Lineare Algebra II

Sommersemester 2018

Mohamed Barakat

Department Mathematik, Universität Siegen
mohamed.barakat@uni-siegen.de

Stand: 18. Juli 2018

Vorwort

Dies ist die geTEXte Version meiner Vorlesungsnotizen, die ich fortlaufend aktualisieren werde. Habt bitte Verständnis dafür, wenn Stand der Vorlesung und der Notizen nicht immer übereinstimmen werden. Daher gilt: Kommt zur Vorlesung und macht Eure eigenen Notizen. *Die sind sowieso besser als jedes Skript*. Die Form eines Skriptes erreichen diese Notizen vermutlich erst gegen Ende der Vorlesung, dies kann ich aber nicht garantieren. Die aktuelle Version ist unter der folgenden Adresse zu finden:

https://www.mathb.rwth-aachen.de/~barakat/Lehre/SS18/LAII/Skript/LA_II.pdf

Als Vorlage benutz(t)e ich das online-verfügbare Skript von Prof. Gabriele Nebe, das sie mir freundlicherweise zur Verfügung gestellt hat.

Das Skript zu LA I befindet sich unter

https://www.mathb.rwth-aachen.de/~barakat/Lehre/WS17/LAI/Skript/LA_I.pdf

Für Korrektur- und Verbesserungsvorschläge bin ich stets dankbar
mohamed.barakat@uni-siegen.de

Inhaltsverzeichnis

Kapitel 6

Moduln

6.1 Moduln

Definition 6.1.1. Sei R ein Ring. Eine abelsche Gruppe $(M, +)$ heißt R-**Modul** (genauer R-Linksmodul), falls eine Abbildung

$$R \times M \to M : (r, m) \mapsto rm$$

gegeben ist mit

$$\begin{aligned} r(m+n) &= rm + rn, \\ (rs)m &= r(sm), \\ (r+s)m &= rm + sm \\ 1m &= m \end{aligned}$$

für alle $r, s \in R, m, n \in M$.
Ist M ein R-Modul, so heißt eine Teilmenge $T \subseteq M$ ein **Teilmodul** von M, in Zeichen $T \leq M$, falls $T \neq \emptyset$ und für alle $t_1, t_2 \in T$, $a \in R$ auch $at_1 + t_2 \in T$ gilt.

Man ist versucht zu sagen, dass Moduln Vektorräume über Ringen sind. Richtig ist natürlich, dass Vektorräume Moduln über Körpern sind.

Übung 6.1.1. Zeigen Sie, dass Teilmoduln genau die Teilmengen T von M sind, die mit der Einschränkung der Addition und Skalarmultiplikation von M auf T wieder zu R-Moduln werden.

Beispiel 6.1.2.

(1) Jede abelsche Gruppe $(M, +)$ ist ein $\mathbb{Z}$-Modul mit

$$am := \begin{cases} \underbrace{m + \ldots + m}_{a}, & a \geq 0 \\ -\underbrace{(m + \ldots + m)}_{-a}, & -a \geq 0. \end{cases}$$

(2) Sei $\mathcal{V}$ ein K-Vektorraum für einen Körper K und $\varphi \in \mathrm{End}_K(\mathcal{V})$. Dann wird $\mathcal{V}$ zu einem $K[x]$-Modul durch die Setzung

$$p(x)v := (p(\varphi))(v) \text{ für alle } p(x) \in K[x], v \in \mathcal{V}.$$

Übung 6.1.2. Sei $\psi : R \to S$ ein Ringhomomorphismus und M ein S-Modul. Dann wird M zu einem R-Modul, durch $rm := \psi(r)m$ für $r \in R, m \in M$.

Ist ψ injektiv (also $R \cong \psi(R)$ ein Teilring von S), so nennt man den R-Modul M auch die **Einschränkung** des S-Moduls M. Ist ψ surjektiv (also $S \cong R/\operatorname{Kern}(\psi)$), so nennt man den R-Modul M auch die **Aufblasung** (Inflation) des S-Moduls M.

Bemerkung 6.1.3.

(1) Sei M ein R-Modul und $\mathcal{T}$ eine Menge von Teilmoduln von M. Dann gilt:

$$\bigcap_{T \in \mathcal{T}} T \leq M$$

ist wieder ein Teilmodul von M.

(2) Für $X \subseteq M$ so ist das **Erzeugnis** $\langle X \rangle := \bigcap_{X \subset T \leq M} T$ der kleinste Teilmodul von M, welcher X enthält.

(3) Es gilt für $\emptyset \neq X \subseteq M$

$$\langle X \rangle = \{m \in M | \text{ es existieren } k \in \mathbb{N}, a \in R^k, v \in X^k \text{ mit } m = a_1 v_1 + \cdots + a_k v_k\}$$

und $\langle \emptyset \rangle = \{0\}$.

Beweis. zu 3. Man rechnet leicht nach, dass die Menge $\overline{X}$ auf der rechten Seite ein R-Teilmodul von M ist.[1] Es gilt $X \subseteq \overline{X}$, da der Ring R eine Eins besitzt. Also nach Definition gilt somit $\langle X \rangle \leq \overline{X}$. Aber andererseits ist $\overline{X}$ in jedem Teilmodul von M enthalten, der X enthält. Das liefert die Gleichheit. □

Definition 6.1.4. Eine Abbildung $\varphi : M \to N$ von R-Moduln M, N heißt R-**Modulhomomorphismus**, falls

$$\varphi(rm_1 + sm_2) = r\varphi(m_1) + s\varphi(m_2)$$

für alle $r, s \in R$ und alle $m_1, m_2 \in M$ gilt. In diesem Fall heißt die Faser über 0

$$\operatorname{Kern}(\varphi) := \{m \in M | \varphi(m) = 0\}$$

der **Kern** von φ.

Bemerkung 6.1.5.

(1) Die Komposition von R-Modulhomomorphismen ist ein R-Modulhomomorphismus.

(2) Ist $\varphi : M \to N$ ein R-Modulhomomorphismus, so ist $\operatorname{Kern}(\varphi)$ ein Teilmodul von M und $\operatorname{Bild}(\varphi) := \{\varphi(m) \mid m \in M\}$ ein Teilmodul von N.

(3) φ ist injektiv genau dann wenn $\operatorname{Kern}(\varphi) = \{0\}$ ist.

(4) φ ist surjektiv genau dann wenn $\operatorname{Bild}(\varphi) = N$ ist.

(5) Ist φ bijektiv (also ein **Isomorphismus**), so ist die Umkehrabbildung φ^{-1} wieder ein R-Modulisomorphismus. Insbesondere ist Isomorphie von Moduln eine Äquivalenzrelation.

(6) R-Modulhomomorphismen von M in sich selbst heißen **Endomorphismen**. $\operatorname{End}_R(M) := \{\varphi : M \to M \mid \varphi \ R\text{-Modulhomomomorphismus } \}$ heißt der **Endomorphismenring** des R-Moduls M. Es ist $\operatorname{End}_R(M)$ ein Ring, im Fall dass R kommutativ ist, sogar eine R-Algebra.

[1] Die Konvention, dass die leere Linearkombination gleich 0 ist, führt zu einer Vereinheitlichung der beiden Fälle.

Beweis. Übungsaufgabe. □

Beispiel 6.1.6. Ein wichtiger Struktursatz für Vektorräume war der Steinitzsche Austauschsatz, der uns vielfältige Möglichkeiten eröffnete, Basen zu konstruieren. Dieser ist für allgemeine R-Moduln falsch:
Sei dazu $R = \mathbb{Z}$ und $M = \mathbb{Z}^{3\times 1}$. Dann wird M durch Einschränken der $\mathbb{Q}$-Vektorraumstruktur von $\mathbb{Q}^{3\times 1}$ zu einem $\mathbb{Z}$-Modul. Jedes Element von M ist eindeutige $\mathbb{Z}$-Linearkombination von Elementen aus der Standardbasis $E = (e_1, e_2, e_3)$, in diesem Sinn ist E eine $\mathbb{Z}$-Basis von M. Jedoch ist $(e_1, e_2, 2e_3)$ eine linear unabhängige Teilmenge von M, die sich nicht zu einer $\mathbb{Z}$-Basis ergänzen lässt. Ebensowenig $B = (e_1 + e_2, e_2 + e_3, e_1 + e_3)$, denn die Linearkombinationen von Elementen in B sind genau die

$$\{ \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \in M \mid a_1 + a_2 + a_3 \text{ gerade } \}$$

Übung 6.1.3. Die Spalten der Matrix A in $\mathbb{Z}^{n\times n}$ bilden genau dann ein Erzeugendensystem des $\mathbb{Z}$-Moduls $\mathbb{Z}^{n\times 1}$ wenn

$$A \in \mathrm{GL}_n(\mathbb{Z}) = (\mathbb{Z}^{n\times n})^* = \{g \in \mathbb{Z}^{n\times n} \mid \det(g) \in \{\pm 1\}\}.$$

Die Tatsache, dass R ein Ring ist, erlaubt es uns, beliebige Moduln als Faktormoduln freier Moduln zu beschreiben und so einen ersten Rahmen zu bekommen, wie man Moduln konstruiert.

Bemerkung 6.1.7. Sei R ein Ring.

(1) $M = R$ kann als R-Modul aufgefasst werden durch

$$R \times M \to M : (r, m) \mapsto rm \qquad \text{(Produkt in } R\text{)}.$$

Diesen Modul bezeichnen wir mit ${}_R R$. Er heißt der **reguläre R-Modul**. Seine Teilmoduln nennt man auch **Linksideale**.

(2) Ist M irgendein R-Modul und $m \in M$, dann gibt es genau einen R-Modulhomomorphismus

$$\varphi_m : {}_R R \to M \text{ mit } \varphi_m(1) = m.$$

(3) Sind M und N R-Moduln, so auch die direkte Summe $M \oplus N$ (entspricht dem direkten Produkt bei abelschen Gruppen in additiver Schreibweise) durch die R-Operation

$$r(m, n) := (rm, rn) \text{ für alle } m \in M, n \in N, r \in R.$$

$M \oplus N$ heißt die **direkte Summe** der R-Moduln M und N.

(4) Ist A eine beliebige Menge, so ist R^A ein R-Modul mit werteweiser Addition und Produkt:

$$R \times R^A \to R^A : (r, f) \mapsto (a \mapsto rf(a)).$$

Im Falle von $A = \underline{n}$ schreiben wir R^n statt $R^{\underline{n}}$.

Beweis.

(1) Klar.

(2) Existenz:

$$\varphi : {}_RR \to M : r \mapsto rm$$

ist wohldefiniert und hat die gewünschte Eigenschaft.
Eindeutigkeit: Sei ψ ein weiterer Homomorphismus mit dieser Eigenschaft. Dann gilt für alle $r \in {}_RR$:

$$\psi(r) = \psi(r1) = r\psi(1) = rm = \varphi_m(r), \text{ also } \psi = \varphi_m.$$

(3) Übung.

(4) Übung. □

Übung 6.1.4. Zeige: Sind A und B disjunkte Mengen, so gilt: $R^A \oplus R^B \cong R^{A \cup B}$ als R-Moduln.

Ende Vorl. 1

Bemerkung 6.1.8. Sei A eine Menge und für jedes $a \in A$ sei $e_a \in R^A$ die charakteristische Funktion von $\{a\}$, definiert durch

$$e_a(b) := \begin{cases} 0, & b \neq a \\ 1, & b = a \end{cases}$$

Dann ist der von den e_a mit $a \in A$ erzeugte R-Teilmodul von R^A gegeben durch

$$\mathrm{Fr}_R(A) := \langle e_a | a \in A \rangle_R = \{f \in R^A \mid |\{a \in A | f(a) \neq 0\}| < \infty\} \leq R^A.$$

$\mathrm{Fr}_R(A)$ heißt der **freie R-Modul auf** A. Ist $|A| < \infty$ so ist offensichtlich $\mathrm{Fr}_R(A) = R^A$.

Satz 6.1.9. *Sei R ein Ring und A eine Menge. Der Modul $\mathrm{Fr}_R(A) := \langle e_a | a \in A \rangle_R \leq_R R^A$ hat folgende Eigenschaft: Für jeden R-Modul M und jede Abbildung $\psi : A \to M$ gibt es genau einen R-Modulhomomorphismus*

$$\begin{array}{rcll} \widetilde{\psi} : \mathrm{Fr}_R(A) & \to & M & \textit{mit} \\ \widetilde{\psi}(e_a) & = & \psi(a) & \textit{für alle } a \in A. \end{array}$$

Moduln, die isomorph zu $\mathrm{Fr}_R(A)$ sind, heißen **frei** *auf dem Erzeugendensystem, welches $(e_a)_{a \in A}$ vermöge $\widetilde{\psi}$ entspricht. Ein freies Erzeugendensysteme heißt auch* **Basis**, *genauer R-Modulbasis.*

Beweis. Jedes Element aus $\mathrm{Fr}_R(A)$ hat eine eindeutige Darstellung als

$$\sum_{a \in A} r_a e_a$$

mit $r_a \in R$ und $r_a = 0$ für alle bis auf endlich viele $a \in A$. Daher ist

$$\widetilde{\psi} : \mathrm{Fr}_R(A) \to M : \sum_{a \in A} r_a e_a \mapsto \sum_{a \in A} r_a \psi(a)$$

eine wohldefinierte Abbildung, von der man leicht zeigt, dass sie ein Modulhomomorphismus ist. Sie erfüllt sicher die Bedingung $\widetilde{\psi}(e_a) = \psi(a)$ für alle $a \in A$ und ist somit auch der einzige Modulhomomorphismus mit dieser Eigenschaft. □

Beispiel 6.1.10.

(1) ${}_RR$ ist frei auf $\{1\}$.

(2) Der Spaltenmodul $R^{n \times 1}$ ist frei auf den Einheitsspalten $(e_1, \ldots, e_n)$.

Übung 6.1.5. Sei R ein kommutativer Ring. Dann gilt

$$\mathrm{End}_R(R^n) \cong R^{n\times n},$$

wobei wir $R^{n\times n}$ durch komponentenweise Addition und übliche Multiplikation zu einem Ring machen. Genauer: Identifiziere R^n mit $R^{n\times 1}$. Dann liefert das Heranmultiplizieren von Matrizen aus $R^{n\times n}$ eine eindeutige Darstellung der Endomorphismen von $R^{n\times 1}$ durch Matrizen. Man setzt
$\mathrm{GL}_n(R) := (R^{n\times n})^*$. (Beachte, der Fall $n = 1$ ist schon interessant.)
(Hinweis:

$$\widetilde{A} : R^{n\times 1} \to R^{n\times 1} : X \mapsto AX$$

ist für jedes $A \in R^{n\times n}$ ein R-Modulendomorphismus und jede Matrix induziert einen anderen Endomorphismus, da ein Endomorphismus durch die Bilder der (freien) Erzeuger $e_1, \ldots, e_n$ festgelegt ist, die wie in der linearen Algebra in den Spalten der beschreibenden Matrix stehen. Da diese Bilder beliebig vorgegeben werden können, folgt die Behauptung, wenn man beachtet, dass der Summe und dem Matrixprodukt gerade die Summe und die Hintereinanderausführung der Endomorphismen entsprechen.)

6.2 Homomorphiesätze und der chinesische Restsatz.

6.2.a Der Homomorphiesatz für Moduln

Bemerkung 6.2.1. Sei R Ring und M ein R-Modul mit Teilmodul $U \leq M$.

(1) Für $m \in M$ heißt

$$m + U := \{m + u | u \in U\}$$

die **Restklasse** von m nach U. Die Menge

$$M/U := \{m + U | m \in M\}$$

aller Restklassen nach U in M bilden den **Faktormodul** M/U von M nach U vermöge der folgenden Verknüpfungen:

$$+ : M/U \times M/U \to M/U : (m_1 + U, m_2 + U) \mapsto (m_1 + m_2) + U$$

und

$$\cdot : R \times M/U \to M/U : (r, m + U) \mapsto rm + U.$$

(2) Der **natürliche Epimorphismus**

$$\nu = \nu_U : M \to M/U : m \mapsto m + U$$

ist ein R-Modulepimorphismus mit $\mathrm{Kern}(\nu_U) = U$.

Beweis.

(1) Wir müssen zeigen, dass $+$ und $\cdot$ wohldefiniert sind. Die Addition lassen wir als Übung. Für das Produkt sei $m + U = n + U$. Wir zeigen: $rm + U = rn + U$. Mit $m - n \in U$ ist auch $r(m - n)$, also $rm + U = rn + U$. Die R-Modulaxiome müssen verifiziert werden. Z.B. ist $U = 0 + U$ das Nullelement von M/U. Den Rest lassen wir als Übung.

(2) Dies folgt direkt aus der Definition des Produktes und der Addition von Restklassen.

□

Der nächste Schritt in der allgemeinen Modultheorie ist der **Homomorphiesatz**, dessen Beweis genau so einfach ist wie bei Vektorräumen.

Satz 6.2.2. *Sei $\varphi : M \to N$ ein R-Modulhomomorphismus. Dann faktorisiert φ als*

$$\varphi = \widetilde{\varphi} \circ \nu_{\mathrm{Kern}(\varphi)}$$

über $M/\operatorname{Kern}(\varphi)$ mit $\nu_{\mathrm{Kern}(\varphi)}$ ein R-Modulepimorphismus und den R-Modulmonomorphimus

$$\widetilde{\varphi} : M/\operatorname{Kern}(\varphi) \to N : m + \operatorname{Kern}(\varphi) \mapsto \varphi(m).$$

Bemerkung 6.2.3. Ist M ein endlich erzeugter R-Modul, so gibt es ein $n \in \mathbb{N}$ und einen R-Modulepimorphismus $\varepsilon : R^{n\times 1} \to M$. Insbesondere $M \cong R^{n\times 1}/\operatorname{Kern}(\varepsilon)$.

Beweis. Sei $\psi : \underline{n} \to M$ gegeben, so dass $\operatorname{Bild}(\psi)$ ein Erzeugendensystem von M ist. Der nach Satz 6.1.9 eindeutige R-Modulhomomorphismus $\varepsilon : R^{n\times 1} \to M$ mit $\varepsilon \circ S = \psi$, wo S die Standardbasis von $R^{n\times 1}$ ist, ist dann ein Epimorphismus. □

Definition 6.2.4. R-Moduln, die von einem Element erzeugt werden heißen **zyklisch**.

Beispiel 6.2.5.

(1) Jede zyklische Abelsche Gruppe ist von der Form $\mathbb{Z}/K$, wobei K ein Linksideal von $\mathbb{Z}$ ist (siehe später).

(2) Jede Abelsche Gruppe, die von n Elementen erzeugt wird, ist von der Form $\mathbb{Z}^n/K$ wobei K ein $\mathbb{Z}$-Teilmodul von $\mathbb{Z}^n = \mathrm{Fr}_{\mathbb{Z}}(\underline{n})$ ist.

