. Demnach hätten wir aber mit $r = b - q \cdot a \in I$ ein Element gefunden, welches der Minimalität von $\delta(a) = k$ widerspricht. Somit folgt $\langle a \rangle = I$. □

Mit diesem Rüstzeug können wir nun folgern, dass der Polynomring $\mathbb{K}[t]$ faktoriell ist. Dies besagt unter anderem das folgende

Korollar 6.4.8. Sei $\mathbb{K}$ ein Körper, und $\mathbb{K}[t]$ der Polynomring in einer Variablen.

(i) Sind die Polynome f und g teilerfremd, d.h. f und g besitzen keinen nichtkonstanten gemeinsamen Teiler, dann gibt es Polynome $r, s \in \mathbb{K}[t]$ mit $rf + sg = 1$.

(ii) Zu jedem Polynom $0 \neq f \in \mathbb{K}[t]$ gibt es eine ganze Zahl $k \geq 0$ und eine Produktzerlegung der Form

$$f = c \cdot p_1 \cdot \ldots \cdot p_k,$$

wobei $c \in \mathbb{K} \setminus \{0\}$ ist und p_i normierte irreduzible Polynome aus $\mathbb{K}[t]$ sind. Diese Produktdarstellung ist bis auf die Reihenfolge der Faktoren eindeutig.

Bevor wir noch ein für den Beweis notwendiges Lemma formulieren können, brauchen wir die

Definition. Sei R ein Integritätsbereich und $a_1, \ldots, a_n \in R$. Das Element $d \in R$ wird als ein größter gemeinsamen Teiler der a_i bezeichnet und mit

$$\mathrm{ggT}(a_1, \ldots, a_n) = d$$

abgekürzt, falls d alle a_i teilt, und für jedes $c \in R$, welches alle a_i teilt, stets auch d teilt.

Lemma 6.4.9. Seien $f, g \neq 0 \in R$ und R ein Hauptidealring. Dann gilt für $d = \mathrm{ggT}(f, g)$

$$\langle f \rangle + \langle g \rangle = \langle d \rangle.$$

Beweis. Das Ideal $\langle f \rangle + \langle g \rangle$ wird von einem Element d' erzeugt. Es existieren somit $r, s, \in R$ mit $rf + sg = d'$, jeder gemeinsame Teiler von f, g ist somit auch Teiler von d' ist. Da $f, g \in \langle d' \rangle$ ist d' auch ein Teiler von $d = \mathrm{ggT}(f, g)$, insgesamt ergibt sich also $d|d'$ und $d'|d$. Daraus folgern wir mit 6.4.1 die gewünschte Identität. □

Nun zum

Beweis von Korollar 6.4.8. (i) ergibt sich mit $\mathrm{ggT}\,(f, g) = 1$ unmittelbar aus Lemma 6.4.9.
(ii) folgt sofort aus Satz 6.4.6 und Satz 6.4.7, da $\mathbb{K}[t]$ ein Euklidischer Ring ist (*Euklidischer Algorithmus* 2.4.1). Der Beweis dieser Tatsache könnte auch ohne Verwendung des Hauptidealargumentes geführt werden (Vgl. hierzu [Art98] S. 448). □

6.5. Moduln

Hauptsächlich entstand dieser Abschnitt unter Verwendung von [Art98], sowie in kleinerem Ausmaß von [Bos14].

Die sogenannten Moduln[3] werden eine außerordentliche Rolle im nächsten Kapitel spielen. Sie stellen gewissermaßen eine Verallgemeinerung der Vektorräume über einem Körper $\mathbb{K}$ dar.

Definition. Sei R ein Ring. Ein **R-Modul** V ist eine abelsche Gruppe, deren Verknüpfung $+$ als Addition (auch: *innere Verknüpfung*) bezeichnet wird, zusammen mit einer Skalarmultiplikation $R \times V \to V, (r, v) \mapsto r \cdot v$ (auch als *äußere Verknüpfung* bezeichnet), die folgende Axiome erfüllt:

(i) $1 \cdot v = v$.

(ii) $(r \cdot s) \cdot v = r \cdot (s \cdot v)$,

(iii) $(r + s) \cdot v = r \cdot v + s \cdot v$,

(iv) $r \cdot (v + v') = r \cdot v + r \cdot v'$

für alle $r, s \in R$ und $v, v' \in M$.

Sei $\mathbb{K}$ ein Körper, im Falle $R = \mathbb{K}$ entsprechen diese Axiome genau den $\mathbb{K}$-Vektorraum-Axiomen. Moduln können somit als Vektorräume über Ringen interpretiert werden; trotzdem stellt, durch die Tatsache (oder eher Konstruktion), dass Ringelemente im Allgemeinen kein Inverses besitzen, ein Modul eine kompliziertere Struktur dar. Dies werden wir anhand der *Torsion* später noch sehen.
Vergleiche hierzu [Bos14, S. 189].

Satz 6.5.1. Die Untermoduln des R-Moduls R sind die Ideale von R.

Vergleiche hierfür [Art98, S. 517]. □

Definition. Seien M und M' R-Moduln. Ein **R-Modulhomomorphismus** $\varphi : M \to M'$ ist eine Abbildung, die mit den Verknüpfungen verträglich ist, d.h. folgende Eigenschaften besitzt:

$$\varphi(m + \tilde{m}) = \varphi(m) + \varphi(\tilde{m}) \quad \text{und} \quad \varphi(r \cdot m) = r \cdot \varphi(m)$$

[3] Im Gegensatz zur üblichen Betonung des Begriffes *Modul* wird *der* Modul in der Mathematik auf der ersten Silbe betont, und die Pluralform lautet *Moduln*.

für alle $m, \tilde{m} \in M$ und $r \in R$. Ein *R-Modulhomomorphismus* ist somit das Analogon zu einer linearen Abbildung zwischen Vektorräumen.

Die Definitionen von **Kern** und **Bild** können ohne Abänderung (bezüglich der Definition für lineare Abbildungen) übernommen werden. Auch *Kern* und *Bild* eines *R-Modulhomomorphismus* bilden Untermoduln von M bzw. M' (entsprechend den Unterräumen bei linearen Abbildungen).

Ein **R-Modulisomorphismus** von Ringen ist ein bijektiver R-Modulhomomorphismus.

Gewisse Begrifflichkeiten, die wir bereits aus Vektorräumen kennen, lassen sich auf Moduln übertragen. So sind die Definitionen *Linearkombination* oder *Basis* auf für Moduln sinnvoll und einfach aus den Vektorräumen über $\mathbb{K}$ bedeutungserhaltend übertragbar.

Definition. Ein Modul, der eine endliche Basis besitzt heißt **endlich frei**.

Die Existenz einer Basis für einen Modul ist im Gegensatz zu Vektorräumen nicht gewährleistet.