(3) Sei K ein Körper und $\mathcal{V}$ ein endlich erzeugter K-Vektorraum mit $\varphi \in \mathrm{End}_K(\mathcal{V})$, so dass das Minimalpolynom und das charakteristische Polynom von φ beide gleich $p(x) \in K[x]$ sind, so ist $\mathcal{V}$ ein zyklischer $K[x]$-Modul und es gilt (vgl. LA I, Begleitmatrix von $p(x)$):
$$\mathcal{V} \cong_{K[x]\ K[x]} K[x]/\langle p(x)\rangle.$$

6.2.b Ideale

Wir kommen zu dem Homomorphiesatz von Ringen. Zuerst sieht die Definition eines Ideals etwas sonderbar aus, wird aber einsichtig, wenn man sich vorstellt, dass ein Ideal etwas ist, was man gleich Null setzen kann, um einen neuen Ring zu bekommen.

Definition 6.2.6. Sei R ein Ring.

(1) $I \subseteq R$ heißt **(zweiseitiges) Ideal** von R, in Zeichen $I \trianglelefteq R$, falls

- $I \neq \emptyset$ und
- $a, b \in I$ und $r, s \in R$ impliziert $ra + bs \in I$.

(2) Sind $I_1, I_2 \trianglelefteq R$ so heißt das kleinste Ideal $I_1 + I_2$, welches I_1 und I_2 enthält, die **Summe** von I_1 und I_2.

Beispiel 6.2.7.

(1) Für $R = \mathbb{Z}$ ist $3\mathbb{Z} = \langle 3\rangle = \{3z \mid z \in \mathbb{Z}\}$ ein Ideal: $\langle 3\rangle \lhd \mathbb{Z}$.

(2) Ist K ein Körper und $a \in K$, dann ist

$$\{p(x) \in K[x] \mid p(a) = 0\} = \langle x - a\rangle := \{p(x)(x - a) \mid p(x) \in K[x]\} \lhd K[x].$$

(3) Ist R ein kommutativer Ring mit Eins, so sind die Ideale in R genau die R-Teilmoduln von ${}_RR$, sprich die Linksideale.

(4) Der Durchschnitt einer Menge von Idealen ist wieder ein Ideal.

(5) Ist $M \subseteq R$, so heißt
$$\langle M\rangle := \bigcap_{M\subseteq I\trianglelefteq R} I$$
das von M **erzeugte Ideal**. Ist $M = \{a_1, \dots a_n\}$ so schreibt man auch $\langle a_1, \dots, a_n\rangle$ statt $\langle M\rangle$.

(6) Ist R ein kommutativer Ring mit Eins, so heißt
$$\langle a\rangle := \{ra \mid r \in R\}$$
das von $a \in R$ erzeugte **Hauptideal**: $\langle a\rangle \trianglelefteq R$.

(7) Ist $\varphi : R \to S$ ein Ringhomomorphismus, dann ist
$$\mathrm{Kern}(\varphi) := \{r \in R \mid \varphi(r) = 0\}$$
ein Ideal von R:
$$\mathrm{Kern}(\varphi) \trianglelefteq R.$$

Übung 6.2.1. Das von M erzeugte Ideal ist
$$\langle M\rangle = \left\{ \sum_{i=1}^{n} a_i m_i b_i \mid n \in \mathbb{N}_0, a_i, b_i \in R, m_i \in M \right\}.$$
Benutze diese Beschreibung um zu zeigen, dass die Ideale von $R = \mathbb{Z}^{n\times n}$ genau die Teilmengen aR sind mit $a \in \mathbb{Z}$.

Bei Ringen können wir Restklassenringe nach Idealen bilden. Der neue Punkt ist die Wohldefiniertheit der vertreterweisen Multiplikation.

Satz 6.2.8. *Sei R ein Ring und $I \trianglelefteq R$ ein Ideal von R. Dann ist $R/I := \{r + I | r \in R\}$ ein Ring mit den vertreterweisen Verknüpfungen*
$$\begin{aligned} + : R/I \times R/I \to R/I &: (r + I, s + I) \mapsto (r + s) + I, \\ \cdot : R/I \times R/I \to R/I &: (r + I, s + I) \mapsto rs + I, \end{aligned}$$
und $\nu = \nu_I : R \to R/I : r \mapsto r + I$ ist ein Ringepimorphismus mit I als Kern. R/I heißt **Restklassenring** *von R nach I und ν der* **natürliche Epimorphismus**. *(Insbesondere ist jedes Ideal Kern eines Ringepimorphismus.) Ist R kommutativ, so auch R/I.*

Beweis. Da $I \leq R$ ein R-Teilmodul von R ist, ist R/I wieder ein R-Modul und wir brauchen uns nur um die Wohldefiniertheit der Multiplikation zu kümmern. Seien also $r + I = r' + I$ und $s + I = s' + I$ für gewisse $r, r', s, s' \in R$. Dann existieren $a, b \in I$ mit $r' = r + a, s' = s + b$, und wir bekommen
$$\begin{aligned} r's' - rs &= (r + a)(s + b) - rs \\ &= rs + rb + as + ab - rs \\ &= rb + as + ab \in I \end{aligned}$$
d.h. $(r + I)(s + I)$ ist wohldefiniert. Die Assoziativ- und Distributivgesetze übertragen sich von R. Dass ν ein Epimorphismus ist, ist gerade die Definition der Operationen im Restklassenring. □

Folgerung 6.2.9. *(Homomorphiesatz für Ringe) Seien R, S Ringe und $\varphi : R \to S$ ein Ringhomomorphismus. Dann ist $I := \mathrm{Kern}(\varphi)$ ein Ideal von R, $\mathrm{Bild}(\varphi)$ ein Teilring von S und*

$$\overline{\varphi} : R/I \to \mathrm{Bild}(\varphi), r + I \mapsto \varphi(r)$$

ein wohldefinierter Ringisomorphismus.

Bemerkung 6.2.10. Sei M ein R-Modul. Dann ist der **Annihilator** von M

$$\mathrm{Ann}_R(M) := \{r \in R \mid rm = 0 \text{ für alle } m \in M\}$$

ein Ideal von R, der Kern des Ringhomomorphismus

$$R \to \mathrm{End}_{\mathbb{Z}}(M), r \mapsto (m \mapsto rm).$$

Weiter ist M ein $R/\mathrm{Ann}_R(M)$-Modul.

Ende Vorl. 2

Bemerkung 6.2.11. Seien R und S Ringe, M ein R-Modul und N ein S-Modul. Dann ist $M \oplus N$ ein $R \times S$-Modul durch

$$(r, s) \cdot (m, n) := (rm, sn) \text{ für alle } r \in R, s \in S, m \in M, n \in N.$$

Sei umgekehrt L ein $R \times S$-Modul. Dann sind

$$M := (R \times \{0\})L = (1, 0)L,$$
$$N := (\{0\} \times S)L = (0, 1)L$$

Teilmoduln von L, so dass $L \cong M \oplus N$. Es ist $\mathrm{Ann}_{R\times S}(M) \supseteq \{0\} \times S$ und somit M ein $R \times S/(\{0\} \times S) \cong R$-Modul und ebenso $(R \times \{0\})N = \{0\}$ und N ist ein S-Modul.
Die $R \times S$-Moduln sind also genau die direkten Summen von R-Moduln und S-Moduln.

Definition 6.2.12. Sei R ein kommutativer Ring mit Eins. Eine R-**Algebra** A ist ein Ring[2] mit Eins, der gleichzeitig R-Modul ist, so dass die Multiplikation R-bilinear ist, d.h.

$$(ra)b = a(rb) = r(ab) \text{ für alle } r \in R, a, b \in A.$$

Ein R-Algebrenhomomorphismus ist ein Ringhomomorphismus, der gleichzeitig R-Modulhomomorphismus ist.

Übung 6.2.2. Sei A eine R-Algebra. Dann ist $R \to A : r \mapsto r1_A$ ein Ringhomomorphismus, sogar ein R-Algebrenhomomorphismus.

Beispiel 6.2.13.

(1) $\mathbb{C}[x]$ ist eine $\mathbb{C}$-Algebra. Sei

$$^- : \mathbb{C} \to \mathbb{C} : a + bi \mapsto a - bi \qquad a, b \in \mathbb{R},$$

die **komplexe Konjugation**. Dann ist

$$\mathbb{C}[x] \to \mathbb{C}[x] : \sum_{k=0}^{n} a_k x^k \mapsto \sum_{k=0}^{n} \overline{a_k} x^k$$

ein Ringhomomorphismus, jedoch kein $\mathbb{C}$-Algebrenhomomorphismus. (Man kann ihn jedoch als $\mathbb{R}$-Algebrenhomomorphismus auffassen.)

(2) Ist A eine R-Algebra und $I \trianglelefteq A$ ein Ideal, so ist A/I eine R-Algebra und ν_I ein R-Algebrenepimorphismus.

(3) Jeder Ring ist eine $\mathbb{Z}$-Algebra.

[2] in unserem Kontext meistens kommutativ

6.2.c Euklidische Ringe

Definition 6.2.14. Sei R ein kommutativer Ring mit Eins.

(1) R heißt **Integritätsbereich** oder auch **nullteilerfrei**, falls $1 \neq 0$ in R gilt und für alle $a, b \in R$, $ab = 0 \implies a = 0$ oder $b = 0$.

(2) R heißt **Hauptidealbereich**, falls R ein Integritätsbereich ist und jedes Ideal von R ein Hauptideal ist (siehe Beispiel 6.2.7.(6)).

(3) R heißt **Euklidischer Bereich** oder **Euklidischer Ring**, falls eine Abbildung

$$\nu : R \longrightarrow \mathbb{Z}_{\geq 0}$$

mit folgenden Eigenschaften existiert:

(a) $\nu(r) = 0$ genau dann, wenn $r = 0$;

(b) $\nu(r_1 r_2) = \nu(r_1)\nu(r_2)$;

(c) für $a \in R$ und $b \in R \setminus \{0\}$ existiert ein $q \in R$ und ein $r \in R$, so dass $a = qb + r$ mit $\nu(r) < \nu(b)$.

Beispiel 6.2.15.

(1) Jeder Teilring eines Körpers ist ein Integritätsbereich.

(2) Offenbar ist jeder Körper sowohl ein Hauptidealbereich als auch ein EUKLIDischer Ring (mit $\nu(a) = 1$ für alle $a \in K^*$).

(3) $\mathbb{Z}$ ist EUKLIDischer Bereich mit dem gewöhnlichen Absolutbetrag:

$$\nu(a) := |a| := \begin{cases} a & a \geq 0 \\ -a & a < 0. \end{cases}$$

(4) Sei K ein Körper. Dann ist $K[x]$ ein EUKLIDischer Bereich mit einer multiplikativen Variante des Grades:

$$\nu(a) := \begin{cases} 0 & a = 0 \\ 2^{\mathrm{Grad}(a)} & a \neq 0. \end{cases}$$

(5) $\mathbb{Z}[x] := \{p(x) \in \mathbb{Q}[x] \mid p(x) = \sum_{i=0}^{n} a_i x^i \text{ für } n \in \mathbb{N}, a_i \in \mathbb{Z}\}$ ist ein Integritätsbereich, da er ein Teilring des Körpers $\mathbb{Q}(x) = \{p(x)/q(x) \mid p(x), q(x) \in \mathbb{Q}[x], q(x) \neq 0\}$ ist. Jedoch ist $\mathbb{Z}[x]$ kein Hauptidealbereich, da z.B. $\langle 2, x\rangle$ kein Hauptideal ist.

Übung 6.2.3. Zeigen Sie der Ring $\mathbb{Z}[i] := \mathbb{Z}[x]/\langle x^2 + 1\rangle$ ist EUKLIDischer Bereich mit $\nu(a + bi) := a^2 + b^2$ für $a, b \in \mathbb{Z}$.

Satz 6.2.16. *Sei R ein Integritätsbereich. Dann gibt es einen Körper K, so dass $R \subseteq K$ Teilring von K ist und*

$$K = \{ab^{-1} | a \in R, b \in R \setminus \{0\}\}.$$

Dieser Körper ist bis auf Isomorphie eindeutig bestimmt und heißt **Quotientenkörper** *von R. Bezeichnung: $K = \mathrm{Quot}(R)$.*

Man beachte: In K sind alle Elemente $\neq 0$ von R zu Einheiten geworden, weshalb es Sinn macht ab^{-1}, $b^{-1}a$, oder a/b für $b \neq 0$ zu schreiben.

Beweis.

Existenz: Definiere $\widetilde{K} = \{(a,b) \mid a,b \in R, b \neq 0\}$ und eine Äquivalenzrelation $\approx$ auf $\widetilde{K}$: $(a,b) \approx (c,d)$ genau dann, wenn $ad = cb$. (Zur Veranschaulichung kann man die Tupel (a,b) als ungekürzte Brüche $\frac{a}{b}$ betrachten.) Die Menge der Äquivalenzklassen $\widetilde{K}/\approx =: K$ bildet einen Ring. Definiere nun $\frac{a}{b} :=$ Äquivalenzlasse von (a,b).
Addition und Multiplikation werden folgendermaßen definiert:

$$\begin{aligned} \frac{a}{b} + \frac{c}{d} &:= \frac{ad+bc}{bd} \\ \frac{a}{b} \cdot \frac{c}{d} &:= \frac{ac}{bd}, \text{ falls } b,d \neq 0. \end{aligned}$$

Zeige, dass die Addition wohldefiniert ist: Da $b,d,bd \neq 0$ sind, sind die Ausdrücke auf beiden Seiten wohldefiniert. Zum Beweis der Vertreterunabhängigkeit sei $\frac{a}{b} = \frac{a'}{b'}$ und $\frac{c}{d} = \frac{c'}{d'}$. Also ist $ab' = ba'$ und $cd' = dc'$. Behauptung: Es ist $\frac{ad+bc}{bd} = \frac{a'd'+b'c'}{b'd'}$. Dies gilt aber genau dann, wenn $bd(a'd' + b'c') = (ad+bc)b'd'$. Nach Ausmultiplizieren erhalte: $ba'dd' + bb'c'd = ab'dd' + bb'cd'$ dies gilt, da ja nach Voraussetzung $ba' = ab'$ und $cd' = dc'$. Genauso kann man zeigen, dass die Multiplikation auf K wohldefiniert ist.
Zeige nun, dass $(K,+)$ eine abelsche Gruppe ist mit Nullelement $0 := \frac{0}{1} = \frac{0}{c}, c \neq 0$ ($\frac{0}{a} = \frac{0}{b}$ für alle $a,b \in R$ mit $a,b \neq 0$). Wegen des Distributivgesetzes in R vereinfacht sich die Summe bei gleichen Nennern, also $\frac{a}{c} + \frac{b}{c} = \frac{a+b}{c}$. Dann ist also $\frac{b}{c} + \frac{0}{c} = \frac{b+0}{c} = \frac{b}{c}$. Negative: $-\frac{b}{c} := \frac{-b}{c}$. Die Kommutativität von $(K,+)$ folgt aus der Kommutativtät von R.
Zeige, dass $(K \setminus \{0\}, \cdot)$ eine abelsche Gruppe bildet mit $1 := \frac{1}{1} \neq \frac{0}{1} =: 0$ (also $K \setminus \{0\} \neq \emptyset$). Das Assoziativgesetz gilt, da $(R,\cdot)$ für Zähler und Nenner assoziativ ist, ebenso das Kommutativgesetz. Das Einselement ist $1 := \frac{1}{1} = \frac{a}{a}, a \neq 0$, das inverse Element zu $\frac{a}{b} \neq 0$ ist $\left(\frac{a}{b}\right)^{-1} := \frac{b}{a}$. Es bleibt noch zu zeigen, dass auch das Distributivgesetz gilt! (Übung)
Die Abbildung $\mu : R \longrightarrow K : r \to \frac{r}{1}$ ist ein Ringmonomorphismus, denn μ ist Ringhomomorphismus und aus $r \in \operatorname{Kern}\mu$ folgt $\frac{r}{1} = \frac{0}{1}$. Es gilt also: $r = r \cdot 1 = 0 \cdot 1 = 0$. Also identifiziere $r \in R$ mit $\frac{r}{1} \in K$. Beachte: Für $r \in R$ und $0 \neq s \in R$ gilt:

$$\frac{r}{s} = \frac{r}{1} \cdot \left(\frac{s}{1}\right)^{-1} = r \cdot s^{-1}$$

mit dieser Identifikation.

Eindeutigkeit: Sei K' ein weiterer Körper mit $R \subseteq K'$. Dann ist die Abbildung $\varepsilon : K \longrightarrow K' : \frac{a}{b} \mapsto a \cdot b^{-1}$ ein wohldefinierter Homomorphismus: Sei dazu $\frac{a}{b} = \frac{a'}{b'}$. Dann gilt: $ab' = ba'$ in R, also $ab^{-1} = a'b'^{-1}$ in K'. Aus der Homomorphie von ε folgt sofort: ε ist Monomorphismus, also $K \cong \operatorname{Bild}\varepsilon$. □

An dieser Stelle ist auf ein Problem hinzuweisen: Wir haben zwar eine Beschreibung der Elemente des Quotientenkörpers, mit der man Gleichheit testen kann. Man hat aber zunächst und im Allgemeinen keine Normalform für die Elemente, die natürlich viel effektiver wäre.

Beispiel 6.2.17.

(1) $\operatorname{Quot}(\mathbb{Z}) = \mathbb{Q}$. Hier lernt man eine Normalform in der Schule kennen.

(2) Sei K ein Körper. Dann heißt $K(x) := \operatorname{Quot}(K[x])$ der Körper der **rationalen Funktionen** in einer Variablen.

(3) Sei K ein Körper. Dann heißt $K(x_1,\ldots,x_n) := \operatorname{Quot}(K[x_1,\ldots,x_n])$ der Körper der **rationalen Funktionen** in n Variablen. Zum Beispiel ist

$$\frac{x_1^3 - x_2^3}{x_1^2 - x_2^2} = \frac{x_1^2 + x_1x_2 + x_2^2}{x_1 + x_2}.$$

Wir haben also Integritätsbereiche als die Teilringe von Körpern charakterisiert. Wir werden gleich sehen, dass in Hauptidealbereiche größte gemeinsame Teiler stets existieren.

Definition 6.2.18. Sei R ein Integritätsbereich und $a, b \in R$.

(1) $d \in R$ **teilt** a genau dann, wenn ein $q \in R$ existiert, mit $a = qd$, also genau dann wenn $a \in \langle d \rangle$. Bezeichnung: $d|a$.

(2) $a \in R \setminus (R^* \cup \{0\})$ heißt **prim**, falls

$$a|(b_1 b_2) \text{ impliziert } a|b_1 \text{ oder } a|b_2 \text{ für alle } b_1, b_2 \in R.$$

(3) Eine Zahl $d \in R$ heißt **größter gemeinsamer Teiler** $\mathrm{ggT}(a, b)$ von a und b, wenn

$$d \mid a, b \text{ und } c \mid a, b \implies c \mid d,$$

d.h. wenn d ein Teiler von a und ein Teiler von b ist und für jedes $c \in R$, welches a und b teilt, auch gilt, dass c ein Teiler von d.

(4) Eine Zahl $v \in R$ heißt **kleinstes gemeinsames Vielfaches** $\mathrm{kgV}(a, b)$ von a und b wenn

$$a, b \mid v \text{ und } a, b \mid w \implies v \mid w,$$

d.h. wenn v sowohl durch a als auch durch b teilbar ist und jedes $w \in R$, welches durch a und b teilbar ist, auch durch v teilbar ist.

Bemerkung 6.2.19. Sei R ein Integritätsbereich und $a, b \in R$. Dann gilt

$$a \mid b \iff b \in \langle a \rangle \iff \langle b \rangle \subseteq \langle a \rangle.$$

Satz 6.2.20. *Sei R ein Hauptidealbereich und $a, b \in R$. Dann existieren $\mathrm{ggT}(a, b) \in R$ und $\mathrm{kgV}(a, b) \in R$ und sind eindeutig bis auf Multiplikation mit Einheiten.*

Ende Vorl. 3

Beweis. Wir betrachten das Erzeugnis $\langle a \rangle + \langle b \rangle = \langle a, b \rangle \trianglelefteq R$. Da R ein Hauptidealbereich ist, ist dieses Ideal ein Hauptideal, also gibt es ein $d \in R$ mit $\langle a \rangle + \langle b \rangle = \langle d \rangle$. Dann gilt $a \in \langle d \rangle$, also d teilt a und ebenso d teilt b. Ist umgekehrt $c \in R$ ein Teiler von a und von b, so heißt dies, dass $a \in \langle c \rangle$ und $b \in \langle c \rangle$, also

$$\langle d \rangle = \langle a \rangle + \langle b \rangle \subseteq \langle c \rangle$$

und somit $c \mid d$.
Für das kleinste gemeinsame Vielfache gilt $\langle \mathrm{kgV}(a, b) \rangle = \langle a \rangle \cap \langle b \rangle$. Der Beweis hierfür und für die Eindeutigkeit bis auf Einheiten ist eine Übungsaufgabe. □

Bemerkung 6.2.21. Sei R ein EUKLIDischer Bereich. Dann ist R ein Hauptidealbereich.

Beweis. Wir zeigen zunächst, dass R nullteilerfrei ist. Seien $a, b \in R$ mit $ab = 0$. Dann ist $0 = \nu(ab) = \nu(a)\nu(b)$ also $\nu(a) = 0$ oder $\nu(b) = 0$, da $\mathbb{Z}$ als Teilring des Körpers $\mathbb{Q}$ nullteilerfrei ist. Somit folgt $a = 0$ oder $b = 0$.
Nun zeigen wir, dass jedes Ideal von R ein Hauptideal ist. Sei $\langle 0 \rangle \neq I \trianglelefteq R$. Wähle ein $a \in I \setminus \{0\}$ mit $\nu(a)$ minimal. (Dies ist möglich, da $\mathbb{Z}_{\geq 0}$ wohlgeortnet ist.)
Behauptung: $I = \langle a \rangle$. Ist nämlich $b \in I$, dann folgt: $b = aq + r$ mit $\nu(r) < \nu(a)$ für $q, r \in R$ und $r \in I$. Daraus folgt aber $r = 0$, d.h. $b = aq \in \langle a \rangle$. □

Man kann sagen, dass die EUKLIDischen Ringe besonders konstruktive Versionen der Hauptidealbereiche sind. Man hat nämlich den EUKLIDischen Algorithmus, um den ggT und damit auch Erzeuger von endlich erzeugten Idealen konstruktiv auszurechnen.

Übung 6.2.4. Man formuliere den EUKLIDischen Algorithmus (inklusive der Darstellung des größten gemeinsamen Teilers) für EUKLIDische Bereiche und zeige, wie man ihn zur Beschreibung von endlich erzeugten Idealen als Hauptideale benutzen kann.

6.2.d Der chinesische Restsatz

Wenn nicht anders angekündigt, bedeutet Ring ein kommutativer Ring mit Eins.

Satz 6.2.22. *(***Chinesischer Restsatz***) Sei R ein Ring und $I_1, \ldots, I_n$ paarweise* **teilerfremde** *Ideale von R, d.h. $I_i \trianglelefteq R$ und $I_i + I_j = R$ für $i, j = 1, \ldots, n$ mit $i \neq j$. Dann gilt:*

$$\begin{aligned} R/\bigcap_{i=1}^{n} I_i &\xrightarrow{\sim} R/I_1 \times R/I_2 \times \ldots \times R/I_n \\ r + \bigcap_{i=1}^{n} I_i &\mapsto (r + I_1, r + I_2, \ldots, r + I_n) \end{aligned}$$

ist ein Isomorphismus.

Bei Anwendungen will man häufig den zu dem obigen Isomorphismus inversen Isomorphismus ausrechnen. Man nennt den Satz auch den **Hauptsatz über das Lösen von simultanen Kongruenzen**. Dies ist wie folgt zu verstehen: Für $I \trianglelefteq R$ schreibt man für Elemente $r, s \in R$ statt $r + I = s + I$ auch schon mal $r \equiv s \pmod{I}$. Der obige Satz sagt also: Für beliebige $r_1, \ldots, r_n \in R$ gibt es ein $x \in R$ mit

$$x \equiv r_i \pmod{I_i} \qquad \text{für alle } i = 1, \ldots n$$

und die Lösungen sind eindeutig $\pmod{\bigcap_{i=1}^{n} I_i}$.

Beweis. Der Fall $n = 2$ ist klar: Der offensichtliche Homomorphismus

$$R \to R/I_1 \times R/I_2 : r \mapsto (r + I_1, r + I_2)$$

hat Kern $I_1 \cap I_2$ und ist wegen $I_1 + I_2 = R$ surjektiv (leichte Übung: Zeige, dass insbesondere $(1, 0)$ und $(0, 1)$ im Bild sind). Die Behauptung folgt aus dem Homomorphiesatz.
Der allgemeine Beweis erfolgt nun durch Induktion, die wir als Übung lassen. Der wesentliche Schritt steckt schon im Fall $n = 3$:
Behauptung: $I_1 + (I_2 \cap I_3) = R$. Zum Beweis beachte: Es existieren $e_1, e_1' \in I_1, e_2 \in I_2, e_3 \in I_3$ mit

$$1 = e_1 + e_2 = e_1' + e_3, \text{ also } 1 = 1 \cdot 1 = \underbrace{e_1 e_1' + e_1 e_3 + e_2 e_1'}_{=:e_{11} \in I_1} + \underbrace{e_2 e_3}_{=:e_{23} \in I_2 \cap I_3}.$$

Damit ist klar: $R = R \cdot 1 \cdot R = Re_{11}R + Re_{23}R \subseteq I_1 + (I_2 \cap I_3)$.
Wir wenden nun den Fall $n = 2$ nun zweimal an:

$$R/(I_1 \cap I_2 \cap I_3) \cong R/I_1 \times R/(I_2 \cap I_3) \cong R/I_1 \times R/I_2 \times R/I_3$$

mit den entsprechenden Isomorphismen. □

Übung 6.2.5. Sei q die Quersumme und a die alternierende Quersumme von natürlichen Zahlen. Dann gilt für jedes $n \in \mathbb{N}$:

- $n \equiv q(n) \pmod{9}$;
- $n \equiv a(n) \pmod{11}$.

Beispiel 6.2.23. Durch Kombination der Neunerprobe mittels Quersumme $\pmod 9$, der Zehnerprobe mittels letzter Stelle $\pmod{10}$ und der Elferprobe vermöge der alternierenden Quersumme $\pmod{11}$ kann man Rechnungen mit ganzen Zahlen, die nur Multiplikationen und Additionen involvieren $\pmod{990}$ überprüfen.

Sei $s := 124$, $t := 351$. Wir wollen $(s+t)t = 166275$ verifizieren. Modulo 10 ist diese Rechnung richtig (letzte Ziffer stimmt).
Modulo 9: $q(s) = 7$, $q(t) = 0$ also $q((s+t)t) = 0$.
Modulo 11: $a(s) = 3$, $a(t) = -1$, also $a((s+t)t) = -2$.
Es ist aber $a(166275) = 5 - 7 + 2 - 6 + 6 - 1 = -1$ also ist an der Rechnung etwas falsch. Die richtige Antwort ist $(s+t)t = 166725$ mit $a(166725) = 5 - 2 + 7 - 6 + 6 - 1 = 9 \equiv -2 \pmod{11}$.

Der Chinesische Restsatz 6.2.22 kann für Euklidische Ringe wie folgt formuliert und auch bewiesen werden:

Satz 6.2.24. *(**Chinesischer Restsatz für Euklidische Ringe**) Sei R ein Euklidischer Ring und $a_1, \ldots, a_n$ paarweise* **teilerfremde** *Elemente von R, d.h. $\mathrm{ggT}(a_i, a_j) = 1$ für $i, j = 1, \ldots, n$ mit $i \neq j$. Dann gilt:*

$$\begin{aligned} \bigcap_{i=1}^{n} \langle a_i \rangle &= \left\langle \prod_{i=1}^{n} a_i \right\rangle \\ \varphi : R/\left\langle \prod_{i=1}^{n} a_i \right\rangle &\xrightarrow{\sim} R/\langle a_1 \rangle \times R/\langle a_2 \rangle \times \ldots \times R/\langle a_n \rangle \\ r + \left\langle \prod_{i=1}^{n} a_i \right\rangle &\mapsto (r + \langle a_1 \rangle, r + \langle a_2 \rangle, \ldots, r + \langle a_n \rangle) \end{aligned}$$

ist ein Isomorphismus dessen Umkehrabbildung mit dem Euklidischen Algorithmus berechnet werden kann.

Beweis. Da $\mathrm{ggT}(a_i, a_j) = 1$ für alle $i \neq j$ gilt, ist $\mathrm{kgV}(a_1, \ldots, a_n) = \prod_{i=1}^{n} a_i$ und somit $\cap_{i=1}^{n} \langle a_i \rangle = \langle \prod_{i=1}^{n} a_i \rangle$. Es genügt also den Algorithmus anzugeben. Ein allgemeines Element von $R/\langle a_1 \rangle \times R/\langle a_2 \rangle \times \ldots \times R/\langle a_n \rangle$ ist von der Form $X := (r_1 + \langle a_1 \rangle, \ldots, r_n + \langle a_n \rangle)$. Gesucht ist ein $r \in R$ mit $r + \langle a_i \rangle = r_i + \langle a_i \rangle$ für alle $i = 1, \ldots, n$, also $\varphi(r + A) = X$ mit $A := \langle \prod_{i=1}^{n} a_i \rangle$. Dazu setze $B_i := \prod_{j \neq i} a_j$. Da $\mathrm{ggT}(a_i, a_j) = 1$ für alle $i \neq j$ gilt, folgt auch $\mathrm{ggT}(a_i, B_i) = 1$, es gibt also $x_i, y_i \in R$ mit

$$1 = x_i a_i + y_i B_i.$$

Setze $e_i := y_i B_i$. Dann ist

$$e_i + \langle a_i \rangle = 1 + \langle a_i \rangle \text{ und } e_i + \langle a_j \rangle = 0 + \langle a_j \rangle \text{ für alle } j \neq i$$

also $\varphi(e_i + A) = (0, \ldots, 0, 1, 0, \ldots, 0)$ mit 1 an der i-ten Stelle. Da φ ein R-Ringhomomorphismus ist, ergibt sich

$$\varphi^{-1}(X) = \sum_{i=1}^{n} r_i e_i + A. \qquad \square$$

Beispiel 6.2.25. Wir wollen das simultane Kongruenzensystem

$$\begin{aligned} x &\equiv 1 \pmod{3} \\ x &\equiv 2 \pmod{4} \\ x &\equiv 3 \pmod{5} \end{aligned}$$

lösen. Wir haben den Isomorphismus

$$\varphi : \mathbb{Z}/\langle 60 \rangle \xrightarrow{\sim} \mathbb{Z}/\langle 3 \rangle \times \mathbb{Z}/\langle 4 \rangle \times \mathbb{Z}/\langle 5 \rangle$$

und haben das Problem gelöst, wenn wir

$$\begin{aligned} e_1 + \langle 60\rangle &:= \varphi^{-1}((\overline{1},\overline{0},\overline{0})),\\ e_2 + \langle 60\rangle &:= \varphi^{-1}((\overline{0},\overline{1},\overline{0})),\\ e_3 + \langle 60\rangle &:= \varphi^{-1}((\overline{0},\overline{0},\overline{1})) \end{aligned}$$

kennen, denn die Lösungsmenge ist dann $e_1 + 2e_2 + 3e_3 + \langle 60\rangle$. Mit Hilfe des EUKLIDischen Algorithmus für 3 und $4 \cdot 5$ etc. bekommen wir dann e_1 etc. :

$$\begin{aligned} 1 &= 7 \cdot 3 + (-1) \cdot 20, \quad \text{also} \quad e_1 = -20,\\ 1 &= 4 \cdot 4 + (-1) \cdot 15, \quad \text{also} \quad e_2 = -15,\\ 1 &= 5 \cdot 5 + (-2) \cdot 12, \quad \text{also} \quad e_3 = -24. \end{aligned}$$

Also $x \in -122 + \langle 60\rangle = -2 + \langle 60\rangle$.

Ende Vorl. 4

Beispiel 6.2.26 (Lagrangeinterpolation). Sei K ein Körper und $p(x) \in K[x]$. Wie wir schon wissen gilt für $a \in K$

$$p(a) = 0 \text{ genau dann, wenn } p(x) \in \langle x - a\rangle \trianglelefteq K[x],$$

(was man auch durch Entwickeln nach Potenzen von $x - a$ sehen kann). Dies liefert auch

$$p(x) \equiv p(a) \pmod{\langle x - a\rangle}.$$

Sind nun $a_1, \ldots, a_n \in K$ paarweise verschiedene Elemente, so sind die Ideale $\langle x - a_i\rangle$ paarweise teilerfremd und der Chinesische Restsatz liefert

$$K[x]/\left\langle \prod_{i=1}^{n}(x - a_i)\right\rangle \cong \bigtimes_{i=1}^{n} \underbrace{K[x]/\langle x - a_i\rangle}_{\cong K}.$$

Die Restklassen der Elemente $\widetilde{e}_i(x) := \prod_{j\neq i}(x - a_j) \in K[x]$ liefern auf der rechten Seite Tupel, deren Komponenten alle Null sind, außer der i-ten, welche gleich $\widetilde{e}_i(a_i) = \prod_{j\neq i}(a_i - a_j)$ ist. Hieraus ergibt sich sofort die LAGRANGEsche Interpolationsformel: Eine Abbildung $f : K \to K$ wird interpoliert an den Stützstellen a_i durch das Polynom (vom Grad $< n$):

$$\sum_{i=1}^{n} f(a_i)\frac{\widetilde{e}_i(x)}{\widetilde{e}_i(a_i)}.$$

Wenn man im Falle $K = \mathbb{R}$ auch noch Ableitungen an den Stellen a_i vorgeben will, muss man mit $\langle (x-a_i)^k\rangle \trianglelefteq K[x]$ arbeiten statt mit $\langle x-a_i\rangle$ und kommt zur **Hermiteinterpolation**.

Die ringdirekten Summen enthalten etwas suspekte Elemente, die sich aber oben schon als sehr nützlich erwiesen haben.

Definition 6.2.27. Sei R ein Ring (kmE) und $a \in R$ mit $a \neq 0$. Man nennt a einen **Nullteiler**, wenn ein $b \in R$ existiert mit $b \neq 0$ und $ab = 0$. Weiter heißt a **nilpotent**, falls ein $n \in \mathbb{N}_{>1}$ existiert mit $a^n = 0$.

Klar, nilpotente Elemente sind Nullteiler, aber oben haben wir schon Nullteiler gesehen, die nicht nilpotent sind.

Beispiel 6.2.28.

(1) Seien $m, k \in \mathbb{Z}$ beide größer als 1. Dann ist $m + \langle m^k\rangle \in \mathbb{Z}/\langle m^k\rangle$ nilpotent.

(2) (Wurzel ziehen). Aus dem Chinesischen Restsatz bekommen wir

$$\begin{aligned} \pi_1 \times \pi_2 : \mathbb{R}[x]/\langle x^2-2\rangle &\xrightarrow{\sim} \underbrace{\mathbb{R}}_{\cong \mathbb{R}[x]/\langle x-\sqrt{2}\rangle} \times \underbrace{\mathbb{R}}_{\cong \mathbb{R}[x]/\langle x+\sqrt{2}\rangle} : \\ \overline{x} := x + \langle x^2-2\rangle &\mapsto (\sqrt{2}, -\sqrt{2}) = (\pi_1(\overline{x}), \pi_2(\overline{x})). \end{aligned}$$

Wir wollen auf der rechten Seite rechnen, können aber nur auf der linken Seite Nullteiler erkennen. Unser Ziel ist eine numerische Approximation von $\sqrt{2}$ zu bestimmen. Klar: Die einzigen Nullteiler der Form $a + \overline{x}$ mit $a \in \mathbb{R}$ sind $\sqrt{2} + \overline{x}$ und $-\sqrt{2} + \overline{x}$. Die Idee ist nun so: Sei $b \in \mathbb{Q}_{\geq 1}$ so gewählt, dass $a := \overline{x} - b$ auf der rechten Seite einem (a_1, a_2) entspricht mit $|a_1| < 1$ und $|a_2| > 1$. (Betragstriche markieren Absolutbeträge.) Es ist $a^2 = 2 - 2b\overline{x} + b^2$. Dann ist $|a_1^2| < |a_1| < 1$, und $\pi_1(a^2/(-2b)) = \pi_1\left(\overline{x} - \frac{b^2+2}{2b}\right)$ hat einen noch kleineren Absolutbetrag, so dass $(b^2+2)/(2b)$ eine noch bessere Näherung an $\sqrt{2}$ ist als b. Durch fortgesetztes Quadrieren verdoppelt sich immer die Anzahl der signifikanten Dezimalstellen, wir haben quadratische Konvergenz. Fängt man also mit $b = 1$ an, so erhält man die folgende Folge die gegen $\sqrt{2}$ konvergiert:

$$1, \frac{3}{2}, \frac{17}{12} = 1.416666667, \frac{577}{408} = 1.414215686, \frac{665857}{470832} = 1.414213562, ...$$

Übung 6.2.6. Vergleiche obige Methode des Wurzelziehens mit dem NEWTONverfahren aus der Numerik.

6.2.e Der chinesische Restsatz und die Hauptraumzerlegung.

Die Hauptraumzerlegung aus Bemerkung 5.5.6 im LA I Skript erinnert stark an den chinesischen Restsatz. Sei dazu K ein Körper, $\mathcal{V}$ ein e.e. K-Vektorraum und $\alpha \in \mathrm{End}(\mathcal{V})$ mit Minimalpolynom

$$\mu_\alpha = p_1^{n_1} \cdots p_k^{n_k}$$

wobei die p_i paarweise verschiedene normierte irreduzible Polynome in $K[x]$ sind und $n_i \in \mathbb{N}$. Unter den obigen Voraussetzungen lässt sich $\mathcal{V}$ eindeutig schreiben als direkte Summe α-invarianter Teilräume $\mathcal{U}_i$:

$$\mathcal{V} = \bigoplus_{i=1}^{k} \mathcal{U}_i,$$

so dass das Minimalpolynom der Einschränkung von α auf $\mathcal{U}_i$ genau $p_i^{n_i}$ ist. Setzt man $q_i := \prod_{j \neq i} p_j^{n_j}$ so ist $\mathcal{U}_i = \mathrm{Bild}(q_i(\alpha))$.