Naheliegenderweise kann man das in Abschnitt 2.1 vorgestellte Konzept zur Beschreibung eines Homomorphismus (bezüglich zweier Basen) auch bezüglich freier Moduln anwenden, d.h. ein R-Homomorphismus lässt sich (bezüglich gewählter Basen) durch eine Matrix mit Koeffizienten aus R beschreiben.

7. Jordansche Normalform - abstrakt

Das folgende Kapitel basiert - sofern nicht anders vermerkt - ausschließlich auf [Bos14].

Vorweg sollen an dieser Stelle einige Worte fallen, die Aufschluss, über das in diesem Kapitel behandelte Programm, geben.

Wie wir bereits gesehen haben, können wir aus einer entsprechenden direkten Summe invarianter Teilräume (bezüglich eines linearen Endomorphismus) die *Jordansche Normalform* entwickeln. Somit wird der Herleitung einer derartigen Zerlegung erneut unser Interesse gelten.

Wir zeigen mit Hilfe der Abschnitte 7.1, 7.2 und 7.3, dass jeder Modul über einem Hauptidealring isomorph zu einem freien Modul plus Torsionsmodul ist. Der Torsionsmodul ist wiederum isomorph zu einer direkten Summe von Restklassenringen. In Abschnitt 7.4 erfahren wir dann, dass jeder Vektorraum (über einem Körper $\mathbb{K}$) als Modul über dem Hauptidealring der Polynome $\mathbb{K}[t]$ aufgefasst werden kann und somit auch einen isomorphen Zusammenhang zu den gerade erwähnten Restklassenringen besitzt. Im Speziellen schließen wir dann, dass ein Vektorraum V als $\mathbb{K}[t]$-Modul zu einem Torsionsmodul isomorph ist. Die Untermoduln dieses Torsionsmodul entsprechen den - in Kapitel 4 genauer beschriebenen - Haupträumen.

Zu guter Letzt wollen wir zeigen, dass die in diesem Kapitel hergeleitete Zerlegung prinzipiell denen aus Kapitel 4 und Kapitel 5 entspricht bzw. demonstrieren, dass die Herangehensweisen an die *Jordanzerlegung* aus den vorigen Ansätzen in der Theorie dieses Kapitels wiederzufinden sind.

7.1. Homomorphiesatz

Satz 7.1.1. Homomorphiesatz für Vektorräume. Es sei U ein Unterraum eines $\mathbb{K}$-Vektorraumes V und $\pi : V \to \bar{V}$ der kanonische Epimorphismus auf den Restklassenraum $\bar{V} = V/U$ (also $\pi : V \to \bar{V}$ mit $v \mapsto v + U$, siehe Abschnitt 6.3). Dann existiert zu jeder linearen Abbildung $\alpha : V \to V'$ mit $U \subset \ker(\alpha)$ genau eine lineare Abbildung $\bar{\alpha} : \bar{V} \to V'$, sodass $\alpha = \bar{\alpha} \circ \pi$ gilt. Außerdem ist $\bar{\alpha}$ genau dann injektiv, wenn $U = \ker(\alpha)$ gilt und genau dann surjektiv, wenn α surjektiv ist.

Beweis. Existenz: Vorweg zur Erinnerung: $\ker(\pi) = U$ (vergleiche Satz 6.3.1).
Zunächst zeigen wir, dass man unabhängig von der Wahl zweier Urbilder $x, \tilde{x} \in \pi^{-1}(\bar{x})$ das gleiche Bild unter α erhält: Sei $\bar{x}$ eine Restklasse aus $\bar{V}$ und $x, \tilde{x} \in \pi^{-1}(\bar{x}) \subseteq V$ zwei (beliebige) Urbilder. Dann folgt

$$x - \tilde{x} \in \ker(\pi) = U \subset \ker(\alpha)$$

und auf Grund der Linearität von α, dass $0 = \alpha(x - \tilde{x}) = \alpha(x) - \alpha(\tilde{x})$, also $\alpha(x) = \alpha(\tilde{x})$ gilt. Damit ist es möglich, eine Abbildung $\bar{\alpha} : \bar{V} \to V'$ durch $\bar{x} \mapsto \alpha(x)$ zu erklären, wobei x ein beliebiger Repräsentant aus $\bar{x}$ (also ein π-Urbild) ist. Nach dieser Definition ist dann $\bar{\alpha}(\pi(x)) = \alpha(x)$ für jedes $x \in V$, erfüllt somit die Forderung $\alpha = \bar{\alpha} \circ \pi$.
Die Abbildung $\bar{\alpha}$ ist linear:
Für zwei (beliebige) Restklassen $\bar{x}, \bar{y} \in \bar{V}$ und entsprechenden (beliebigen) Repräsentanten $x, y \in V$ ist $x + y$ ein π-Urbild (oder Repräsentant) für $\bar{x} + \bar{y}$, sowie $\mu \cdot x$ ($\mu \in \mathbb{K}$) ein π-Urbild für $\mu \cdot \bar{x}$. Demgemäß ist

$$\bar{\alpha}(\bar{x} + \bar{y}) = \alpha(x + y) = \alpha(x) + \alpha(y) = \bar{\alpha}(\bar{x}) + \bar{\alpha}(\bar{y})$$

und

$$\bar{\alpha}(\mu \cdot \bar{x}) = \alpha(\mu \cdot x) = \mu \cdot \alpha(x) = \mu \cdot \bar{\alpha}(\bar{x}).$$

Sei $\bar{\alpha}$ injektiv. Dann gilt für ein Kern-Element $x \in \ker(\alpha)$ die Beziehung $\bar{\alpha}(\pi(x)) = \alpha(x) = 0$ und somit liegt x auch im Kern von π, es gilt nämlich $\pi(x) = 0$ (für jede lineare Abbildung f ist Injektivität äquivalent zu $\ker(f) = \{0\}$). Das heißt $\ker(\alpha) \subset \ker(\pi) = U$, und die umgekehrte Inklusion gilt sowieso ($\alpha(\pi(y)) = \alpha(0) = 0$ für $y \in \ker(\pi)$), daher insgesamt $\ker(\alpha) = \ker(\pi) = U$.
Gelte umgekehrt $U = \ker(\alpha)$. Dann erhält man für ein beliebig gewähltes $\bar{x} \in \ker(\bar{\alpha})$ und einem (dazugehörenden) Repräsentanten x die Gleichung $\alpha(x) = \bar{\alpha}(\pi(x)) = \bar{\alpha}(\bar{x}) = 0$, also $x \in \ker(\alpha) = U$. Wegen $U = \ker(\pi)$ können wir mit $\bar{x} = \pi(x) = 0$ folgern, dass $0 \in \bar{V}$ das einzige Element im Kern von $\bar{\alpha}$ ist (d.h. $\ker(\alpha) = \{0\}$) und somit die Injektivität von $\bar{\alpha}$ gilt.
Sei $\bar{\alpha}$ surjektiv, dann ist auch $\alpha = \bar{\alpha} \circ \pi$ surjektiv. Umgekehrt folgt aus der Surjektivität von α auch die von $\bar{\alpha}$.
Eindeutigkeit: $\bar{\alpha}(x + U) = \bar{\alpha}(\pi(x)) = \bar{\alpha} \circ \pi(x) = \alpha(x)$. □