Bemerkung 6.2.29.

(1) Das Minimalpolynom von α war definiert als normierter Erzeuger des Kerns des Einsetzungshomomorphismus

$$\varepsilon_\alpha : K[x] \to \mathrm{End}(\mathcal{V}), p \mapsto p(\alpha)$$

$\mathrm{Kern}(\varepsilon_\alpha) = \langle \mu_\alpha \rangle$, $\mathrm{Bild}(\varepsilon_\alpha) = K[\alpha]$. Also ist nach dem Homomorphiesatz für Ringe $K[x]/\langle \mu_\alpha \rangle \cong K[\alpha]$.

(2) Setzt man $R := K[x]/\langle \mu_\alpha \rangle$, so wird $\mathcal{V}$ ein R-Modul durch

$$(p + \langle \mu_\alpha \rangle)V := \varepsilon_\alpha(p)(V) = p(\alpha)(V).$$

(3) Die Struktur von R bekommen wir aus dem Chinesischen Restsatz:

$$R = K[x]/\langle p_1^{n_1} \cdots p_k^{n_k} \rangle \cong K[x]/\langle p_1^{n_1} \rangle \times K[x]/\langle p_2^{n_2} \rangle \times \ldots \times K[x]/\langle p_k^{n_k} \rangle$$

und Urbilder e_i der **Idempotente** $\pi_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ (mit 1 an der i-ten Stelle) können mit dem Algorithmus in 6.2.24 ermittelt werden, indem wir für alle i mit dem Euklidischen Algorithmus $1 = x_i p_i^{n_i} + y_i q_i$ schreiben, und $e_i = \varepsilon_\alpha(y_i q_i) \in K[\alpha]$ setzen. Aus einer Übung wissen wir, dass das Bild von e_i gleich dem von $q_i(\alpha)$ ist.

(4) Also ist $\mathcal{V}$ ein Modul für dem Produktring $K[x]/\langle p_1^{n_1} \rangle \times K[x]/\langle p_2^{n_2} \rangle \times \ldots \times K[x]/\langle p_k^{n_k} \rangle$ und die Hauptraumzerlegung $\mathcal{V} \cong \bigoplus_{i=1}^k \mathcal{U}_i$ ist nichts anderes als die Zerlegung von $\mathcal{V}$ gemäß Bemerkung 6.2.11, wobei $\mathcal{U}_i$ in natürlicher Weise ein $K[x]/\langle p_i^{n_i} \rangle$-Modul ist.

6.3 Elementare Teilbarkeitstheorie für Ringe

Definition 6.3.1. Sei R Ring (kmE) und $a, b \in R$.

(1) a **teilt** b (in R) oder b ist ein **Vielfaches** von a, in Zeichen $a \mid b$, falls ein $r \in R$ existiert mit $ar = b$.

(2) a ist **assoziiert** zu b (in R), in Zeichen $a \sim b$, falls $a \mid b$ und $b \mid a$.

Klar: Die Vielfachen von a bilden gerade das Hauptideal $\langle a \rangle$. Teilen ist eine transitive Relation auf R und $\sim$ ist eine Äquivalenzrelation auf R. Auch klar für $a, b \in R$: Falls ein $e \in R^*$ existiert mit $a = eb$, dann gilt $a \sim b$. Wenn wir versuchen die Umkehrung zu beweisen, stoßen wir auf eine kleine Schwierigkeit. Diese verschwindet, wenn wir verlangen, dass R ein Integritätsbereich ist.

Bemerkung 6.3.2. In einem Integritätsbereich sind die Assoziertenklassen der Elemente gegeben durch R^*a mit $a \in R$. Assoziiertheit ist also genau die Äquivalenzrelation die durch die Operation der Einheitengruppe auf R (durch Multiplikation) definiert ist:

$$\sim = \sim_{R^*} .$$

Beweis. Es ist klar, dass jedes Element der Form ua mit $u \in R^*$ zu $a = u^{-1}ua$ assoziiert ist. Sei nun $b \in R$ mit $b \mid a$ und $a \mid b$. Dann gibt es $r, s \in R$ mit $a = br$ und $b = as$.
Ist $a = 0$, so ist $b = as = 0 \in R^*a = \{0\}$.
Sei also $a \neq 0$. Wir wollen zeigen, dass r und s Einheiten sind, sogar $r = s^{-1}$ also $rs = 1$: Denn es ist $a = br = asr$ also $a(1-sr) = 0$. Da $a \neq 0$ und R nullteilerfrei, folgt jetzt $1-sr = 0$ also $sr = 1$. □

Ende Vorl. 5

Bemerkung 6.3.3. Ist R ein Integritätsbereich, so ist der Polynomring $R[x_1, \ldots, x_n]$ auch ein Integritätsbereich. Zum Beweis definieren wir den **Grad** eines Polynoms

$$p = p(x_1, \ldots, x_n) = \sum a_{i_1 \ldots i_n} x_1^{i_1} \cdots x_n^{i_n} \in R[x_1, \ldots, x_n]$$

als

$$\operatorname{Grad} p := \max\{i_1 + \cdots + i_n | a_{i_1 \ldots i_n} \neq 0\} \text{ für } p \neq 0 \quad \text{und} \quad \operatorname{Grad} 0 := -\infty.$$

Für $p, q \in R[x_1, \ldots, x_n]$ gilt offenbar

$$\operatorname{Grad} pq = \operatorname{Grad}(p) + \operatorname{Grad}(q),$$

woraus man leicht sieht, dass man keine Nullteiler hat.

Unsere nächste Frage lautet: Welche Restklassenringe sind Integritätsbereiche?

Definition 6.3.4. Sei R ein Ring (kmE).

(1) Ein Ideal $I \trianglelefteq R$ heißt **Primideal**, falls $I \neq R$ für alle $r, s \in R$ gilt:

$$r, s \notin I \text{ impliziert } rs \notin I.$$

(2) Ein Ideal I heißt **maximales Ideal**, falls $I \neq R$ und für jedes Ideal J von R, welches I enthält gilt $J = I$ oder $J = R$.

Bemerkung 6.3.5. Sei R Ring (kmE) und $I \trianglelefteq R$.

(1) R/I ist genau dann Integritätsbereich, wenn I Primideal ist.

(2) R/I ist genau dann ein Körper, wenn I ein maximales Ideal ist.

Beweis.

(1) R/I ist Integritätsbereich, genau dann wenn für alle $r, s \in R$ mit $(r + I)(s + I) = 0$ gilt, dass entweder $r + I = 0$ ist oder $s + I = 0$, also entweder $r \in I$ oder $s \in I$. Da $(r + I)(s + I) = rs + I$ ist dies gleichbedeutend mit ($rs \in I \implies r \in I$ oder $s \in I$), also damit, dass I ein Primideal ist.

(2) R/I ist ein Körper, genau dann wenn jedes $r + I \in R/I \setminus \{0 + I\}$ ein multiplikatives Inverses hat. Sei also $r \in R \setminus I$. Dann ist $\langle r, I\rangle = \langle r\rangle + I$ ein Ideal von R, welches I echt enthält. Ist I ein maximales Ideal, so ist $\langle r\rangle + I = R$ und es gibt $a \in R$, $i \in I$ mit $ar + i = 1$. Dann ist $(a + I)(r + I) = 1 + I$ und $(a + I)$ ein multiplikatives Inverses von $r + I$. Die Umkehrung geht genauso. □

Beispiel 6.3.6.

(1) Sei R Integritätsbereich. Das Hauptideal $\langle x\rangle \trianglelefteq R[x]$ ist ein Primideal, da $R[x]/\langle x\rangle \cong R$ Integritätsbereich.

(2) Sei R ein Ring (kmE) und $I \trianglelefteq R$ maximal, dann ist I prim.

(3) Sei ein R Ring (kmE). Genau dann ist $\{0\}$ Primideal, wenn R Integritätsbereich ist.

Nun kommen wir zur Teilbarkeitstheorie in Integritätsbereichen. Es wird in dem Sinne ganz elementar, als dass wir wieder mehr von Elementen als von Idealen sprechen. Zuerst eine Ernüchterung: Die Begriffe "prim" und "irreduzibel" für Elemente in einem Integritätsbereich fallen im Allgemeinen auseinander:

Definition 6.3.7. Sei R ein Integritätsbereich.

(1) Ein Element $a \in R \setminus (R^* \cup \{0\})$ heißt **irreduzibel** oder **unzerlegbar**, falls jede Faktorisierung von a in R trivial ist, das heißt, falls gilt:

$$a = a_1 a_2 \text{ für } a_i \in R \text{ impliziert } a_1 \in R^* \text{ oder } a_2 \in R^*.$$

Sonst **reduzibel** oder **zerlegbar**.

(2) Ein Element $a \in R \setminus (R^* \cup \{0\})$ heißt **prim**, falls

$$a|(b_1 b_2) \text{ impliziert } a|b_1 \text{ oder } a|b_2 \text{ für alle } b_i \in R.$$

Beispiel 6.3.8.

(1) In $R = \mathbb{Z}$ ist 2 prim und irreduzibel.

(2) Sei K ein Integritätsbereich und $R = K[x]$. Dann ist x irreduzibel (da vom Grad 1) und prim: Falls nämlich $x|(a(x)b(x))$ und $x \nmid a(x)$, so folgt $a(0) \neq 0$, also $b(0) = 0$ und $x|b(x)$.

Bemerkung 6.3.9. Sei R ein Integritätsbereich und $a \in R$.

(1) Ist a prim, so ist a auch irreduzibel.

(2) a ist prim genau dann, wenn $\langle a \rangle \trianglelefteq R$ ein Primideal $\neq 0$ ist.

Beweis.

(1) Sei a prim und reduzibel, d.h. $a = a_1 a_2$ mit $a_i \in R \setminus R^*$. Dann gilt: $a|a_1 a_2$ und $a \nmid a_1$ (sonst wäre $a_2 \in R^*$) und $a \nmid a_2$ (sonst wäre $a_1 \in R^*$). Dies ist aber ein Widerspruch zu der Voraussetzung, dass a prim ist.

(2) Sei $a \in R$ prim. Dann ist $\langle a \rangle \neq 0$. Weiter gilt für $r, s \in R$:

$$rs \in \langle a \rangle \Leftrightarrow a \mid rs \Leftrightarrow a \mid r \text{ oder } a \mid s \Leftrightarrow r \in \langle a \rangle \text{ oder } s \in \langle a \rangle.$$

Die Umkehrung folgt ebenso. □

Übung 6.3.1. Es gibt irreduzible Elemente, die nicht prim sind:
Sei $R = \mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} \mid a, b \in \mathbb{Z}\}$. Es gilt: $3 \cdot 2 = (1 + \sqrt{-5})(1 - \sqrt{-5})$. Zeige: Die Elemente $2, 3, (1 + \sqrt{-5}), (1 - \sqrt{-5})$ sind allesamt in R irreduzibel, aber nicht prim.
Hinweis: Betrachte die multiplikative Norm $\nu(a + b\sqrt{-5}) := a^2 + 5b^2$.

An dieser Stelle sei angemerkt, dass DEDEKIND, der sicher als einer der Urväter der modernen Algebra einzuschätzen ist, wegen der schlechten Eigenschaften des Teilbarkeitsbegriffes für Elemente, den Idealbegriff erfunden hat, um in bestimmten Situationen der Zahlentheorie doch noch zu einer befriedigenden Teilbarkeitstheorie, diesmal aber für Ideale (sprich ideale Zahlen) zu kommen. Im obigen Beispiel kann man nämlich alle vier Zahlen als Ideale noch weiter zerlegen:

$$\begin{aligned}
\langle 3 \rangle &= \langle 3, 1 + \sqrt{-5} \rangle \langle 3, 1 - \sqrt{-5} \rangle \\
\langle 2 \rangle &= \langle 2, 1 + \sqrt{-5} \rangle \langle 2, 1 - \sqrt{-5} \rangle \\
\langle 1 + \sqrt{-5} \rangle &= \langle 3, 1 + \sqrt{-5} \rangle \langle 2, 1 + \sqrt{-5} \rangle \\
\langle 1 - \sqrt{-5} \rangle &= \langle 3, 1 - \sqrt{-5} \rangle \langle 2, 1 - \sqrt{-5} \rangle
\end{aligned}$$

Satz 6.3.10. *Ist R Hauptidealbereich, so ist jedes irreduzible Element prim.*

Beweis. Sei $a \in R \setminus (R^* \cup \{0\})$ irreduzibel. Wir wissen, dass a genau dann prim ist, wenn $\langle a \rangle$ ein Primideal $\neq 0$ ist. Wir können sogar zeigen, dass $\langle a \rangle$ ein maximales Ideal ist, denn sei $\langle a \rangle \subseteq I \trianglelefteq R \neq I$. Da R Hauptidealbereich ist, folgt $I = \langle d \rangle$ für ein $d \in R \setminus (R^* \cup \{0\})$. Somit $a \in \langle d \rangle$, also $a = dd'$ für ein $d' \in R$. Das bedeutet $d' \in R^*$, denn a ist irreduzibel und $d \notin R^*$. Da $a \sim d$ ist, haben wir $I = \langle a \rangle$. □

Übung 6.3.2. Sei R Hauptidealbereich und $a, b \in R \setminus \{0\}$. Zeige $\langle a \rangle + \langle b \rangle = \langle a, b \rangle = \langle \mathrm{ggT}(a, b) \rangle$ und $\langle a \rangle \cap \langle b \rangle = \langle \mathrm{kgV}(a, b) \rangle$ und schließlich $\langle a \rangle \langle b \rangle = \langle ab \rangle$.

Aus Bemerkung 6.3.9.(2) und dem letzten Beweis folgern wir:

Folgerung 6.3.11. *Ist R ein Hauptidealbereich, so ist jedes Primideal, das ungleich dem 0-Ideal ist, ein maximales Ideal.*

Beispiel 6.3.12.

(1) $\mathbb{Z}[x]$ ist kein Hauptidealbereich, denn das Primideal $\langle x\rangle$ ist kein maximales Ideal.

(2) $K[x,y]$ (sogar für ein Körper K) ist ebenfalls kein Hauptidealbereich.

(3) $\mathbb{Z}$ ist ein Hauptidealbereich. Die Primzahlen sind bis auf Multiplikation mit Einheiten $\{\pm 1\}$ die irreduziblen Elemente von $\mathbb{Z}$. Die $\mathbb{F}_p := \mathbb{Z}/\langle p\rangle$ für Primzahlen p sind die einzigen Restklassenkörper von $\mathbb{Z}$.

(4) Für jeden Körper K ist $K[x]$ ein Hauptidealbereich. Die irreduziblen Polynome in $K[x]$ sind gleich den Primelementen in $K[x]$ und gleich den irreduziblen Elementen in $K[x]$. Die Restklassenkörper von $K[x]$ sind alle von der Form $K[x]/\langle p(x)\rangle$, wo $p(x) \in K[x]$ irreduzibel ist.

Satz 6.3.13. *Sei R ein Hauptidealbereich, $a \in R \setminus (R^*\{0\})$. Dann gibt es im Wesentlichen eindeutige Primelemente $a_1,\dots,a_n \in R$ mit $a = a_1\cdots a_n$. Die Eindeutigkeit bedeutet: Falls $a = a_1\cdots a_n = b_1\cdots b_m$ mit $a_i, b_j \in R$ irreduzibel, so folgt $n = m$ und nach Umnummerierung $a_i \sim b_i$ für $i = 1,\dots,n$.*

Beweis. Wir können zunächst genauso vorgehen wie für $R = \mathbb{Z}$: Ist a irreduzibel, so ist a schon prim und wir sind fertig. Ist a nicht irreduzibel, so können wir $a = bc$ schreiben mit $b, c \in R \setminus R^*$ und dann mit b und c weitermachen. Wieso hört dieser Prozess auf? Falls nicht, so konstruiert man eine Folge von echten Teilern $\dots \mid b_{i+1} \mid b_i \mid \dots \mid b_1 = b \mid a$, so dass

$$\langle a\rangle \subsetneq \langle b_1\rangle \subsetneq \langle b_2\rangle \dots$$

Sei $I := \bigcup_{i\in\mathbb{N}}\langle b_i\rangle$. Da die Ideale $\langle b_i\rangle$ eine Kette bilden, ist I ein Ideal von R. Nun ist R ein Hauptidealbereich also gibt es ein $d \in R$ mit $I = \langle d\rangle$. Da $d \in I = \bigcup_{i\in\mathbb{N}}\langle b_i\rangle$ existiert also ein $i \in \mathbb{N}$ mit $d \in \langle b_i\rangle$. Aber dann ist $\langle b_j\rangle = \langle b_i\rangle = I$ für alle $j \geq i$ ein Widerspruch dazu, dass b_{i+1} ein echter Teiler von b_i ist.
Die Eindeutigkeit zeigt man genauso wie für $\mathbb{Z}$: Sei $a = a_1\cdots a_n = b_1\cdots b_m$ mit $a_i, b_j \in R$ irreduzibel. Dann ist a_1 prim, $a_1 \mid a = b_1\cdots b_m$, also gibt es ein i mit $a_1 \mid b_i$. Da b_i irreduzibel ist, gilt also $a_1 \sim b_i$, usw. mit Induktion über die Anzahl m von Faktoren. □

Ende Vorl. 6

6.4 Moduln über Hauptidealbereichen.

6.4.a Der Struktursatz

Endlich erzeugte Moduln über Hauptidealbereichen haben eine sehr schöne Struktur. Um algorithmisch einen solchen Modul auf Normalform zu bringen benötigt man lediglich die algorithmische Berechenbarkeit der Bézout Identität. Die „algorithmisch zugänglichen" HIB sind also die Euklidischen Ringe.

Definition 6.4.1. Sei R ein Integritätsbereich und M ein R-Modul. Ein $m \in M$ heißt **Torsionselement**, falls das **Annihilatorideal**[3]

$$\mathrm{Ann}_R(m) := \{r \in R \mid rm = 0\}$$

von $\langle 0\rangle$ verschieden ist. Der **Torsionsteilmodul**, also der Teilmodul aller Torsionselemente von M, wird mit $T(M)$ bezeichnet. Der Modul M heißt **torsionsfrei**, falls $T(M) = \{0\}$ und ein **Torsionsmodul**, falls $T(M) = M$ ist.

[3]Es gilt: $\mathrm{Ann}_R(M) = \cap_{m\in M}\mathrm{Ann}_R(m)$.

Man konkretisiert sich die Begriffe mit Hilfe der folgenden Beispiele, wobei man $R = \mathbb{Z}$ und $R = K[x]$ betrachtet.

Beispiel 6.4.2. Sei R ein Hauptidealbereich mit $K := \mathrm{Quot}(R) \neq R$.

(1) K ist ein nicht endlich erzeugter, torsionsfreier R-Modul. (Beweis später.)

(2) K/R ist ein nicht endlich erzeugter R-Torsionsmodul. (Zu kompliziert für uns.)

(3) Jeder freie R-Modul ist torsionsfrei, insbesondere ${}_RR = R$ selbst.

(4) Ist M beliebiger R-Modul, so ist $M/T(M)$ torsionsfrei (Übung).

(5) Jeder **zyklische Modul** (sprich, von einem Element erzeugt), ist entweder isomorph zu ${}_RR \equiv R$ oder zu ${}_RR/Ra$ für ein $a \in R \setminus \{0\}$. Im letzteren Fall haben wir einen Torsionsmodul.

Obwohl ${}_RR/Ra$ und $R/\langle a\rangle$ als abelsche Gruppen und auch als R-Moduln[4] identisch sind, wollen wir doch im Kontext von Moduln lieber die erste und im Kontext von Restklassenringen die zweite Notation benutzen. Und da

$$\mathrm{Ann}_R({}_RR/Ra) := \cap_{x\in {}_RR/Ra}\,\mathrm{Ann}_R(x) = \langle a\rangle$$

können wir den R-Modul ${}_RR/Ra$ auch als $R/\langle a\rangle$-Modul auffassen. Als letzter ist er sogar frei.

(6) Endliche direkte Summen der Moduln aus (5) sind typische endlich erzeugte R-Moduln. Nur scheinbar allgemeiner sind die folgenden:

(7) Faktormoduln von $R^{n\times 1}$ (freier Modul auf n Erzeugern) nach Teilmoduln.

Unser Ziel wird sein, zu zeigen, dass die Moduln aus (7) nicht allgemeiner sind als die Moduln aus (6). Wir machen zwei kleine Schritte in diese Richtung:

Lemma 6.4.3. *Sei R Integritätsbereich und $(e_1 = (1,0,\ldots,0)^{tr}, e_2, \ldots, e_n)$ die Standardbasis von $R^{n\times 1}$, also freies Erzeugendensystem des freien R-Moduls $\mathrm{Fr}_R(\underline{n}) \equiv R^{n\times 1}$. Sei $k \leq n$ und $d_1, \ldots, d_k \in R \setminus \{0\}$. Dann gilt*

$$\begin{aligned} R^{n\times 1}/\langle d_1e_1, \ldots, d_ke_k\rangle_R &= \left(\bigoplus_{i=1}^{n} Re_i\right) / \left(\bigoplus_{i=1}^{k} Rd_ie_i\right) \\ &\cong \bigoplus_{i=1}^{k} Re_i/Rd_ie_i \oplus \bigoplus_{i=k+1}^{n} Re_i \\ &\cong \bigoplus_{i=1}^{k} {}_RR/Rd_i \oplus R^{(n-k)\times 1}. \end{aligned}$$

Falls in dieser Situation noch zusätzlich d_i teilt d_{i+1} für $i = 1, \ldots, k-1$ gilt, so nennt man $(e_1, \ldots, e_n)$ und $(d_1e_1, \ldots, d_ke_k)$ **kompatible Basen**, *genauer ein Paar kompatibler Basen. (Den Begriff* **Basis** *benutzen wir als Synonym für freies Erzeugendensystem.)*

Beweis. Übung.
Hinweis: Bestimme den offensichtlichen Epimorphismus

$$R^{n\times 1} \to \bigoplus_{i=1}^{k} {}_RR/Rd_i \oplus R^{(n-k)\times 1}$$

[4] wenn man die R-Modulstruktur auf $R/\langle a\rangle$ wie erwartet definiert!

und wende den Homomorphiesatz für Moduln an. Beachte:

$$R^{k\times 1} \to \bigoplus_{i=1}^{k} Rd_ie_i : (a_1, \ldots, a_k)^{tr} \mapsto (a_1d_1, \ldots, a_kd_k, 0, \ldots, 0)^{tr}$$

ist ein R-Modulisomorphismus. □

Bislang wissen wir nicht einmal, dass Teilmoduln freier endlich erzeugter Moduln über Hauptidealbereichen frei sind, geschweige denn, ob kompatible Basen existieren. Hier ein erstes Indiz.

Lemma 6.4.4. *Sei R ein Hauptidealbereich und $M \cong R^{n\times 1}$ ein freier R-Modul von Rang n. Dann ist jeder R-Teilmodul von M endlich erzeugter freier R-Modul auf $k \leq n$ freien Erzeugern.*

Beweis. Wir führen den Beweis durch Induktion über n. Für $n = 1$ ist die Sache klar, da Teilmoduln von ${}_RR$ Ideale von R sind, also Hauptideale. Sei nun $T \leq_R M$ und $(e_1, \ldots, e_n)$ die Standard-R-Basis von M. Zu der Zerlegung

$$M = Re_1 \oplus \langle e_2, \ldots, e_n\rangle_R$$

gehört die Projektion $\pi : M \to Re_1$. Dann ist $\pi(T) \leq Re_1$. Im Falle $\pi(T) = \{0\}$ greift die Induktionsvoraussetzung sofort, da dann $T = \operatorname{Kern}\pi_{|T} \leq \langle e_2, \ldots, e_n\rangle$. Sonst ist $\pi(T) = Rde_1$ für ein $d \in R \setminus \{0\}$. Beachte, $\pi(T)$ ist frei auf de_1. Wähle $t \in T$ mit $\pi(t) = de_1$. Dann definiert

$$\iota : Rde_1 \to T : de_1 \mapsto t$$

einen R-Modulmonomorphismus und es gilt (Übung)

$$T = Rt \oplus \operatorname{Kern}\pi_{|T}.$$

Wegen $\operatorname{Kern}\pi_{|T} \leq \langle e_2, \ldots, e_n\rangle_R$ können wir die Induktionsvoraussetzung benutzen und bekommen unsere Behauptung. □

Wir wollen jetzt die Existenz kompatibler Basen beweisen, indem wir sowohl beim Teilmodul $R^{k\times 1} \cong T$ als auch bei $R^{n\times 1} \cong M$ Basistransformationen vornehmen.

Bemerkung 6.4.5. Sei R ein Ring (kmE).

(1) Die Beschreibung von Homomorphismen von freien R-Moduln in freie R-Moduln (alle endlich erzeugt) geschieht wie bei Vektorräumen durch Matrizen bezüglich Basen. (Alle relevanten Formeln aus der linearen Algebra bleiben gültig.)

(2) Automorphismen von freien R-Moduln (vom Rang n) und Basistransformationen werden beschrieben durch Matrizen aus

$$(R^{n\times n})^* = \mathrm{GL}_n(R) = \{g \in R^{n\times n} \mid \det(g) \in R^*\}.$$

Sei nun R ein Hauptidealbereich.

(3) Für $a, b \in R \setminus \{0\}$ mit $\mathrm{ggT}(a, b) = d$ gibt es $s, t \in R$ mit $sa + tb = d$. Es gilt

$$U_{(a,b)} := \begin{pmatrix} s & t \\ -\frac{b}{d} & \frac{a}{d} \end{pmatrix} \in \mathrm{GL}_2(R) \text{ und } U_{(a,b)} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} d \\ 0 \end{pmatrix}.$$

Satz 6.4.6. *Sei R Hauptidealbereich und $C \in R^{k\times n}$. Dann existieren Matrizen $A \in \mathrm{GL}_k(R)$ und $B \in \mathrm{GL}_n(R)$, so dass*

$$D = ACB = \begin{pmatrix} \mathrm{Diag}(d_1, \ldots, d_l) & 0 \\ 0 & 0 \end{pmatrix}$$

mit $d_i \in R \setminus \{0\}, l \leq \min(k, n)$ und $d_i \mid d_{i+1}$ für $i = 1, \ldots, l-1$. Die Matrix ACB nennt man die **Smith-Form** *von C.*

Beweis. Setze $C_1 := C$. Wir führen reversible Zeilenoperationen durch, und erhalten $C_2 := A_1C_1, C_3 := A_2C_2, \ldots, C_r := A_{r-1}C_{r-1}$, so dass die erste Spalte von C_r gleich $(d, 0, \ldots, 0)^{tr}$. Dabei sind die A_i Permutationsmatrizen oder von der Form $\mathrm{Diag}(U_{(a,b)}, I_{k-2})$, wenn die obersten zwei Einträge a, b der jeweils ersten Spalte von C_i von Null verschieden sind. Beachte, d ist der grösste gemeinsame Teiler der Einträge der ersten Spalte von C.
Danach führen wir reversible Spaltenoperationen durch und erhalten $C_{r+1} := C_rB_1, \ldots,$ $C_{r+s} := C_{r+s-1}B_s$, so dass die erste Zeile von C_{r+s} gleich $(d', 0, \ldots, 0)$ ist. Dabei sind die B_i Permutationsmatrizen oder von der Form $\mathrm{Diag}(U^{tr}_{(a,b)}, I_{n-2})$, wenn die ersten zwei Einträge a, b der jeweils ersten Zeile von C_i von Null verschieden sind. Klar: $d' \mid d$. Aber leider ist jetzt die erste Spalte nicht mehr notwendig ausgeräumt, so dass man den ersten Schritt wiederholen muss. Da aber R keine echten unendlichen Teilerketten zulässt, hat man nach endlich vielen Wiederholungen die Matrix $C_t := \mathrm{Diag}(d_1, C') \in R^{k\times n}$.
Rekursives Anwenden der Methode auf C' liefert nach endlich vielen Schritten schliesslich Matrizen $A' \in \mathrm{GL}_n(R), B' \in \mathrm{GL}_k(R)$, so dass

$$C' := A'CB' = \begin{pmatrix} \mathrm{Diag}(d_1, \ldots, d_l) & 0 \\ 0 & 0 \end{pmatrix}$$

mit $d_i \in R \setminus \{0\}$. Sollte für ein i noch die Bedingung $d_i \mid d_{i+1}$ verletzt sein, sind noch weitere Umformungen durchzuführen. Es genügt, diese für 2×2-Matrizen zu demonstrieren:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \mathrm{Diag}(d_i, d_{i+1}) = \begin{pmatrix} d_i & d_{i+1} \\ 0 & d_{i+1} \end{pmatrix}$$

also eine Matrix, die man mit den anfänglichen Methoden wieder auf die Form

$$\mathrm{Diag}(\mathrm{ggT}(d_i, d_{i+1}), \mathrm{kgV}(d_i, d_{i+1}))$$

transformieren kann. Nach endlich vielen Schritten hat man die geforderte Gestalt. □

Übung 6.4.1. Gib ein effektives Verfahren für Matrizen über EUKLIDischen Bereichen an, welches versucht, immer das Matrixelement einer festen Spalte bzw. einer festen Zeile mit dem größten ν-Wert abzubauen, bis die Spalte oder Zeile ausgeräumt ist.

Beispiel 6.4.7. Sei $R := \mathbb{Z}$ und $A := \begin{pmatrix} 6 & 2 \\ 8 & 7 \end{pmatrix} \in \mathbb{Z}^{2\times 2}$.

Mit $U_1 := \begin{pmatrix} -1 & 1 \\ -4 & 3 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z})$ folgt $U_1A = \begin{pmatrix} 2 & 5 \\ 0 & 13 \end{pmatrix}$.

Mit $W_1 := \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z})$ folgt $U_1AW_1 = \begin{pmatrix} 1 & 0 \\ -13 & 26 \end{pmatrix}$.

Mit $U_2 := \begin{pmatrix} 1 & 0 \\ 13 & 1 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z})$ folgt $U_2U_1AW_1 = \mathrm{Diag}(1, 26)$.

Ende Vorl. 7

Beispiel 6.4.8. Simultane Kongruenzen

$$\begin{aligned} x_1 + x_2 &\equiv 0 \pmod{\mathbb{Z}} \\ x_1 - x_2 &\equiv 1/4 \pmod{\mathbb{Z}} \end{aligned}$$

sind für $x_1, x_2 \in \mathbb{R}$ zu lösen. Wir schreiben dies als erweiterte Matrix $(A|b)$ und führen Zeilenumformungen auf die erweiterte Matrix (die wir uns nicht merken müssen und) und Spaltenumformungen auf den linken Teil der erweiterten Matrix, die wir uns aber merken müssen: $Ax = b \iff UAW \underbrace{(W^{-1}x)}_{y} = Ub$.

$$\left(\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & -1 & 1/4 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & -2 & 1/4 \end{array}\right) \xrightarrow{W:=\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}} \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 2 & 1/4 \end{array}\right)$$

Aus der letzten Matrix lesen wir die Zwischenlösung

$$y = \begin{pmatrix} z_1 \\ 1/8 + 1/2 z_2 \end{pmatrix} \qquad \text{mit } z_1, z_2 \in \mathbb{Z}$$

ab und erhalten als endgültige Lösung durch Multiplikation mit W:

$$x = \begin{pmatrix} 1/8 + z_1 + 1/2 z_2 \\ -1/8 - 1/2 z_2 \end{pmatrix} \qquad \text{mit } z_1, z_2 \in \mathbb{Z}$$

Lineare Differentialgleichungssysteme mit konstanten Koeffizienten erweisen sich auch als Kongruenzsysteme über $\mathbb{R}[x]$ oder $\mathbb{C}[x]$. Solange die rechte Seite Null ist, sind keine Probleme der Analysis involviert. Im inhomogenen Fall braucht man aus der Analysis die Methode der Variation der Konstanten. Wir beschränken uns auf den homogenen Fall.

Beispiel 6.4.9. Seien x_1, x_2, x_3 unendlich oft differenzierbare Funktionen auf $\mathbb{R}$ oder Elemente von $\mathbb{R}[[t]]$. Gesucht sind die Lösungen des Differentialgleichungssystems

$$\begin{aligned} x_1' - x_2' + x_3'' &= b_1 \\ x_1 + x_2'' + x_3' &= b_2 \end{aligned}$$

mit $b_1 = b_2 = 0$. In Matrizen, wobei D dann die Ableitung induzieren soll:

$$C \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 \text{ mit } C := \begin{pmatrix} D & -D & D^2 \\ 1 & D^2 & D \end{pmatrix},$$

Zeilenumformungen:

$$\left(\begin{array}{ccc|c} D & -D & D^2 & b_1 \\ 1 & D^2 & D & b_2 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & D^2 & D & b_2 \\ 0 & D^3 + D & 0 & -b_1 + Db_2 \end{array}\right)$$

Spaltenumformungen mit

$$W := \begin{pmatrix} 1 & -D^2 & -D \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ liefert } \begin{pmatrix} 1 & 0 & 0 \\ 0 & D^3 + D & 0 \end{pmatrix}$$

als Matrix der linken Seite. Wegen $D^3 + D = D(D^2 + 1)$ löst sich dieses System sehr leicht:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ a + b\sin(t) + c\cos(t) \\ f(t) \end{pmatrix}$$

mit $a, b, c \in \mathbb{R}$ und f eine unendlich oft differenzierbare Funktion $\mathbb{R} \to \mathbb{R}$. Durch Heranmultiplizieren von W erhalten wir die Lösungen des ursprünglichen Systems:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b\sin(t) + c\cos(t) - f'(t) \\ a + b\sin(t) + c\cos(t) \\ f(t) \end{pmatrix}$$

Hauptsatz 6.4.10. *Sei R ein Hauptidealbereich.*

(1) Ist M ein endlich erzeugter R-Modul, so gibt es $s, t \in \mathbb{Z}_{\geq 0}$ und $d_1, \ldots, d_t \in R \setminus (R^* \cup \{0\})$ mit $d_i \mid d_{i+1}$ für alle i, so dass

$$M \cong_R R^{s\times 1} \oplus \underbrace{\bigoplus_{i=1}^{t} {}_R R/Rd_i}_{\cong T(M)}.$$

(2) Gilt

$$R^{s\times 1} \oplus \bigoplus_{i=1}^{t} {}_R R/Rd_i \cong_R R^{s'\times 1} \oplus \bigoplus_{i=1}^{t'} {}_R R/Rd'_i$$

mit $s, t, s', t' \in \mathbb{Z}_{\geq 0}$ und $d_1, \ldots, d_t, d'_1, \ldots, d'_{t'} \in R\setminus(R^*\cup\{0\})$ mit $d_i \mid d_{i+1}$ für $i = 1, \ldots, t-1$ und $d'_i \mid d'_{i+1}$ für $i = 1, \ldots, t' - 1$, so gilt $s = s', t = t'$ und $d_i \sim d'_i$ für $i = 1, \ldots, t$. Man nennt s den **Rang** *(genauer torsionsfreien Rang) von M und d_i den i-ten* **Elementarteiler** *von M.*

Beweis.