Dieser Satz ist ohne viel Aufwand für Ringe zu formulieren:

Satz 7.1.2. Homomorphiesatz für Ringe. Es sei $\varphi : R \to R'$ ein Ringhomomorphismus, $I \subset R$ ein Ideal mit $I \subset \ker(\varphi)$ und $\pi : R \to R/I$. Dann existiert ein eindeutig bestimmter Ringhomomorphismus $\bar{\varphi} : R/I \to R'$, sodass $\varphi = \bar{\varphi} \circ \pi$ gilt. Dabei ist $\bar{\varphi}$ genau dann injektiv, wenn $I = \ker(\varphi)$ gilt und genau dann surjektiv, wenn φ surjektiv ist.

Beweis. Bei näherer Betrachtung des Beweises für den Homomorphiesatz für Vektorräume (7.1.1) bemerkt man, dass keine Eigenschaften der Vektorraum-Elemente benötigt wurden, d.h. ob wir uns wie oben einen Vektor oder im Bezug auf diesen Satz ein Element aus einem Ring „vorstellen“ ist unerheblich.
Ein kleines Detail noch: Da in Satz 7.1.2 die Relation $I \subset \ker(\varphi)$ gleichermaßen gefordert wird, ist die Unabhängigkeit von $\varphi(x)(= \bar{\varphi}(\bar{x}))$ bezüglich des Repräsentanten x auch für die Ring-Version des Homomorphiesatzes gewährleistet. □

An dieser Stelle bietet es sich an, den fehlenden Teil des Beweises von Satz 3.1.1 nachzuholen. Die darin verwendete *induzierte Abbildung* erinnert nämlich stark an die gerade behandelten Homomorphismen.

Zu einem Homomorphismus α gibt es einen eindeutigen Homomorphismus $\underline{\alpha} : V/W \to V/W$, sodass $\pi \circ \alpha = \underline{\alpha} \circ \pi$ gilt, wobei π den kanonischen Epimorphismus bezeichnet. Man kann ohne größeren Aufwand direkt nachprüfen, dass $\underline{\alpha}(v+W) = \alpha(v) + W$ die gewünschten Eigenschaften (Wohldefiniertheit, Linearität und obige Identität) erfüllt. Diesen Homomorphismus nennt man *induzierte Abbildung*.

Nun zum fehlenden Teil des

Beweises von 3.1.1. Wir bilden die Basiselemente $v_{k+1}, \ldots, v_n$ durch π ab und erhalten eine Basis $\mathcal{B}'' = \{\pi(v_{k+1}), \ldots, \pi(v_n)\}$ für V/W. Mit

$$\dim(\underbrace{V/W}_{=\mathrm{Im}(\pi)}) = \dim(V) - \dim(\underbrace{W}_{=\ker(\pi)}) = n - k,$$

bzw. auf Grund der Surjektivität von π folgt nämlich, dass $\mathcal{B}''$ ein minimales Erzeugendensystem, also eine Basis für V/W, ist.

Die Form von $M_{\mathcal{B}}(\alpha) = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ liefert für $j = 1, \ldots, n-k$ die Identität

$$\alpha(v_{k+j}) = b_{1,j}v_1 + \cdots + b_{k,j}v_k + c_{1,j}v_{k+1} + \cdots + c_{n-k,j}v_n,$$

wobei $B = (b_{ij})$ und $C = (c_{ij})$. Somit ist

$$\underline{\alpha}(\pi(v_{k+j})) = \pi(\alpha(v_{k+j})) = c_{1,j}\pi(v_{k+1}) + \cdots + c_{n-k,j}\pi(v_n),$$

also C die Matrixdarstellung von $\underline{\alpha}$ bezüglich $\mathcal{B}''$. □

7.2. Elementarteilersatz

Satz 7.2.1. Elementarteilersatz. Es sei R ein Hauptidealring und F ein endlich 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 $\nu_1, \dots, \nu_s \in R \setminus \{0\}$, sodass gilt:

1. $\nu_1 x_1, \dots, \nu_s x_s$ bilden eine Basis von M.
2. $\nu_i | \nu_{i+1}$ für $1 \leq i < s$.

Dabei sind $\nu_1, \dots, \nu_s$ bis auf Assoziiertheit, sowie die Anzahl s eindeutig durch M bestimmt, unabhängig von der Wahl von $x_1, \dots x_s$.

Beweisskizze. Man zeigt zunächst, dass im Falle der Existenz einer endlichen Basis für den Modul F, auch Untermoduln $M \subset F$ endlich erzeugt sind: Ein Erzeugendensystem von $M \cap \ker(\pi)$ und ein Erzeugendensystem von $\mathrm{Im}(M)$ bilden insgesamt ein Erzeugendensystem von M, wobei π die Projektion auf das Erzeugnis eines Basiselementes y_i von $F = \langle y_1, \dots, y_m \rangle =: \langle \mathcal{Y} \rangle$ bezeichnet.
Sei nun $\{z_1, \dots, z_n\}$ ein Erzeugendensystem von M. Wir definieren eine lineare Abbildung f durch $f : R^n \to F$ mit $e_j \mapsto z_j$, wobei $\mathcal{E} = \{e_1, \dots, e_n\}$ die kanonische Basis des R^n bezeichnet. Gilt dann:

$$z_j = \sum_{i=1}^{m} \nu_{ij} y_i, \qquad \text{für } j = 1, \dots, n,$$

dann ist in der Nomenklatur von Abschnitt 2.1 die Matrixdarstellung von f gegeben durch $[f]_{\mathcal{Y}}^{\mathcal{E}} = (\nu_{ij})_{i,j} \in R^{m \times n}$.
Das Lemma 7.2.2 liefert uns dann eine Darstellung von f bezüglich einer anderen Basis $\{e'_1, \dots, e'_n\}$ für R^n und einer anderen Basis $\{x_1, \dots, x_m\}$ für F. Infolgedessen gilt, dass M als Bild von f durch $\{\nu_1 x_1, \nu_2 x_2, \dots, \nu_s x_s\}$ erzeugt wird. Da die Elemente $x_1, \dots, x_m$ linear unabhängig sind, bildet $\{\nu_1 x_1, \dots, \nu_s x_s\}$ sogar eine Basis für M.