(1) Folgt sofort aus den vorangegangenen Lemmata und Satz 6.4.6:
M ist nach Bemerkung 6.2.3, Lemma 6.4.4 und Beispiel 6.4.2.(7) gegeben als Faktormodul $R^{k\times 1}/N$, wobei der Teilmodul $N \leq R^{k\times 1}$ durch die Spalten einer $k \times n$ Matrix C gegeben ist. Invertierbare Spaltenumformungen von C sind Übergänge zu anderen Erzeugendensystemen vom Teilmodul N. Invertierbare Zeilenumformungen von C sind Übergänge zu anderen Basen (=freien Erzeugendensystemen) von $R^{k\times 1}$. Sei D die zu C gehörige Smith-Normalform. Dann gilt:

$$M \cong R^{k\times 1}/\langle C_{-,1}, \ldots, C_{-,n}\rangle \cong R^{k\times 1}/\langle D_{-,1}, \ldots, D_{-,n}\rangle \cong \underbrace{\bigoplus_{i=1}^{t} {}_R R/Rd_i}_{\cong T(M)} \oplus R^{s\times 1}.$$

(2) Bezeichne die linke Seite mit M und die rechte mit N. Aus $M \cong N$ folgt

$$R^{s\times 1} \cong M/T(M) \cong N/T(N) \cong R^{s'\times 1}.$$

Sei $p \in R$ ein Primelement und $F := R/\langle p\rangle$ der zugehörige Restklassenkörper. Dann gilt:

$$s = \dim_F(\underbrace{R^{s\times 1}/pR^{s\times 1}}_{\cong (R/Rp)^{s\times 1} = (R/\langle p\rangle)^{s\times 1} = F^{s\times 1}}) = \dim_F(R^{s'\times 1}/pR^{s'\times 1}) = s'.$$

Weiter folgt $T(M) \cong T(N)$, also auch

$$\langle d_t \rangle = \mathrm{Ann}_R(T(M)) = \mathrm{Ann}_R(T(N)) = \langle d_{t'} \rangle \text{ d.h. } d_t \sim d'_{t'}.$$

Um die d_i zu rekonstruieren, brauchen wir nur die p-Potenzen zu testen, welche in den d_i aufgehen, wobei p die Primteiler von d_t durchläuft. Man betrachtet hierzu die Zahlenfolge

$$\dim_F p^i T(M)/p^{i+1}T(M) = \dim_F p^i T(N)/p^{i+1}T(N)$$

für $i = 0, 1, \ldots k(p)$, wo $p^{k(p)}$ die höchste p-Potenz ist, die in d_t aufgeht. Es ist klar, wie man (durch Übergang zur sogenannten konjugierten Partition[5]) die p-Potenzanteile der d_i und der d'_i rekonstruieren kann und so $d_i \sim d'_i$ und $t = t'$ beweist. Beachte: $t = t'$ ist die minimale Erzeugendenzahl von $T(M) \cong T(N)$ und gleich dem Maximum der Dimensionen der $T(M)/pT(M) \cong T(N)/pT(N)$ als $R/\langle p \rangle$-Vektorräume, wo p die Primteiler von d_t durchläuft. □

Ende Vorl. 8

Folgerung 6.4.11. *Sei R ein Hauptidealbereich und $M = T(M)$ ein endlich erzeugter R-Torsionsmodul mit $\mathrm{Ann}(M) = \langle d \rangle$. Ist $d \sim \prod_i p_i^{n_i}$ die Faktorisierung in Potenzen von Primelementen. Dann ist M ebenfalls ein Modul über dem Restklassenring $R/\mathrm{Ann}(M) = R/\langle d \rangle = R/\langle \prod_i p_i^{n_i} \rangle$, der sich sich nach dem chinesischen Restsatz wie folgt zerlegt:*

$$R/\langle d \rangle \cong \bigtimes_i R/\langle p_i^{n_i} \rangle.$$

Und entsprechend zerlegt sich M nach Bemerkung 6.2.11 als

$$M \cong \bigoplus_i (R/\langle p_i^{n_i} \rangle)\, M = \bigoplus_i M/p_i^{n_i} M$$

und $M/p_i^{n_i}M$ ist $R/\langle p_i^{n_i} \rangle$-Modul.
Ist $p \in R$ prim, so sind die endlich erzeugten $R/\langle p^n \rangle$-Moduln durch endliche monoton steigende Folgen a natürlicher Zahlen $\leq n$ charakterisiert: a liefert den Modul

$$\bigoplus_i {}_R R/Rp^{a_i}.$$

Beispiel 6.4.12. Sei $M = \mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}/6\mathbb{Z} \oplus \mathbb{Z}/24\mathbb{Z} = T(M)$. Dann ist $\mathrm{Ann}(M) = \langle 24 \rangle$ und M ist ein Modul über dem Restklassenring $\mathbb{Z}/\mathrm{Ann}(M) = \mathbb{Z}/\langle 24 \rangle \cong \mathbb{Z}/\langle 3 \rangle \times \mathbb{Z}/\langle 8 \rangle$. Entsprechend zerlegt sich M als

$$\begin{aligned}
(\mathbb{Z}/\langle 3 \rangle)M \oplus (\mathbb{Z}/\langle 8 \rangle)M &= M/3M \oplus M/8M \\
&= (\mathbb{Z}/(3\mathbb{Z}+3\mathbb{Z}) \oplus \mathbb{Z}/(6\mathbb{Z}+3\mathbb{Z}) \oplus \mathbb{Z}/(24\mathbb{Z}+3\mathbb{Z})) \\
&\quad \oplus (\mathbb{Z}/(3\mathbb{Z}+8\mathbb{Z}) \oplus \mathbb{Z}/(6\mathbb{Z}+8\mathbb{Z}) \oplus \mathbb{Z}/(24\mathbb{Z}+8\mathbb{Z})) \\
&= (\mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}) \\
&\quad \oplus (\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/8\mathbb{Z}).
\end{aligned}$$

Folgerung 6.4.13. *Sei R ein Hauptidealbereich und und M ein e.e. torsionsfreier R-Modul. Dann ist M frei.*

Achtung: Dies ist falsch für nicht e.e. torsionsfreie R-Moduln. Ist z.B. $R = \mathbb{Z}$, so ist $\mathbb{Q}$ ein torsionsfreier $\mathbb{Z}$-Modul. Jedoch ist $\mathbb{Q}$ nicht frei, denn für je zwei Elemente $a/b, c/d \in \mathbb{Q}$ gilt $(bc)a/b - (ad)c/d = 0$.

Folgerung 6.4.14. *Sei R ein Hauptidealbereich mit $K := \mathrm{Quot}(R) \neq R$. Dann ist K ist ein nicht endlich erzeugter, torsionsfreier R-Modul. (Übung)*

[5] Siehe Definition 7.2.12.

6.4.b Der Hauptsatz über endlich erzeugte abelsche Gruppen

Wir wenden unsere Erkenntnisse jetzt speziell für den Ring $R = \mathbb{Z}$ an. Die $\mathbb{Z}$-Moduln sind genau die abelschen Gruppen, endlich erzeugte $\mathbb{Z}$-Moduln also endlich erzeugte abelsche Gruppen und wir erhalten den folgenden Struktursatz.

Folgerung 6.4.15. *(Hauptsatz über endlich erzeugte abelsche Gruppen) Sei $G = \langle g_1, \ldots, g_n\rangle$ eine endlich erzeugte abelsche Gruppe. Dann gibt es $r, t \in \mathbb{Z}_{\geq 0}$, $f_1, \ldots, f_r, h_1, \ldots, h_t \in G$ und $d_1, \ldots, d_s \in \mathbb{Z}$ mit $d_1|d_2|\ldots|d_s$, so dass*

$$G = \langle f_1\rangle \times \ldots \times \langle f_r\rangle \times \langle h_1\rangle \times \ldots \times \langle h_t\rangle \cong \mathbb{Z}^r \oplus \mathbb{Z}/d_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/d_s\mathbb{Z}.$$

Die Abelsche Gruppe G ist genau dann endlich, wenn ihr Rang als $\mathbb{Z}$-Modul Null ist. In dem Fall ist $|G| = d_1 \cdots d_t$ und G ist ein Modul über $\mathbb{Z}/\langle d_t\rangle$ bzw. über $\mathbb{Z}/\langle d_1 \cdots d_t\rangle$.

Folgerung 6.4.16.

(1) Sei G eine endliche abelsche Gruppe der Ordnung $|G| = p_1^{n_1} \cdots p_s^{n_s}$. Dann ist G ein $\mathbb{Z}/\langle \prod_{i=1}^s p_i^{n_i}\rangle$-Modul. Dieser Ring ist nach dem chinesischen Restsatz isomorph zu $\bigtimes_{i=1}^s \mathbb{Z}/\langle p_i^{n_i}\rangle$. Dementsprechend lässt sich G eindeutig schreiben als

$$G = \bigoplus_{i=1}^{s} P_i$$

wo $|P_i| = p_i^{n_i}$ ist. (P_i nennt man auch die p_i-Sylowgruppe von G).

(2) Sei P eine abelsche Gruppe von Primzahlpotenzordnung $|P| = p^n > 1$ (p-Gruppe). Dann gibt es eindeutiges $t \in \mathbb{N}$, $a_1 \leq \ldots \leq a_t \in \mathbb{N}$ mit $\sum_i a_i = n$ und $P \cong \mathbb{Z}/p^{a_1}\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/p^{a_t}\mathbb{Z}$.

Beispiel 6.4.17. Die abelschen Gruppen der Ordnung $24 = 2^3 \cdot 3$ sind:

(1) $\mathbb{Z}/24\mathbb{Z} = \mathbb{Z}/8\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$ entsprechend der Partitionen (3=3, 1=1),

(2) $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/4\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$ entsprechend der Partitionen (3=2+1, 1=1),

(3) $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$ entsprechend der Partitionen (3=1+1+1, 1=1).

Folgerung 6.4.18. *Sei G eine endliche abelsche Gruppe, $a \in G$. Das $n \in \mathbb{N}$ mit $\langle n\rangle = \mathrm{Ann}_{\mathbb{Z}}(a)$ nennt man auch die* **Ordnung**[6] *von a. Es gilt $\mathrm{ord}(a) = |\langle a\rangle|$ und $\mathrm{ord}(a)$ teilt*[7] *$|G|$.*
Existiert ein $a \in G$, so dass $\langle a\rangle = G$, so heißt G **zyklische Gruppe**. *Von Ordnung $m \in \mathbb{Z}_{\geq 0}$ ist $C_m := (\mathbb{Z}/m\mathbb{Z}, +)$ bis auf Isomorphie die einzige zyklische Gruppe von Ordnung m.*

Wir wollen unsere Ergebnisse jetzt auf Einheitengruppen von Restklassenringen von $\mathbb{Z}$ anwenden, nämlich

$$(\mathbb{Z}/\langle m\rangle)^* = \{a + m\mathbb{Z} \in \{1, \ldots, m-1\} \mid \mathrm{ggT}(a, m) = 1\}.$$

Definition 6.4.19. Die Eulersche φ-Funktion:

$$\varphi : \mathbb{N} \implies \mathbb{N};\ \ \varphi(m) := |(\mathbb{Z}/\langle m\rangle)^*|$$

Bemerkung 6.4.20.

(1) (Kleiner Satz von Fermat) Ist p eine Primzahl und $a \in \mathbb{Z}$, dann $a^p \equiv a \pmod p$, denn $\varphi(p) = p - 1$.

[6] n ist also die kleinste positive natürliche Zahl mit $na = 0$ (falls G additiv geschrieben wird) bzw. $a^n = 1$ (falls G multiplikativ geschrieben wird.

[7] da $\langle |G|\rangle \subset \mathrm{Ann}_{\mathbb{Z}}(G) \subset \mathrm{Ann}_{\mathbb{Z}}(a) = \langle \mathrm{ord}(a)\rangle$.

(2) Ist p eine Primzahl, so ist $(\mathbb{Z}/\langle p^a\rangle)^* = \mathbb{Z}/\langle p^a\rangle \setminus p(\mathbb{Z}/\langle p^a\rangle)$ also $\varphi(p^a) = p^a - p^{a-1} = p^{a-1}(p-1)$.

(3) Ist $m = \prod_{j=1}^s p_j^{\alpha_j}$ mit paarweise verschiedenen Primzahlen p_j und $\alpha_j \geq 1$, dann ist $\mathbb{Z}/\langle m\rangle = \bigtimes_j \mathbb{Z}/\left\langle p_j^{\alpha_j}\right\rangle$ nach dem chinesischen Restsatz, also auch

$$(\mathbb{Z}/\langle m\rangle)^* = \bigtimes_j (\mathbb{Z}/\left\langle p_j^{\alpha_j}\right\rangle)^*.$$

(4) Es ist $\varphi(m) = \prod_{j=1}^s \varphi\left(p_j^{\alpha_j}\right) = \prod_{j=1}^s p_j^{\alpha_j - 1}(p_j - 1)$.

Lemma 6.4.21. $\sum_{d|m} \varphi(d) = m$.

Beweis. Sei $m = \prod_{j=1}^s p_j^{\alpha_j}$. Induktion über $\sum_{j=1}^s \alpha_j =: n$.
$n = 1$: Klar (beachte $\varphi(1) = 0$).
Induktionsschritt:

$$\sum_{d|m} \varphi(d) = \sum_{d|(m/p_1)} \varphi(d) + \sum_{d|\left(m/p_1^{\alpha_1}\right)} \varphi(p_1^{\alpha_1} d) =$$

$$m/p_1 + p_1^{\alpha_1 - 1}(p_1 - 1)m/p_1^{\alpha_1} = m. \qquad \square$$

Der folgende Satz verallgemeinert den kleinen Satz von Fermat.

Ende Vorl. 9

Satz 6.4.22. *Sei K ein endlicher Körper. Dann ist seine Einheitengruppe zyklisch.*

Beweis. Betrachte $(K^*, \cdot)$ als $\mathbb{Z}$-Modul über $z \cdot a := a^z$ für $z \in \mathbb{Z}$ und $a \in K^*$. Wir zeigen: Für jeden Teiler d von $|K| - 1$ hat K^* genau $\varphi(d)$ Elemente der Ordnung d.
Denn: Sei $\psi(d) := |\{a \in K^* \mid \operatorname{ord}(a) = d\}|$. Dann gilt mit Folgerung 6.4.18 die Implikation: $d \nmid (|K| - 1) \implies \psi(d) = 0$.
Ist $a \in K^*$ und $\operatorname{ord}(a) = d$, so ist $\langle a\rangle$ die Menge der d verschiedenen Nullstellen von $X^d - 1$. Die Elemente der Ordnung d in $\langle a\rangle$ sind genau die a^m mit $\operatorname{ggT}(m, d) = 1$, also ist ihre Anzahl genau $\varphi(d)$. Also ist $\varphi(d) \leq \psi(d)$. Deshalb ist nach dem vorigen Lemma

$$|K| - 1 = \sum_{d||K|-1} \varphi(d) \leq \sum_{d||K|-1} \psi(d) = |K^*| = |K| - 1.$$

Also $\psi(d) = \varphi(d)$. Insbesondere ist $\psi(|K| - 1) \neq 0$ und $K^* = \langle a\rangle$ für jedes Element $a \in K^*$ mit $\operatorname{ord}(a) = |K| - 1$. $\square$

Satz 6.4.23. *Sei p eine Primzahl, $\alpha \geq 1$.*

(1) Ist $p > 2$, so ist $(\mathbb{Z}/\langle p^\alpha\rangle)^* \cong (\mathbb{Z}/p^{\alpha-1}\mathbb{Z}, +) \oplus (\mathbb{Z}/(p-1)\mathbb{Z}, +)$.

(2) Ist $\alpha \geq 2$, so ist $(\mathbb{Z}/\langle 2^\alpha\rangle)^* \cong (\mathbb{Z}/2\mathbb{Z}, +) \oplus (\mathbb{Z}/2^{\alpha-2}\mathbb{Z}, +)$.

Beweis.

(1) Zeige: Es reicht zu zeigen, dass die Gruppe zyklisch ist. Die Behauptung folgt dann aus dem chinesischen Restsatz angewandt auf $\mathbb{Z}/\langle(p-1)p^{\alpha-1}\rangle$ mit der additiven Gruppe $\mathbb{Z}/(p-1)p^{\alpha-1}\mathbb{Z}$. Mit Hilfe des binomischen Lehrsatzes zeigt man: Ist $\beta \geq 1$, $b \in \mathbb{Z}$, $p \nmid b$, dann ist

$$(1 + p^\beta b)^p = 1 + p^{\beta+1} c \text{ für ein } c \in \mathbb{Z} \text{ mit } p \nmid c. \qquad (*)$$

Ist $a \in \mathbb{Z}$ mit $\langle a + p\mathbb{Z}\rangle = (\mathbb{Z}/\langle p\rangle)^*$, so ist nach Fermat $a^{p-1} = 1 + pb$ für ein $b \in \mathbb{Z}$. Gilt $p \nmid b$, so ist $\langle a + p^\alpha\mathbb{Z}\rangle = (\mathbb{Z}/\langle p^\alpha\rangle)^*$ wegen (*). Falls $p \mid b$, dann ersetze a durch $a + p$ und erhalte $\langle a + p + p^\alpha\mathbb{Z}\rangle = (\mathbb{Z}/\langle p^\alpha\rangle)^*$ (kleine Rechnung als Übung).

(2) Übung. □

Damit haben wir die Struktur der Einheitengruppe von $\mathbb{Z}/\langle m\rangle$ bestimmt, denn dieser Ring ist nach dem chinesischen Restsatz isomorph zu

$$\mathbb{Z}/\langle p_1^{\alpha_1}\rangle \times \ldots \times \mathbb{Z}/\langle p_s^{\alpha_s}\rangle,$$

falls $m = p_1^{\alpha_1} \cdots p_s^{\alpha_s}$ eine Primfaktorzerlegung von m ist. Also ist seine Einheitengruppe das direkte Produkt der Einheitengruppen

$$(\mathbb{Z}/\langle m\rangle)^* \cong (\mathbb{Z}/\langle p_1^{\alpha_1}\rangle)^* \times \ldots \times (\mathbb{Z}/\langle p_s^{\alpha_s}\rangle)^*,$$

was wir bereits in der letzten Bemerkung gebraucht haben.

Kapitel 7

Normalformen für Matrizen.

7.1 Ähnlichkeit von Matrizen

Wir wollen über das Klassifikationsproblem der Endomorphismen eines endlich dimensionalen K-Vektorraumes sprechen oder, was äquivalent hierzu ist, über eine Normalform der Matrizen des Endomorphismus, die durch eine gewisse Basiswahl erreicht wird.

Definition 7.1.1. Sei $\mathcal{V}$ ein endlich erzeugter Vektorraum über dem Körper K der Dimension $n \in \mathbb{N}$.

(1) Zwei Endomorphismen $\alpha, \beta \in \mathrm{End}(\mathcal{V})$ heißen **ähnlich** oder **konjugiert** unter $\mathrm{GL}(\mathcal{V})$, falls ein $\gamma \in \mathrm{GL}(\mathcal{V})$ existiert mit $\alpha = \gamma \circ \beta \circ \gamma^{-1}$, d.h. α, β liegen in derselben Bahn unter der Operation

$$\mathrm{GL}(\mathcal{V}) \times \mathrm{End}(\mathcal{V}) \to \mathrm{End}(\mathcal{V}) : (\gamma, \zeta) \mapsto \gamma \circ \zeta \circ \gamma^{-1}.$$

(2) Zwei Matrizen $A, B \in K^{n\times n}$ heißen **ähnlich** oder **konjugiert** unter $\mathrm{GL}_n(K)$, falls ein $g \in \mathrm{GL}_n(K)$ existiert mit $A = gBg^{-1}$, d.h. A, B liegen in derselben Bahn unter der Operation

$$\mathrm{GL}_n(K) \times K^{n\times n} \to K^{n\times n} : (g, C) \mapsto gCg^{-1}.$$

(3) Sei R ein Integritätsbereich. Zwei Matrizen $A, B \in R^{n\times m}$ heißen **äquivalent** (über R), wenn es $g \in \mathrm{GL}_n(R)$, $h \in \mathrm{GL}_m(R)$ gibt mit $gAh = B$.

Klar: Äquivalenz von Matrizen ist eine Äquivalenzrelation.

Den Struktursatz 6.4.6 kann man auch so formulieren: Ist R ein Hauptidealbereich, so ist jede Matrix äquivalent zu einer Diagonalmatrix (sogar mit den Teilbarkeitsbedingungen). Ähnliche Matrizen sind äquivalent über K. Die Umkehrung gilt jedoch nicht. Z.B. ist jede Matrix in $\mathrm{GL}_n(K)$ ist äquivalent über K zur Einheitsmatrix, aber dies ist nicht richtig für Ähnlichkeit.
Die Matrizen, die einen festen Endomorphismus beschreiben, bilden eine Ähnlichkeitsklasse, wenn man die Basis variieren lässt. Umgekehrt sind zwei Endomorphismen genau dann durch dieselbe Matrix beschreibbar, wenn sie konjugiert sind.

Bemerkung 7.1.2. μ_α, χ_α sind Invarianten (aus denen man z.B. Spur und Determinante ablesen kann). Für jedes Polynom $p(x) \in K[x]$ ist

$$\mathrm{End}(\mathcal{V}) \to \mathbb{Z}_{\geq 0} : \alpha \mapsto \mathrm{Dim}(\mathrm{Kern}(p(\alpha)))$$

eine Invariante.

Wir wollen in diesem Abschnitt für einen gegebenen Endomorphismus $\alpha \in \mathrm{End}(\mathcal{V})$ eine möglichst einfache Form für die Matrix ${}^B\alpha^B$ finden. Bislang haben wir nur Teilergebnisse und Spezialfälle erledigt, an die wir uns kurz erinnern wollen. Für den ganzen Abschnitt sei $\mathcal{V}$ ein endlich erzeugter K-Vektorraum. Hier ist eine Zusammenfassung einiger relevanter Ergebnisse aus dem ersten Semester.

Bemerkung 7.1.3. Ist $\mu_\alpha(x) = \prod_{i=1}^{\ell} p_i^{m_i}$ die Zerlegung des Minimalpolynoms in normierte irreduzible und paarweise verschiedene Polynome p_i, dann gilt:

(1) Das charakteristische Polynom ist gegeben durch $\chi_\alpha(x) = \prod_{i=1}^{\ell} p_i^{c_i}$ mit $c_i \geq m_i$.

(2) Man hat eine kanonische Zerlegung von $\mathcal{V}$ in die p_i-**Haupträume** $\mathcal{V}_i := \mathrm{Kern}(p_i^{m_i}(\alpha))$, die alle α-invariant sind:
$$\mathcal{V} = \bigoplus_{i=1}^{\ell} \mathcal{V}_i.$$
Man hat $\mathcal{V}_i = \mathrm{Kern}(p_i^{m_i}(\alpha)) = \mathrm{Bild}(q_i(\alpha))$ mit $q_i := \prod_{j\neq i}^{\ell} p_j^{m_j}$.

(3) Die Projektionen der Zerlegung sind gegeben durch $\pi_i = (a_i q_i)(\alpha)$ wobei $a_i \in K[x]$ mit
$$1 = a_1 q_1 + \ldots + a_\ell q_\ell$$
gegeben sind.

(4) Für die Dimension der Haupträume gilt
$$\mathrm{Dim}(\mathrm{Kern}(p_i^{m_i}(\alpha))) = c_i \,\mathrm{Grad}(p_i).$$

(5) Im Falle $m_i = 1$ ist $\mathrm{Kern}(p_i^{m_i}(\alpha))$ ein $K[x]/\langle p_i\rangle$-Vektorraum der Dimension c_i, und aus einer $K[x]/\langle p_i\rangle$-Basis konstruiert man leicht eine K-Basis, die für die Einschränkung von α auf $\mathrm{Kern}(p_i^{m_i}(\alpha))$ die Matrix $\mathrm{Diag}(M_{p_i}, \ldots, M_{p_i})$ liefert, wobei M_{p_i} die Begleitmatrix von p_i ist. (Wichtiger Spezialfall: $p_i(x) = x - a$ für ein $a \in K$. Dann ist $M_{p_i} = (a)$ und wir haben ($m_i = 1$ vorausgesetzt) eine Basis aus Eigenvektoren für den Hauptraum.)

Übung 7.1.1. Sei $\mu_\alpha = p^r$ für ein irreduzibles normiertes Polynom $p \in K[x]$. Zeige, dass der Algorithmus zur Berechnung des Minimalpolynoms einen Vektor $V \in \mathcal{V}$ als Nebenprodukt produziert, für dessen Minimalpolynom gilt: $\mu_{\alpha,V} = \mu_\alpha$.

Mit diesen Vorbemerkungen wollen wir zunächst einen kurzen Blick auf den Zentralisator eines Endomorphimus werfen.

Definition 7.1.4. Sei $\mathcal{V}$ ein endlich erzeugter K-Vektorraum der Dimension n und $\alpha \in \mathrm{End}(\mathcal{V})$. Dann heißt
$$C_{\mathrm{End}(\mathcal{V})}(\alpha) := \{\beta \in \mathrm{End}(\mathcal{V}) | \alpha \circ \beta = \beta \circ \alpha\} \leq \mathrm{End}(\mathcal{V})$$
der **Zentralisator** oder die **Zentralisatoralgebra** von α (in $\mathrm{End}(\mathcal{V})$).
Für $A \in K^{n\times n}$ ist die Zentralisatoralgebra von A definiert als $C_{K^{n\times n}}(A) := \{X \in K^{n\times n} \mid AX = XA\}$

Bemerkung 7.1.5. $C_{\mathrm{End}(\mathcal{V})}(\alpha)$ ist der Kern der linearen Abbildung
$$\mathrm{End}(\mathcal{V}) \to \mathrm{End}(\mathcal{V}) : \gamma \mapsto \alpha \circ \gamma - \gamma \circ \alpha$$
und somit ein Teilraum des K-Vektorraums $\mathrm{End}(\mathcal{V})$. Da $C_{\mathrm{End}(\mathcal{V})}(\alpha)$ auch abgeschlossen ist unter Komposition von Abbildungen, ist $C_{\mathrm{End}(\mathcal{V})}(\alpha)$ eine Teilalgebra von $\mathrm{End}(\mathcal{V})$.

Beispiel 7.1.6. Ist $A = \mathrm{Diag}(0,1) \in K^{2\times 2}$ so ist $C_{K^{2\times 2}}(A) = \{\mathrm{Diag}(a,b) \mid a,b \in K\} = K[A]$.
Für $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbb{F}_2^{2\times 2}$ ist $C_{\mathbb{F}_2^{2\times 2}}(A) = \mathbb{F}_2[A] \cong \mathbb{F}_4$. Betrachtet man die Blockdiagonalmatrix $\mathrm{Diag}(A,A) \in \mathbb{F}_2^{4\times 4}$ so ist ihr Zentralisator isomorph zu $\mathbb{F}_4^{2\times 2}$ (Übung).

Satz 7.1.7. *Sei $\alpha \in \mathrm{End}(\mathcal{V})$ etc. wie in Bemerkung 7.1.3. Setze $\mathcal{V}_i = \pi_i(\mathcal{V})$, $\alpha_i := \alpha_{|\mathcal{V}_i} : \mathcal{V}_i \to \mathcal{V}_i$.* Ende Vorl. 10

(1) Für $\beta \in C_{\mathrm{End}(\mathcal{V})}(\alpha)$ und $1 \le i \neq j \le \ell$ gilt $\pi_i \circ \beta \circ \pi_j = 0$, d.h. β respektiert die Zerlegung von $\mathcal{V}$ in seine Haupträume, und somit gilt insbesondere

$$\beta = \sum_{i=1}^{\ell} \underbrace{\pi_i \circ \beta \circ \pi_i}_{\in C_{\mathrm{End}(\mathcal{V}_i)}(\alpha_i)}$$

und

$$C_{\mathrm{End}(\mathcal{V})}(\alpha) = \bigtimes_{i=1}^{\ell} C_{\mathrm{End}(\mathcal{V}_i)}(\alpha_i)$$

(2) Sei $\mu_\alpha(x) = \chi_\alpha(x)$, also das Minimalpolynom und das charakteristische Polynom seien gleich. Dann ist

$$(\mathrm{id}_{\mathcal{V}}, \alpha, \alpha^2, \ldots, \alpha^{n-1})$$

mit $n = \mathrm{Dim}(\mathcal{V})$ eine K-Vektorraumbasis von $C_{\mathrm{End}(\mathcal{V})}(\alpha)$.

Beweis.

(1) Da $\beta \in C_{\mathrm{End}(\mathcal{V})}(\alpha)$ mit α vertauschbar ist, ist es auch mit jedem Polynom in α vertauschbar, insbesondere mit den π_i. Damit folgen die ersten Aussagen. Und mit $\mathrm{id}_{\mathcal{V}} = \sum_i \pi_i$ ist auch

$$C_{\mathrm{End}(\mathcal{V})}(\alpha) = \bigoplus_i \pi_i C_{\mathrm{End}(\mathcal{V})}(\alpha)$$

und $\pi_i C_{\mathrm{End}(\mathcal{V})}(\alpha) = C_{\mathrm{End}(\mathcal{V}_i)}(\alpha_i)$.

(2) Aus der Vorübung schließen wir, dass jeder Hauptraum $\mathcal{V}_i$ einen Vektor enthält, dessen Minimalpolynom gleich p^{m_i} ist. Wegen $\mu_\alpha(x) = \chi_\alpha(x)$ liefert die Summe dieser Vektoren uns einen Vektor $V \in \mathcal{V}$ mit mit Minimalpolynom $\mu_{\alpha,V}(x) = \mu_\alpha(x)$.
Unter den gegebenen Voraussetzungen ist

$$(V, \alpha(V), \ldots, \alpha^{n-1}(V))$$

eine Basis von $\mathcal{V}$. Sei nun $\beta \in C_{\mathrm{End}(\mathcal{V})}(\alpha)$ und

$$\beta(V) = V' = \sum a_i \alpha^i(V) \text{ für geeignete } a_i \in K.$$

Dann ist $\beta(\alpha(V)) = \alpha(\beta(V)) = \alpha(V')$ und allgemeiner $\beta(\alpha^j(V)) = \alpha^j(V')$, also

$$\beta(\alpha^j(V)) = (\sum_i a_i \alpha^i)(\alpha^j(V)).$$

Also hat β denselben Effekt auf unsere Basis wie $\sum a_i \alpha^i$, somit ist $\beta = \sum a_i \alpha^i$. □

Übung 7.1.2. Zeige: Im Falle $\mu_\alpha = \chi_\alpha$ ist $C_{\mathrm{End}(\mathcal{V})}(\alpha)$ als K-Algebra isomorph zu $K[x]/\langle \mu_\alpha \rangle$.

Übung 7.1.3. Man formuliere den letzten Satz und die zugehörigen Übungen in der Sprache der Matrizen. (Sehr wichtig!)

Die Situation aus Satz 7.1.7.(2) hat einen Namen:

Definition 7.1.8. Sei $\alpha \in \mathrm{End}(\mathcal{V})$. Jeder Vektor $V \in \mathcal{V}$ mit $\langle V \rangle_\alpha := K[\alpha](V) = \mathcal{V}$ heißt ein **zyklischer Vektor** von $\mathcal{V}$ (bezüglich α). Falls ein solcher existiert, heißt $\mathcal{V}$ ein **zyklischer Vektorraum** bezüglich α.

Bemerkung 7.1.9. $\mathcal{V}$ ist genau dann ein zyklischer Vektorraum bezüglich $\alpha \in \mathrm{End}(\mathcal{V})$ wenn $\mu_\alpha = \chi_\alpha$ gilt.

Beweis. Die Rückrichtung ist Teil des Beweis von Satz 7.1.7.(2). Die Hinrichtung ist trivial, denn für jeden zyklischen Vektor V hat $\mu_{\alpha,V}$ den Grad $\dim(\mathcal{V})$. □

Beispiel 7.1.10.

- Sei $A = \mathrm{Diag}(0,1) \in K^{2\times 2}$. Ist $K^{2\times 1}$ ein zyklischer Vektorraum bezüglich $\widetilde{A}$?
- Sei $B = \mathrm{Diag}(1,1) = I_2 \in K^{2\times 2}$. Ist $K^{2\times 1}$ ein zyklischer Vektorraum bezüglich $\widetilde{B}$?

7.2 Normalformen für Matrizen

7.2.a Die rationale kanonische Form

In diesem Abschnitt möchten wir den Struktursatz für Moduln über Hauptidealbereichen anwenden, um eine Normalform für Ähnlichkeitsklassen von Matrizen zu erhalten. Sei dazu K ein Körper.

Satz 7.2.1.

(1) Jede Matrix $A \in K^{n\times n}$ macht $K^{n\times 1}$ zu einem $K[x]$-Modul durch $p(x)V = p(A)V$ für alle $V \in K^{n\times 1}$, $p(x) \in K[x]$. Diesen $K[x]$-Modul bezeichnen wir mit M_A.

(2) Es ist $\mathrm{Ann}_{K[x]}(M_A) = \langle \mu_A \rangle \trianglelefteq K[x]$ das vom Minimalpolynom von A erzeugte Ideal.

(3) $\mathrm{End}_{K[x]}(M_A) \cong C_{K^{n\times n}}(A) = \{X \in K^{n\times n} \mid XA = AX\}$.

(4) Für $A, B \in K^{n\times n}$ gilt $M_A \cong M_B$ genau dann, wenn es ein $g \in \mathrm{GL}_n(K)$ gibt mit $A = g^{-1}Bg$, also genau dann wenn A und B ähnlich sind.

Beweis.

1. & 2. hatten wir schon früher bemerkt.

3. Der $K[x]$-Modul M_A wird durch Einschränkung zu einem K-Modul, also einem K-Vektorraum. Insbesondere sind alle $K[x]$-Modulhomomorphismen auch K-lineare Abbildungen und somit gegeben durch Multiplikation mit einer Matrix. Für $X \in K^{n\times n}$ ist die K-lineare Abbildung $\widetilde{X} \in \mathrm{End}_{K[x]}(M_A)$ genau dann, wenn $\widetilde{X}(xV) = x\widetilde{X}(V)$ für alle $V \in K^{n\times 1}$, also genau dann wenn $XAV = AXV$ für alle $V \in K^{n\times 1}$, d.h. $XA = AX$ und somit $X \in C_{K^{n\times n}}(A)$.

4. Sei $\widetilde{X} : M_A \to M_B$ ein Isomorphismus. Dann ist insbesondere die lineare Abbildung $\widetilde{X}$ bijektiv, also $X \in \mathrm{GL}_n(K)$. Weiter erfüllt $\widetilde{X}$ die Bedingung $\widetilde{X}(AV) = B\widetilde{X}(V)$ für alle $V \in K^{n\times 1}$, also $XA = BX$ und somit $A = X^{-1}BX$. □

Übung 7.2.1. Formulieren Sie obigen Satz und alle weiteren Ergebnisse dieses Abschnitts in der Sprache der Endomorphismen.

Definition 7.2.2. Sei $A \in K^{n\times n}$. Die **charakteristische Matrix** $\mathfrak{X}(A)$ ist definiert als $\mathfrak{X}(A) = xI_n - A \in K[x]^{n\times n}$.

Satz 7.2.3. *Sei $A \in K^{n\times n}$. Der Kern des $K[x]$-Modulepimorphismus*

$$f_A : \underbrace{K[x]^{n\times 1}}_{=\mathrm{Fr}_{K[x]}(\underline{n})} \twoheadrightarrow M_A, f_A(e_i) := e_i$$

ist der Teilmodul $S(\mathfrak{X}(A)) \leq K[x]^{n\times 1}$, der frei auf den Spalten von $\mathfrak{X}(A)$ ist (vom Rang n).

Beachte: e_i hat in diesem Kontext zwei verschiedene Bedeutungen: Als Argument von f_A lebt e_i in $K[x]^{n\times 1}$ und bezeichnet die i-te Einheitsspalte in diesem freien $K[x]$-Modul. Als Wert von f_A lebt e_i in M_A und bezeichnet die i-te Einheitsspalte in $K^{n\times 1}$.

Beweis. Für $p = (p_1, \ldots, p_n)^{tr} \in K[x]^{n\times 1}$ ist

$$f_A(p) = f_A(\sum_{i=1}^{n} p_i e_i) = \sum_{i=1}^{n} p_i(A) \cdot f_A(e_i) = \sum_{i=1}^{n} p_i(A) e_i.$$

Insbesondere haben wir für alle $1 \leq j \leq n$:

$$f_A(xe_j) = Ae_j = A_{-j}$$

und somit $f_A(xe_j - A_{-j}) = f_A(xe_j - \sum_{i=1}^{n} A_{ij}e_i) = A_{-j} - A_{-j} = 0$. Somit ist der von den Spalten $xe_j - A_{-j}$ der charakteristischen Matrix erzeugte Teilmodul von $K[x]^{n\times 1}$ enthalten im Kern von f_A.

Bezeichne $N_A := S(\mathfrak{X}(A))$ diesen Spaltenraum. Da N_A im Kern von f_A liegt, ist die Abbildung $\overline{f}_A : K[x]^{n\times 1}/N_A \to M_A$ wohldefiniert und wie die Ausgangsabbildung f_A ist auch $\overline{f}_A$ K-linear und surjektiv. Die K-Dimension von M_A ist n.
Behauptung: $(e_1 + N_A, \ldots, e_n + N_A)$ ist ein K-Erzeugendensystem von $K[x]^{n\times 1}/N_A$.

Dazu sei $p = (p_1, \ldots, p_n)^{tr} \in K[x]^{n\times 1}$. Wir zeigen durch Induktion über $\mathrm{Grad}(p) := \max\{\mathrm{Grad}(p_i) \mid 1 \leq i \leq n\}$, dass es ein $c \in K^{n\times 1}$ und ein $C \in N_A$ gibt mit $p = c + C$. Dies ist klar, falls $\mathrm{Grad}(p) = 0$, da dann schon $p = c \in K^{n\times 1}$. Ansonsten dividiere alle p_i mit $\mathrm{Grad}(p_i) = \mathrm{Grad}(p)$ sukzessive mit Rest durch $(x - A_{ii})$, also $p_i = q_i(x - A_{ii}) + r_i$ und ersetze p durch $p - \underbrace{q_i\mathfrak{X}(A)_{-i}}_{\in N_A}$. Da $\mathrm{Grad}(q_i) = \mathrm{Grad}(p_i) - 1$ ist, verringert sich der Grad von p nach der sukzessiven Ersetzung um mindestens 1. Nun können wir die Induktionsannahme anwenden.
Also ist die K-Dimension von $K[x]^{n\times 1}/N_A$ höchstens n und somit $\overline{f}_A$ auch injektiv.
Dass N_A frei auf den Spalten von $\mathfrak{X}(A)$ ist, folgt da $\det(\mathfrak{X}(A)) = \chi_A \neq 0$. □

Ende Vorl. 11

Folgerung 7.2.4. *Als $K[x]$-Modul ist*

$$M_A \cong K[x]^{n\times 1}/\operatorname{Kern} f_A = K[x]^{n\times 1}/S(\mathfrak{X}(A)).$$

Nach dem Struktursatz 6.4.6 gibt es $\mathfrak{g}, \mathfrak{h} \in \mathrm{GL}_n(K[x])$ mit

$$\mathfrak{g}\mathfrak{X}(A)\mathfrak{h} = \mathrm{Diag}(f_1(x), \ldots, f_n(x)),$$

so dass $f_i(x) \in K[x]$ normiert, $f_1(x) \mid f_2(x) \mid \ldots \mid f_n(x)$. Ist $d_i := \mathrm{Grad}(f_i)$ und $s = \min\{i \in \underline{n} \mid d_i > 0\}$ so ist

$$\chi_A = \det(\mathfrak{X}(A)) = \prod_{i=1}^{n} f_i(x) = \prod_{i=s}^{n} f_i(x),$$

$$\mu_A = f_n(x)$$

und

$$\begin{aligned} M_A &\cong_{K[x]} K[x]/\langle f_1(x)\rangle \oplus \ldots \oplus K[x]/\langle f_n(x)\rangle \\ &\cong_{K[x]} K[x]/\langle f_s(x)\rangle \oplus \ldots \oplus K[x]/\langle f_n(x)\rangle \cong M_{\mathrm{RKF}(A)} \end{aligned}$$

mit

$$\begin{aligned} \mathrm{RKF}(A) &= \mathrm{Diag}(M_{f_1}, \ldots, M_{f_n}) \\ &= \mathrm{Diag}(M_{f_s}, \ldots, M_{f_n}) \end{aligned}$$

wobei M_{f_i} die Begleitmatrix von f_i bezeichnet (M_{f_i} ist leer, falls $d_i = 0$ ist). Die $f_i(x)$ sind nach Satz 6.4.10 durch A eindeutig bestimmt (als Elementarteiler von $\mathfrak{X}(A)$). Die zu A ähnliche Blockdiagonalmatrix $\mathrm{RKF}(A)$ heißt die **rationale kanonische Form** *oder auch* **Frobenius-Normalform** *von $A \in K^{n\times n}$.*

Bemerkung 7.2.5. Sei $\mathrm{RKF}(A) := \mathrm{Diag}(M_{f_s}, \ldots, M_{f_n})$. Dann ist M_A die direkte Summe von $n - s + 1$ zyklischen $K[x]$-Moduln $K[x]/\langle f_i\rangle$. Jede andere solche Zerlegung von M_A hat mindestens ebensoviele zyklische Summanden.