Die komplette Variante dieser Beweisskizze findet sich in [Bos14, S. 208]. □

Lemma 7.2.2 (Smith-Normalform)**.** Es sei R ein Hauptidealring und $(\nu_{ij})_{i,j} \in R^{m \times n}$ 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

$$S \cdot A \cdot T = \begin{pmatrix} \nu_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \nu_2 & & 0 & \vdots & \dots & 0 \\ \vdots & 0 & \ddots & \vdots & 0 & \dots & 0 \\ \vdots & \vdots & & \nu_s & & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

und Koeffizienten $\nu_1, \dots, \nu_s \in R \setminus \{0\}$ (wobei $0 \leq s \leq \min(m, n)$ gilt), die für $1 \leq i < s$ die Bedingung $\nu_i | \nu_{i+1}$ erfüllen. Dabei sind $\nu_1, \dots, \nu_s$ bis auf Assoziiertheit eindeutig bestimmt.

Ein Beweis findet sich in [Bos14, S. 210]. □

Bemerkung. Die Ring-Elemente $\nu_1, \nu_2, \dots, \nu_s$ nennt man **Elementarteiler**.[1] Diese liefern nicht „nur“ ein Hilfsmittel für Beweise in folgendem Abschnitt, sondern sogar eine numerisch geeignetere Möglichkeit der Klassifikation von Matrizen (anstatt der Bestimmung der Eigenwerte). Siehe hierzu [Bri85, S. 89].

7.3. Strukturs̈atze

Anders ausgedrückt besagt das folgende Lemma:
Ist eine beliebige algebraische Struktur isomorph zu einer direkten Summe $\bigoplus R/\langle \nu_i \rangle$ (R Hauptidealring) mit Nichteinheiten $\nu_1, \dots, \nu_s \in R \setminus \{0\}$, die $\nu_i | \nu_{i+1}$ erfüllen, dann ist diese direkte Summe eindeutig.

Lemma 7.3.1. Sei R ein Hauptidealring und

$$Q \cong \bigoplus_{i=1}^{s} R/\langle \nu_i \rangle$$

ein Isomorphismus von R-Moduln, wobei $\nu_1, \dots, \nu_s \in R \setminus \{0\}$ Nichteinheit mit $\nu_i | \nu_{i+1}$ für $1 \leq i < s$ sind und $\oplus_{i=1}^{s} R/\langle \nu_i \rangle$ die konstruierte direkte Summe der R-Moduln $R/\langle \nu_i \rangle$ bezeichne.
Dann sind $\nu_1, \dots, \nu_s$ bis auf Assoziiertheit eindeutig durch Q bestimmt.

[1] Eigentlich die *Elementarteiler von M.* Die Diagonaleinträge der Matrix $S \cdot A \cdot T$ aus Lemma 7.2.2 nennt man Elementarteiler der Matrix $[f]_{\mathcal{Y}}^{\mathcal{E}} = (\nu_{ij})_{i,j}$, sie sind aber offensichtlich gleich den Elementarteilern von M.

Ein Beweis findet sich in [Bos14] S. 213. □

Ergänzend noch ein Wort zur eben erwähnten *konstruierten direkten Summe*. Es sei R ein Ring und $N_1, \ldots, N_k$ R-Moduln, die weder unbedingt $N_i \cap N_j \neq \{0\}$ erfüllen müssen, noch notwendigerweise Untermoduln eines R-Moduls N sind. Das kartesische Produkt $N_1 \times \cdots \times N_k$, welches dann selbst wieder ein R-Modul ist (unter komponentenweise Addition und skalare Multiplikation), kann nun als *direkte Summe* aufgefasst werden. Wir *konstruieren* mit dem Setzen von

$$\tilde{N}_i := (0 \times \cdots \times N_i \times \cdots \times 0) \subset N_1 \times \cdots \times N_k$$

für $i = 1, \ldots k$, Untermoduln, für die dann gilt:

$$\tilde{N}_1 \oplus \cdots \oplus \tilde{N}_k = N_1 \times \cdots \times N_k.$$

Sind die $N_1, \ldots N_k$ keine Untermoduln des selben R-Moduls N, so spricht man von der *äußeren* direkten Summe.

Diese Konstruktion spielt eine entscheidende Rolle, sie erlaubt nämlich künftig alle Beweise, die *Isomorphie* zu einem kartesischen Produkt liefern, demnach auch als Korrespondenz (=Isomorphie) zu einer direkten Summe aufzufassen.

Für den späteren Hauptsatz dieses Abschnittes benötigen wir weitere Hilfsmittel. Eines ist der folgende

Satz 7.3.2 (Chinesischer Restsatz)**.** Es sei R ein Hauptidealring und

$$a = \epsilon \cdot p_1^{n_1} \cdot \ldots \cdot p_r^{n_r}$$

eine Primfaktorzerlegung in R mit einer Einheit $\epsilon \in R$ und paarweise nicht-assoziierten Primelementen $p_i \in R$. Bezeichne dann

$$\pi_i : R \to R/\langle p_i^{n_i} \rangle \quad \text{für } i = 1, \ldots, r,$$

jeweils die kanonische Projektion (siehe 6.3.1), so ist der Homomorphismus

$$\varphi : R \to R/\langle p_1^{n_1} \rangle \times \cdots \times R/\langle p_r^{n_r} \rangle, \quad b \mapsto (\pi_1(b), \ldots, \pi_r(b))$$

surjektiv und erfüllt $\ker(\varphi) = \langle a \rangle$, induziert demgemäß den isomorphen Zusammenhang

$$R/\langle a \rangle \xrightarrow{\sim} R/\langle p_1^{n_1} \rangle \times \cdots \times R/\langle p_r^{n_r} \rangle.$$

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

Beweis. Wir belegen zunächst die Surjektivität von φ.
Finden wir Elemente $e_1, \ldots, e_r \in R$, die auf die kanonische Basis von $R/\langle p_1^{n_1}\rangle \times \cdots \times R/\langle p_r^{n_r}\rangle$ abbilden, so würde die Surjektivität von φ folgen (eine Basis wäre Teilmenge des Bildes des Homomorphismus φ). Wir hätten also gerne

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

Wir konstruieren nun einfach Elemente, die, wie wir gleich sehen werden, obige Beziehung erfüllen: Die Elemente $p_j^{n_j}$ und $\prod_{i\neq j} p_i^{n_i}$ sind (für ein fest gewähltes j) teilerfremd. Dementsprechend gibt es nach Lemma 6.4.9 ein Element d_j, sodass

$$e_j + d_j = 1 \qquad \text{mit } e_j \in \langle \prod_{i\neq j} p_i^{n_i}\rangle \text{ und } d_j \in \langle p_j^{n_j}\rangle$$

gilt. Wir erhalten damit $\pi_i(e_j) = 0$ für $i \neq j$ und $\pi_i(e_i) = 1$.
Ein Element $b \in R$ liegt genau dann im Kern von φ, wenn es jeweils im Kern der kanonischen Projektion π_i liegt, d.h. $\varphi(b) = 0$ ist gleichwertig zu $\pi_i(b) = 0$ für alle $i = 1, \ldots, r$. Damit wird ein Element b mit $\varphi(b) = 0$ durch $p_1^{n_1} \cdot \ldots \cdot p_r^{n_r}$ geteilt. Das bedeutet $\ker(\varphi) = \langle a\rangle$. Der Homomorphismus φ ist nach dem Homomorphiesatz 7.1.2 bijektiv und induziert somit die behauptete Isomorphie. □

Wir können den Ausdruck $R/\langle p_1^{n_1}\rangle \times \cdots \times R/\langle p_r^{n_r}\rangle$ entsprechend voriger Bemerkung (nach 7.3.1) als direkte Summe auffassen und erhalten mit Satz 7.3.2 dann sogar

$$R/\langle a\rangle \xrightarrow{\sim} R/\langle p_1^{n_1}\rangle \oplus \cdots \oplus R/\langle p_r^{n_r}\rangle.$$

Und das kommt unserem Vorhaben - eine geeignete direkte Zerlegung eines Vektorraumes zu finden - schon recht nahe.

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

$$T = \{a \in M : \text{ es existiert ein } \mu \in R \setminus \{0\} \text{ mit } \mu \cdot a = 0\}.$$

Der Torsionsmodul eines freien Moduls ist trivial, denn für ein Element $a \in M \setminus \{0\}$ (M besitzt eine Basis $\{a_1, \ldots, a_n\}$, ist also frei) existiert stets eine Darstellung $a = \mu_1 \cdot a_1 + \cdots + \mu_n \cdot a_n$ mit mindestens einem $\mu_i \neq 0$, somit kann $\mu a = 0$ nur für $\mu = 0$ gelten.
Freie „Moduln sind daher sozusagen als das ‚Gegenstück' zu den Torsionsmoduln anzusehen". (Vgl. [Bos14, S. 218]).

Andererseits könnte man informell anmerken, dass klarerweise so etwas wie $\mu \cdot a = 0$ für $\mu \in R \setminus \{0\}$ und $a \in N$ „passieren“ kann, wenn der Modul N keine Basis besitzt.

In diesem Sinne zeigt der nächste Satz, dass ein Modul in seinen Torsionsmodul und einen freien Modul (direkt) zerlegt werden kann. Auch die Struktur des Torsionsmodul wird durch den folgenden Satz näher erklärt.

Satz 7.3.3. Es sei M ein endlich erzeugter Modul über einem Hauptidealring R und $T \subset M$ sein Torsionsmodul. Dann gibt es einen endlichen freien Untermodul $F \subset M$, etwa $F \cong R^d$, sowie Nichteinheiten $\nu_1, \dots, \nu_s \in R \setminus \{0\}$ mit $\nu_i | \nu_{i+1}$ für $1 \leq i < s$ und

$$M = F \oplus T, \quad T \cong \bigoplus_{i=1}^{s} R/\langle \nu_i \rangle.$$

Dabei ist d eindeutig bestimmt. Außerdem sind $\nu_1, \dots, \nu_s$ eindeutig bis auf Assoziiertheit bestimmt.

Wir brauchen für den folgenden Beweis noch ein

Lemma 7.3.4. Es seine R ein Ring, M_i, $i \in I \subset \mathbb{N}$, R-Moduln und $U_i \subset M_i$ Untermoduln. Dann gilt:

$$\left(\bigoplus_{i \in I} M_i \right) / \left(\bigoplus_{i \in I} U_i \right) \cong \bigoplus_{i \in I} M_i / U_i.$$

Beweis. Wir definieren den kanonischen Epimorphismus

$$\pi : \bigoplus_{i \in I} M_i \to \left(\bigoplus_{i \in I} M_i \right) / \left(\bigoplus_{i \in I} U_i \right), \quad (v_i)_{i \in I} \mapsto (v_i + U_i)_{i \in I}.$$

Da $\ker(\pi) = \bigoplus_{i \in I} U_i$ gilt, folgt die Behauptung mit dem Homomorphiesatz 7.1.2. □

Beweis von Satz 7.3.3. Es seien $x_1, \dots, x_n \in M$ Erzeugende für M. Wir definieren den surjektiven Homomorphismus

$$\pi \colon R^n \to M, \quad (r_1, \dots, r_n) \mapsto r_1 v_i.$$

Mit dem Homomorphiesatz 7.1.2 wissen wir, dass $M \cong R^n / \ker(\pi)$ gilt. Da $\ker(\pi)$ ein Untermodul von R^n ist, liefert der Elementarteilersatz 7.2.1 Elemente $\nu_1, \dots, \nu_s \in R \setminus \{0\}$ mit $\nu_i | \nu_{i+1}$, sodass für eine Basis $u_1, \dots, u_n$ für R^n gilt:

$$\ker(\pi) = \langle \nu_1 \cdot u_1 \rangle \oplus \dots \oplus \langle \nu_s \cdot u_s \rangle.$$

Seien $\nu_1, \dots, \nu_k \in R^\times$ die Einheiten unter den ν_i. Wir erhalten mit $\langle \nu_i \cdot u_i \rangle = \langle \nu_i \rangle$ und $\langle u_i \rangle = R$ dann

$$\begin{aligned} M &\cong R^n / \ker(\pi) \\ &\cong \left(\langle u_1 \rangle \oplus \cdots \oplus \langle u_n \rangle \right) / \left(\langle \nu_1 \cdot u_1 \rangle \oplus \cdots \oplus \langle \nu_s \cdot u_s \rangle \oplus \langle u_{s+1} \rangle \oplus \cdots \oplus \langle u_n \rangle \right) \\ &\cong \bigoplus_{i=1}^{s} \langle u_i \rangle / \langle \nu_i \cdot u_i \rangle \oplus \bigoplus_{i=s+1}^{n} \langle u_i \rangle / \langle u_i \rangle \\ &\cong \underbrace{\bigoplus_{i=1}^{k} R/\langle \nu_i \rangle}_{=\bar{0} \oplus \cdots \oplus \bar{0}} \oplus \bigoplus_{i=k+1}^{s} R/\langle \nu_i \rangle \oplus R^{n-s} \\ &\cong R^{n-s} \oplus \bigoplus_{i=k+1}^{s} \langle \nu_i \rangle. \end{aligned}$$

Man beachte, dass wir von der zweiten auf die dritte Isomorphie Lemma 7.3.4 in Verwendung brachten.
Wir haben somit die Behauptung belegt: der erste Summand ist offensichtlich ein freier R-Modul und der zweite der Torsionsmodul.
Die Eindeutigkeitsaussage der Elemente $\nu_{k+1}, \dots, \nu_s \in R$ ergibt sich aus Lemma 7.3.1. □

Auf den ersten Blick wurde in diesem Beweis, der auf [Hau13, S. 79] zurückgeht, lediglich die Isomorphie zwischen den Räumen M und $F \oplus T$ bewiesen. Da R^d und $\bigoplus_{i=1}^{s} R/\langle \nu_i \rangle$ disjunkt sind gibt es für jedes Element $x \in R^d \oplus \bigoplus_{i=1}^{s} R/\langle \nu_i \rangle$ eine eindeutige Zerlegung $x = x_1 + x_2$ mit $x_1 \in R^d$ und $x_2 \in \bigoplus_{i=1}^{s} R/\langle \nu_i \rangle$, diese Eindeutigkeit gilt dann auch für $\varphi(x) = \varphi(x_1) + \varphi(x_2) \in F \oplus T$, wobei $\varphi : R^d \oplus \bigoplus_{i=1}^{s} R/\langle \nu_i \rangle \to M$ den entsprechenden Isomorphismus bezeichnet. Mit $F = \varphi(R^d)$ und $T = \varphi(\bigoplus_{i=1}^{s} R/\langle \nu_i \rangle)$ „erwirken“ wir dann die gewünschte Identität.

Für uns von weit tragender Bedeutung und den Höhepunkt dieses Abschnittes darstellend ist der folgende

Satz 7.3.5. Es sei M ein endlich erzeugter Modul über einem Hauptidealring R und $T \subset M$ der Torsionsmodul. Weiter sei $P \subset R$ ein Vertretersystem der Primelemente von R, und für $p \in P$ bezeichne

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

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 $F \subset M$ 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 $1 \leq n(p,1) \leq \cdots \leq n(p,s_p)$ mit

$$M_p \cong \bigoplus_{j_p=1,\ldots,s_p} R/\langle p^{n(p,j_p)} \rangle,$$

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

$$M \cong F \oplus \bigoplus_{p \in P} \bigoplus_{j_p=1,\ldots,s_p} R/\langle p^{n(p,j_p)} \rangle$$

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

Beweis. Es kann bereits viel auf vorherige Erkenntnisse „abgewälzt“ werden: Das Wissen um die Zerlegung $M = F \oplus T$ haben wir bereits aus Satz 7.3.3 (und der Nachbemerkung), dabei ist

$$F \cong R^d \quad \text{und} \quad T \cong \bigoplus_{i=1}^{s} R/\langle \nu_i \rangle.$$

Die ν_i sind (für $i = 1, \ldots, s$) in Primelemente zerlegbar, also etwa

$$\nu_i = \epsilon_i \cdot p_1^{n(1,j)} \cdot \ldots \cdot p_r^{n(r,j)},$$

mit $p_1^{n(1,j)}, \ldots, p_r^{n(r,j)} \in R$, $\epsilon_i \in R^\times$ und $n(j,i) \in \mathbb{N}$. Auf Grund des *Chinesischen Restsatzes* (7.3.2) gilt dann

$$T \cong \bigoplus_{i=1}^{s} \bigoplus_{j=1}^{r} R/\langle p_j^{n(j,i)} \rangle \cong \bigoplus_{j=1}^{r} \bigoplus_{i=1}^{s} R/\langle p_j^{n(j,i)} \rangle.$$

Wir behaupten jetzt, dass $\bigoplus_{i=1}^{s} R/\langle p_j^{n(j,i)} \rangle$ gerade den Untermoduln $M_{p_j} \subset M$ der p_j-Torsion entspricht.
Zunächst eine erforderliche Zwischenbemerkung: Ein p-Torsionselement[2] $x \in M$ kann keine Einheit sein, denn sonst könnte man aus $p^n x = 0$ für ein $n \in \mathbb{N}$ die Beziehung $p^n x x^{-1} = p^n \cdot 1_M = 0$ folgern, also ein Widerspruch (R ist ein Integritätsbereich).

[2] So nennt man die Elemente eines Torsionsmoduls.

Die Restklasse von p_j ist in jedem Restklassenring der Form $R/\langle p_{j'}^n \rangle$ mit $j \neq j'$ eine Einheit (die Primelemente sind teilerfremd). Das bedeutet: einem p_k-Torsionselement $x \in M$ kann nur ein Element aus $\bigoplus_{i=1}^{s} R/\langle p_k^{n(k,i)} \rangle$ entsprechen. Damit folgt dann auch insgesamt

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

Die Eindeutigkeit aus Lemma 7.3.1 lässt sich auf diesen Satz übertragen. □

Wir werden auf diesen Satz im folgenden Abschnitt gleich wieder zurück kommen und ihn mit Überlegungen zu Vektorräumen (über einem Körper $\mathbb{K}$) in Bezug bringen.

7.4. Vektorräume als Moduln über dem Hauptidealring der Polynome

Der vielleicht etwas furchteinflößende Titel verliert rasch seinen Schrecken, wenn wir uns überlegen, dass wir hier nun das bisher Erarbeitete zusammentragen können, und somit relativ entspannt die Ernte der gesamten Arbeit einfahren.

Wir beginnen mit einer bedeutenden Interpretation; diese ermöglicht es uns, die für Moduln formulierten Sätze auf einen (endlich-dimensionalen) $\mathbb{K}$-Vektorraum anzuwenden.

Wir betrachten einen $\mathbb{K}$-Vektorraum V, sowie eine lineare Abbildung $\alpha\colon V \to V$; nun ersetzen wir die übliche skalare Multiplikation $\mathbb{K} \times V \to V, (k, v) \mapsto k \cdot v$ durch folgende:

$$\mathbb{K}[t] \times V \to V, (p, v) \mapsto p \cdot v := p(\alpha)(v),$$

wobei unter $p(\alpha)(v)$ die in Abschnitt 2.5 beschriebene Nomenklatur zu verstehen ist.
Ausgestattet mit dieser Multiplikation wird der Vektorraum V zu einem $\mathbb{K}[t]$-Modul. Dies erlaubt es uns den *Struktursatz* und dessen Ableitungen auch für Vektorräume gültig zu wissen.

Folgender Satz setzt sich mit der Interpretation eines Vektorraumes V als $\mathbb{K}[t]$-Modul näher auseinander und stellt hierbei auch einen Schlusssatz dar.

Satz 7.4.1. Es sei $\alpha : V \to V$ ein Endomorphismus eines endlichdimensionalen $\mathbb{K}$-Vektorraumes V. Man betrachte V als $\mathbb{K}[t]$-Modul bezüglich α. Dann ist äquivalent:

(i) V ist α-zyklisch.

(ii) V ist als $\mathbb{K}[t]$-Modul von einem einzigen Element $a \in V$ erzeugt.

(iii) Es gibt ein normiertes Polynom $p \in \mathbb{K}[t]$, sodass V als $\mathbb{K}[t]$-Modul isomorph zu $\mathbb{K}[t]/\langle p \rangle$ ist.

(iv) Das Polynom p aus (iii) stimmt mit dem Minimalpolynom m_α sowie dem charakteristischen Polynom p_α von α überein.

Ein Beleg für diesen Satz ist in [Bos14, S. 225] zu finden. Wir werden diesen Satz nicht beweisen; (i) $\Leftrightarrow$ (ii) müsste relativ klar sein. Die nachfolgenden Absätze fassen - auch als Art Beweisskizze für die restlichen Punkte dieses Satzes - die bisherigen Resultate zusammen und bringen diese in Zusammenhang mit einem Vektorraum V (über einem Körper $\mathbb{K}$) aufgefasst als ein $\mathbb{K}[t]$-Modul.

Mit Teil (iii) aus Satz 7.4.1 wissen wir um die Isomorphie von V zu $\mathbb{K}[t]/\langle p \rangle$. Demzufolge induziert Satz 7.3.5 die Isomorphie von V zu einem Modul der Form $\bigoplus_{i=1}^{n} \bigoplus_{j=1}^{s} \mathbb{K}[t]/\langle p_i^{r(i,j)} \rangle$ bzw. $V = \bigoplus_{q \in P} M_q$.

Für einen Endomorphismus $\alpha : V \to V$ sind die Elemente $x \in \mathrm{Hau}(\alpha; \lambda_i) = \ker(\alpha - \lambda_i \,\mathrm{id})^{r_i}$ genau die p_i-Torsionselemente von M_{p_i}, wobei $p_i := (\alpha - \lambda_i \,\mathrm{id}) \in P$, denn die irreduziblen Elemente (also die Primelemente) in $\mathbb{K}[t]$ sind genau von der Form $(t + a)$ und es gilt

$$p_i^{r_i} \cdot x = (t - \lambda_i)^{r_i} x = (\alpha - \lambda_i \cdot \mathrm{id})^{r_i}(x) = 0.$$

Jeder Vektor $x \in \mathrm{Hau}(\alpha; \lambda_i)$ ist also ein Element $x \in M_{p_i} = \bigcup_{n \in \mathbb{N}} \ker(p_i^n(\alpha))$, was bedeutet, dass die Zerlegung in die Haupträume (aus den vorigen Zugängen) der Zerlegung $\bigoplus_{p \in P} M_p$ entspricht.
Man kann hiermit auch plausibel machen, dass es sich bei p_i um das charakteristische Polynom von $\alpha|_{\mathrm{Hau}(\alpha;\lambda_i)}$, also einen Faktor des charakteristischen Polynoms p_α handelt.[3]

In Kapitel 4 bzw. explizit in Kapitel 5 wurde die Zerlegung eines Hauptraumes $\mathrm{Hau}(\alpha; \lambda_i)$ in Erzeugnisse der Form $U_{ij} = \langle u_{ij}, \alpha(u_{ij}), \ldots, \alpha^{k-1}(u_{ij}) \rangle$ für einen gewählten Vektor $u_{ij} \in \mathrm{Hau}(\alpha; \lambda_i)$ mit maximaler Periode demonstriert. In Abschnitt 4.3 sind wir hierfür, zwecks entsprechender Anordnung der konstruierten

[3] Wir wollen das hier nicht beweisen, ein Beleg dafür findet sich in [Bos14, Seite 225].

Basiselemente jeweils von unten nach oben „gelaufen“. Es stellt somit jede Reihe einen zyklischen α-invarianten Unterraum von $\mathrm{Hau}(\alpha; \lambda_i)$ dar. Die Dimension eines dieser Unterräume U_{ij} wurde durch den erzeugenden Vektor u_{ij} und der Beziehung $\alpha^k(u_{ij}) = 0$, also auch durch den „Exponenten“ k der nilpotenten Abbildung φ festgelegt.
Das Pendant hierfür in Zugang Drei ist die Zerlegung des Moduls M_{p_i} in Untermoduln mit folgender Eigenschaft: jeder Untermodul, etwa U_{ij}, ist seiner Dimension entsprechend isomorph zu einem Restklassenring $R/\langle p_i^{r(i,j)} \rangle$; nach obigem Satz ist U_{ij} monogen, d.h. von einem einzigen Element erzeugt. Es gilt die Identität $\dim(U_{ij}) = r(i,j)$, d.h. der Exponent $r(i,j)$ des Quotienten $R/\langle p_i^{r(i,j)} \rangle$ bestimmt die Dimension des korrespondierenden Untermoduls U_{ij}.[4] Man muss sich jedoch damit begnügen, dass der Untermodul U_{ij} insofern nicht eindeutig ist, da dieser im Allgemeinen von der Wahl des erzeugenden Vektors u_{ij} abhängt.

Somit sind wir am Ende unserer Bemühungen angelangt. Wir wissen um eine entsprechende Zerlegung von V, und mit der Referenz auf Kapitel 4 oder 5 könnten wir „den Sack zumachen“.

Es bietet sich hier jedoch die Möglichkeit einen alternativen Weg für die letzten Schritte zur Herleitung der *Jordanschen Normalform* einzuschlagen.
Wir wollen uns an dieser Stelle nochmal von der Einschränkung, dass die Polynome p_i wie in obiger Zusammenfassung von der Form $(t - \lambda)$ sind, lösen und werden auf Grund der - im Vergleich zu den vorigen Ansätzen - universelleren „Natur“ der bisher formulierten Sätze imstande sein die *Jordansche Normalform* als Spezialfall der sogenannten *Allgemeinen Normalform* herzuleiten.

Dazu benötigen wir noch die folgende

Definition. Sei $q = t^d + a_{d-1} \cdot t^{d-1} + \cdots + a_0 \in \mathbb{K}[t]$ ein normiertes Polynom vom Grad $d \geq 1$. Dann bezeichnet

$$B(q) := \begin{pmatrix} 0 & & & & & -a_0 \\ 1 & 0 & & & & -a_1 \\ & 1 & \ddots & & & \vdots \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & 0 & \vdots \\ & & & & 1 & -a_{d-1} \end{pmatrix}$$

[4] Vergleiche hierzu [Bos14, Abschnitt 6.5, v.a. Theorem 6] unter Berücksichtigung $\deg(p_i) = 1$.

die **Begleitmatrix** zu q. Insbesondere gilt $B(q) = (-a_0)$ für $d = 1$.

Die Strukturs\"atze aus dem vorigen Abschnitt erlauben nun die Formulierung von

Satz 7.4.2. Jede Matrix $A \in M_n(\mathbb{K})$ ist ähnlich zu einer Matrix in Blockdiagonalform, wobei die Blöcke der Hauptdiagonale A_i jeweils der Begleitmatrix zu einer Potenz $q_i = p_i^{e_i}$ eines normierten Primpolynoms aus $\mathbb{K}[t]$ ist. Die Blöcke $A_1, \ldots, A_r$ sind bis auf die Reihenfolge eindeutig durch A bestimmt. Es gilt $m_A = \mathrm{kgV}(q_1, \ldots, q_r)$ für das Minimalpolynom von A, sowie $p_A = q_1 \cdot \ldots \cdot q_r$ für das charakteristische Polynom von A (bzw. der zugehörigen linearen Abbildung $\alpha\colon V \to V$, wobei V ein endlichdimensionaler Vektorraum über $\mathbb{K}$ ist).
Illustriert bedeutet dies: für den Vektorraum V existiert eine Basis $\mathcal{B}$, sodass

$$M_{\mathcal{B}}(\alpha) = \begin{pmatrix} \boxed{A_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{A_r} \end{pmatrix} \text{ gilt,}$$

wobei für $i = 1, \ldots, r$ die sogenannten *verallgemeinerten Jordan-Kästchen* durch

$$A_i = \begin{pmatrix} \boxed{B(p_i)} & & & 0 \\ 1 & \boxed{B(p_i)} & & \\ & & \ddots & \\ 0 & & 1 & \boxed{B(p_i)} \end{pmatrix},$$

mit e_i-mal $B(p_i)$ auf der Hauptdiagonale und Eintrag 1 jeweils auf der unteren Nebendiagonale (in der Ecke an denen sich zwei Blöcke berühren), gegeben sind.

Beweis. Einen Beleg liefert [Bos14, S. 230] und die Tatsache, dass es die Faktorisierung $q_i = p_i^{e_i}$ in Primpolynome mit passenden Zahlen $e_i \in \mathbb{N}$ erlaubt, die Blöcke A_i entsprechend weiter zu strukturieren.
Im - zugegeben - sehr Groben haben wir diesen Satz aber schon belegt. □

Dieser Satz gilt also auch für Endomorphismen $\alpha\colon V \to V$ eines Vektorraumes V über einem nicht notwendig algebraisch abgeschlossenen Körper $\mathbb{K}$.
Die aus 7.4.2 resultierende Matrixdarstellung bezeichnet man als *Allgemeine Normalform*, weshalb sie sich als die bestmögliche Matrixdarstellung eines Endomorphismus α, dessen Minimalpolynom m_α (bzw. charakteristisches Polynom p_α) nicht in Linearfaktoren zerfällt, ausweist.

Letzten Endes können wir erneut den dritten Ansatz - elegant - vollenden. Falls das charakteristische Polynom in Linearfaktoren der Form $(t-\lambda)$ zerfällt, offenbart sich uns ein mittlerweile sehr wohl bekanntes Bild: Die **Jordansche Normalform**.

Literaturverzeichnis

[Art98] Michael Artin. *Algebra*. Deutsche Erstausgabe. Basel: Birkhäuser, 1998.

[Bos14] Siegfried Bosch. *Lineare Algebra*. 5., überarbeitete und erweiterte Auflage. Berlin: Springer Spektrum, 2014.

[Bri85] Egbert Brieskorn. *Lineare Algebra und Analytische Geometrie II*. Braunschweig: Vieweg, 1985.

[Brö03] Theodor Bröcker. *Lineare Algebra und Analytische Geometrie. Ein Lehrbuch für Physiker und Mathematiker*. Basel: Birkhäuser, 2003.

[Cap10] Andreas Čap. *Einführung in die Lineare Algebra und Geometrie*. Vorlesungsskriptum. Universität Wien: Fakultät für Mathematik, Sommersemester 2010.

[Cap11] Andreas Čap. *Lineare Algebra und Geometrie für LehramtskandidatInnen (Kapitel 6-9)*. Vorlesungsskriptum. Universität Wien: Fakultät für Mathematik, Wintersemester 2010/11.

[Fis10] Gerd Fischer. *Lineare Algebra. Eine Einführung für Studienanfänger*. 17. Auflage. Wiesbaden: Vieweg+Teubner, 2010.

[Fis13] Gerd Fischer. *Lehrbuch der Algebra*. 3., erweiterte Auflage. Wiesbaden: Springer Spektrum, 2013.

[Hau13] Jürgen Hausen. *Lineare Algebra II*. 1. Auflage. Aachen: Shaker, 2013.

[Jän08] Klaus Jänich. *Lineare Algebra*. 11. Auflage. Berlin: Springer, 2008.

[KM09] Christian Karpfinger und Kurt Meyberg. *Algebra. Gruppen - Ringe - Körper*. Heidelberg: Spektrum, 2009.

[Mut06] Herbert J. Muthsam. *Lineare Algebra und ihre Anwendungen*. 1. Auflage. München: Spektrum, 2006.

Anhang

Zusammenfassung

Der Gegenstand dieser Arbeit ist die Präsentation dreier an sich unabhängiger Darstellungen zur Entwicklung bzw. zum Existenzbeleg einer sogenannten *Jordan-Basis*.

In der ersten Herangehensweise wird elementar und auch am ausführlichsten gezeigt, dass die *Jordansche Normalform* die in gewissem Sinne bestmögliche Matrixdarstellung von Endomorphismen ist, deren charakteristische Polynome in Linearfaktoren zerfallen.

Im zweiten Ansatz erfolgt die „grobe“ Hauptraumzerlegung des betrachteten Vektorraums über das Minimalpolynom, die weitere Zerlegung der Haupträume resultiert - wie im ersten Zugang - aus den Betrachtungen des nilpotenten Anteils des entsprechenden Endomorphismus.

Im dritten Zugang zeigt sich wie mit geeigneter Interpretation algebraische Betrachtungen fruchtbar in die Lineare Algebra und in unser Themengebiet überführt werden können und die oben beschriebene Vektorraumzerlegung herzuleiten vermögen.

Abstract

Topic of my diploma thesis is the presentation of three independent approaches to the proof of a *Jordan-Base*. These three approaches are structured in increasing abstraction-amount.

The first approach will thoroughly show that the *Jordan canonical form* is - in a certain way - the best representation of a linear map which *characteristic polynomial* decompose in linear factors.

The decomposition of the pertaining vector space in the second approach results from considerations about the *minimal polynomial* and - analogically to the first approach - the nilpotent part of the regarded endomorphism.

In the course of the last approach the isomorphic relation between certain *quotient spaces* and *modules* - initiated by the formulization of the *homomorphism theorem* - will be shown. Furthermore yields this isomorphic structure and the interpretation of vector spaces as a special case of modules to the desired invariant decomposition of the regarded vector space.