Bemerkung 7.2.6. *Achtung:* Im Gegensatz zur Hauptraumzerlegung ist die Zerlegung in Bemerkung 7.2.5 nur bis auf Isomorphie eindeutig: Ist $\alpha = \widetilde{A}$ und sind B, B' Basen von $\mathcal{V} = K^{n\times 1}$ mit

$$ {}^{B}\alpha^{B} = \mathrm{Diag}(M_{f_s}, \ldots, M_{f_n}) = {}^{B'}\alpha^{B'}, $$

so ist der Endomorphismus $\beta \in \mathrm{End}(\mathcal{V})$ definiert durch $\beta(B_i) := B'_i$ für alle $1 \le i \le n = \mathrm{Dim}(\mathcal{V})$ eine Einheit im Zentralisator von α:

$$\beta \in C_{\mathrm{End}(\mathcal{V})}(\alpha)^* = C_{\mathrm{End}(\mathcal{V})}(\alpha) \cap \mathrm{GL}(\mathcal{V}) =: \mathrm{Aut}_\alpha(\mathcal{V}).$$

Folgerung 7.2.7. *Seien $A, B \in K^{n\times n}$. Äquivalent sind:*

(1) A und B sind ähnlich.

(2) $\mathfrak{X}(A)$ und $\mathfrak{X}(B)$ sind ähnlich.

(3) $\mathfrak{X}(A)$ und $\mathfrak{X}(B)$ sind äquivalent über $K[x]$.

(4) M_A und M_B sind isomorphe $K[x]$-Moduln.

(5) A und B haben dieselbe rationale kanonische Form.

Durch eine Kombination von Hauptraumzerlegung und rationaler kanonischer Form erhält man die sogenannte **primäre rationale Form** einer Matrix A, mit der man M_A in die maximal mögliche Anzahl nicht-trivialer zyklischer $K[X]$-Moduln zerlegt. Diese hat den Vorteil, dass die Blockdiagonalmatrizen kleiner sind als bei der rationalen kanonischen Form.

Bemerkung 7.2.8. Sei $A \in K^{n\times n}$ mit Minimalpolynom $\mu_A = \prod_{i=1}^{\ell} p_i^{m_i}$, charakterisitischem Polynom $\chi_A = \prod_{i=1}^{\ell} p_i^{c_i}$ und $\mathcal{V}_i := \mathrm{Kern}(p_i^{m_i}(A))$ der p_i-Hauptraum. Dann ist $K^{n\times 1} = \bigoplus_{i=1}^{\ell} \mathcal{V}_i$ eine $\widetilde{A}$-invariante Zerlegung. Bezüglich einer an diese Zerlegung angepassten Basis hat $\widetilde{A}$ also eine Matrix $\mathrm{Diag}(A_1, \ldots, A_\ell)$ in Blockdiagonalgestalt. Ist $\mathrm{RKF}(A_i) = \mathrm{Diag}\left(M_{p_i^{a_{i1}}}, \ldots, M_{p_i^{a_{is_i}}}\right)$ die rationale kanonische Form von A_i (also $0 < a_{i1} \le a_{i2} \le \ldots \le a_{is_i}$, $c_i = a_{i1} + \ldots + a_{is_i}$, $a_{is_i} = m_i$) so ist A ähnlich zu

$$\mathrm{PRF}(A) = \mathrm{Diag}(\mathrm{RKF}(A_1), \ldots, \mathrm{RKF}(A_\ell)) = \mathrm{Diag}\left(M_{p_1^{a_{11}}}, \ldots, M_{p_\ell^{a_{\ell s_\ell}}}\right).$$

$\mathrm{PRF}(A)$ heißt die **primäre rationale Form** oder **Weierstraß Form** von A.

Übung 7.2.2. Die a_{ij}'s lassen sich aus der Primfaktorzerlegung der Elementarteiler von $\mathfrak{X}(A)$ bestimmen.

Beispiel 7.2.9. Sei $K := \mathbb{Q}$ und

$$A := \begin{pmatrix} -6 & 6 & 0 & 6 \\ -4 & 6 & -1 & 3 \\ -6 & 12 & -3 & 3 \\ -6 & 12 & -3 & 3 \end{pmatrix} \in \mathbb{Q}^{4\times 4}$$

Dann ist

$$\mathfrak{X}(A) = \begin{pmatrix} x+6 & -6 & 0 & -6 \\ 4 & x-6 & 1 & -3 \\ 6 & -12 & x+3 & -3 \\ 6 & -12 & 3 & x-3 \end{pmatrix} \in \mathbb{Q}[x]^{4\times 4}$$

und $\mathfrak{g}\mathfrak{X}(A)\mathfrak{h} = \mathrm{Diag}(1, 1, x, x^3)$, wobei

$$\mathfrak{g} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1/24x + 1/4 & -3/4 & 0 & 1/4 \\ 1/6x^2 - x - 4 & x + 12 & -4 & x \end{pmatrix},\ \mathfrak{h} = \begin{pmatrix} 0 & 0 & 0 & 6 \\ 0 & 0 & -1 & 1/4x + 3 \\ -1/2 & 1 & x - 3 & -1/4x^2 + 3/4x + 3 \\ -1/6 & 0 & 1 & 3/4x + 3 \end{pmatrix}.$$

Also erhält man

$$\mathrm{RKF}(A) = \mathrm{PRF}(A) = \mathrm{Diag}(M_x, M_{x^3}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Beispiel 7.2.10. Sei

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{Q}^{3\times 3}.$$

Dann findet man $\mathfrak{g}\mathfrak{X}(A)\mathfrak{h} = \mathrm{Diag}(1, 1, x^3 + x^2 - x - 1)$ wobei

$$\mathfrak{g} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ x & x^2 & 1 \end{pmatrix},\ \mathfrak{h} = \begin{pmatrix} 0 & -1 & x+1 \\ 0 & 0 & 1 \\ -1 & -x & x^2 + x - 1 \end{pmatrix}$$

Also ist $\mu_A = \chi_A = x^3 + x^2 - x - 1 = (x+1)^2(x-1)$, $\mathrm{RKF}(A) = M_{\mu_A}$ und $\mathrm{PRF}(A) = \mathrm{Diag}(M_{x-1}, M_{(x+1)^2}) = \mathrm{Diag}((1), \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix})$.

7.2.b Trennende Invarianten

Satz 7.2.11. *Sei $\alpha \in \mathrm{End}(\mathcal{V})$ mit $\mu_\alpha = p^m$ mit $p \in K[x]$ irreduzibel. Setze $\nu := p(\alpha)$, $d := \mathrm{Grad}(p)$. Dann sind folgende Aussagen äquivalent:*

(1) $\mathcal{V}$ ist zyklisch bezüglich α.

(2) $\mathrm{Dim}_K \mathrm{Kern}(\nu) = d$.

(3) $\mathrm{Rang}(\nu) = \mathrm{Dim}_K(\mathcal{V}) - d$.

(4) $\mu_\alpha = \chi_\alpha$.

(5) Sämtliche α-invarianten Teilräume von $\mathcal{V}$ sind gegeben durch

$$\langle \nu^i(\mathcal{V})\rangle_\alpha = \langle \nu^i(\mathcal{V})\rangle = \mathrm{Bild}(\nu^i)$$

für $i = 0, 1, \ldots m$.

Beweis. Zunächst eine allgemeine Vorbemerkung, die nur die gemeinsame Voraussetzung benutzt: Setze $\mathcal{V}_i := \mathrm{Bild}(\nu^i)$. Nach Definition des Minimalpolynoms ist $\mathcal{V}_{m-1} \neq \{0\}$ und $\mathcal{V}_m = \{0\}$. Da $\nu = p(\alpha)$ mit α kommutiert, sind die $\mathcal{V}_i$'s allesamt α-invariant, d.h. $\alpha(\mathcal{V}_i) \subseteq \mathcal{V}_i$. Es ist

$$\mathcal{V} = \mathcal{V}_0 > \mathcal{V}_1 > \ldots > \mathcal{V}_{m-1} > \mathcal{V}_m = \{0\}, \qquad (\star)$$

wobei die Faktoren $\mathcal{V}_i/\mathcal{V}_{i+1} = \nu(\mathcal{V}_{i-1}/\mathcal{V}_i)$ epimorphe Bilder voneinander sind. Insbesondere ist

$$\mathrm{Dim}(\mathcal{V}_{m-1}) \leq \ldots \leq \mathrm{Dim}(\mathcal{V}_i) - \mathrm{Dim}(\mathcal{V}_{i+1}) \leq \mathrm{Dim}(\mathcal{V}_{i-1}) - \mathrm{Dim}(\mathcal{V}_i) \leq \ldots \leq \underbrace{\mathrm{Dim}(\mathcal{V}_0) - \mathrm{Dim}(\mathcal{V}_1)}_{=\mathrm{Dim}(\mathrm{Kern}(\nu))}.$$

Das Minimalpolynom des von α auf $\mathcal{V}_i/\mathcal{V}_{i+1}$ induzierten Endomorphismus α_i ist $\mu_{\alpha_i} = p$. Insbesondere ist nach Satz 5.5.7 (im LA I Skript) die Dimension $\mathrm{Dim}(\mathcal{V}_i/\mathcal{V}_{i+1})$ ein Vielfaches von d.

(2) $\Leftrightarrow$ (3) folgt aus dem Homomorphiesatz.

(1) $\Leftrightarrow$ (4) ist Bemerkung 7.1.9.

(4) $\Rightarrow$ (2) Ist $\mu_\alpha = \chi_\alpha$, so ist $dm = \mathrm{Grad}(\mu_\alpha) = \mathrm{Grad}(\chi_\alpha) = n$ und die echt absteigende Kette von Teilräumen oben hat Länge n/d. Damit muss aber $\mathrm{Dim}(\mathcal{V}_i) - \mathrm{Dim}(\mathcal{V}_{i+1}) = d$ sein für alle i, also auch $\mathrm{Dim}(\mathrm{Kern}(\nu)) = \mathrm{Dim}(\mathcal{V}) - \mathrm{Dim}(\mathrm{Bild}(\nu)) = \mathrm{Dim}(\mathcal{V}_0) - \mathrm{Dim}(\mathcal{V}_1) = d$.

(2) $\Rightarrow$ (4) Ist $\mathrm{Dim}(\mathrm{Kern}(\nu)) = d$, so folgt $\mathrm{Dim}(\mathcal{V}_i) = \mathrm{Dim}(\mathcal{V}_{i+1}) + d$ für alle i und daher $m = n/d$.

(1) $\Rightarrow$ (5) Ist $\mathcal{W} \leq \mathcal{V}$ ein α-invarianter Teilraum, so gibt es ein größtes i mit $\mathcal{W} \leq \mathcal{V}_i$. Aber jedes $W \in \mathcal{W} \setminus \mathcal{V}_{i+1}$ erfüllt bereits $\langle W\rangle_\alpha = \mathcal{V}_i$, da auch $\mathcal{V}_i$ ein zyklischer Modul für $\alpha_{|\mathcal{V}_i}$ und $\mathcal{V}_i/\mathcal{V}_{i+1} \cong K[x]/\langle p\rangle$ ein einfacher Modul ist, sprich, nur die trivialen Teilmoduln hat.

(5) $\Rightarrow$ (1) Klar, da dann alle Vektoren in $\mathcal{V} \setminus \mathcal{V}_1$ (und nur diese) demnach zyklisch sein müssen. □

Definition 7.2.12.

(1) Eine **Partition der natürlichen Zahl** $c \in \mathbb{N}_{>0}$ ist ein k-Tupel $a = (a_1, \ldots, a_k) \in \mathbb{N}_{>0}^k$ für ein $k \in \mathbb{N}_{>0}$ mit $a_1 \geq a_2 \geq \ldots \geq a_k$ und $a_1 + a_2 + \cdots + a_k = c$.

(2) Ist $a = (a_1, \ldots, a_k) \in \mathbb{N}_{>0}^k$ eine Partition von $c \in \mathbb{N}_{>0}$, so ist die **konjugierte Partition** a' von a definiert durch $a'_i := |\{j | a_j \geq i\}|$.

Man visualisiert üblicherweise die Partition durch Kästchen, die man linksbündig in Zeilen untereinander anordnet mit a_i Kästchen in der i-ten Zeile. Dies nennt man das **Young-Diagramm** der Partition. Die YOUNG-Diagramme von a und a' sind transponiert zueinander.

Ende Vorl. 12

Definition 7.2.13. Sei $\mathcal{V}$ ein K-Vektorraum und $\alpha \in \mathrm{End}(\mathcal{V})$ mit $\chi_\alpha = p^c$, $p \in K[x]$ irreduzibel. Dann gibt es eine K-Basis B von $\mathcal{V}$ mit

$$^B\alpha^B = \mathrm{Diag}(M_{p^{a_1}}, \ldots, M_{p^{a_k}})$$

entsprechend dem $K[x]$-Modulisomorphismus

$$\mathcal{V} \cong_{K[x]} K[x]/\langle p^{a_1}\rangle \oplus \cdots \oplus K[x]/\langle p^{a_k}\rangle$$

für eine eindeutig durch α definierte Partition $(a_k, \ldots, a_1)$ von c. Diese Partition heißt die durch α definierte Partition.

Folgerung 7.2.14. *Seien α und p wie in Definition 7.2.13, $d := \mathrm{Grad}(p)$ und $a := (a_k, \ldots, a_1)$ die durch α definierte Partition. Sei $\mathcal{V}_i := \mathrm{Bild}(p(\alpha)^i)$ für $i = 0, \ldots, a_k$. Dann gilt*

$$\mathcal{V} = \mathcal{V}_0 > \mathcal{V}_1 > \ldots > \mathcal{V}_{a_k-1} > \mathcal{V}_{a_k} = \{0\}$$

und $\dim(\mathcal{V}_i/\mathcal{V}_{i+1}) = a'_i d$, wobei a' die zu a konjugierte Partition ist.

Beweis. Wende Satz 7.2.11 auf jeden zyklischen Summanden an. □

Beispiel 7.2.15. Sie $p \in K[x]$ normiert, irreduzibel von Grad d und $A \in K^{5d\times 5d}$ mit $\mu_A = p^3$. Dann ist $\chi_A = p^5$ und man hat 2 Möglichkeiten für die Ähnlichkeitsklasse von A:

$$A \sim \mathrm{Diag}(M_p, M_p, M_{p^3}) \text{ oder } A \sim \mathrm{Diag}(M_{p^2}, M_{p^3})$$

entsprechend den $K[x]$-Modulisomorphismen

$$M_A \cong K[x]/\langle p\rangle \oplus K[x]/\langle p\rangle \oplus K[x]/\langle p^3\rangle \text{ oder } M_A \cong K[x]/\langle p^2\rangle \oplus K[x]/\langle p^3\rangle.$$

Im ersten Fall ist $\dim(\mathrm{Kern}(p(A)) = 3d$ und im zweiten Fall gleich $2d$. Man kann also entscheiden, zu welcher Ähnlichkeitsklasse die Matrix A gehört, indem man nur den Rang von $\nu = p(A)$ berechnet.

Es stellt sich abschließend die Frage nach trennenden Invarianten für die Konjugationsoperation.

Satz 7.2.16. *Zwei Endomorphismen $\alpha, \beta \in \mathrm{End}(\mathcal{V})$ sind genau dann unter $\mathrm{GL}(\mathcal{V})$ konjugiert, wenn gilt*

(1) die Minimalpolynome sind gleich: $\mu_\alpha(x) = \mu_\beta(x)$ und

(2) für jeden normierten irreduziblen Teiler $p \in K[x]$ des Minimalpolynoms sind die Partitionen des p-Hauptraumes von $(\mathcal{V}, \alpha)$ und von $(\mathcal{V}, \beta)$ gleich.

In anderen Worten: Das Minimalpolynom zusammen mit den Partitionen bilden ein System trennender Invarianten für die Ähnlichkeitsklassen.

Beispiel 7.2.17. Ähnlichkeitsklassen in $\mathbb{C}^{3\times 3}$:

$\chi_A(x)$	$\mu_A(x)$	Partitionen	Vertreter
$(x-a)^3$	$x-a$	$(1,1,1)$	$\mathrm{Diag}(a,a,a)$
	$(x-a)^2$	$(2,1)$	$\mathrm{Diag}(a, M_{(x-a)^2})$
	$(x-a)^3$	(3)	$M_{(x-a)^3}$
$(x-a)^2(x-b)$	$(x-a)(x-b)$	$(1,1),(1)$	$\mathrm{Diag}(a,a,b)$
	$(x-a)^2(x-b)$	$(2),(1)$	$\mathrm{Diag}(M_{(x-a)^2}, b)$
$(x-a)(x-b)(x-c)$	$(x-a)(x-b)(x-c)$	$(1),(1),(1)$	$\mathrm{Diag}(a,b,c)$

Beispiel 7.2.18. Ähnlichkeitsklassen in $\mathbb{C}^{4\times 4}$:

$\chi_A(x)$	$\mu_A(x)$	Partitionen	Vertreter
$(x-a)^4$	$x-a$	$(1,1,1,1)$	$\mathrm{Diag}(a,a,a,a)$
	$(x-a)^2$	$(2,1,1)$	$\mathrm{Diag}(a,a,M_{(x-a)^2})$
	$(x-a)^2$	$(2,2)$	$\mathrm{Diag}(M_{(x-a)^2},M_{(x-a)^2})$
	$(x-a)^3$	$(3,1)$	$\mathrm{Diag}(a,M_{(x-a)^3})$
	$(x-a)^4$	(4)	$M_{(x-a)^4}$
$(x-a)^3(x-b)$	$(x-a)(x-b)$	$(1,1,1),(1)$	$\mathrm{Diag}(a,a,a,b)$
	$(x-a)^2(x-b)$	$(2,1),(1)$	$\mathrm{Diag}(a,M_{(x-a)^2},a,b)$
	$(x-a)^3(x-b)$	$(3),(1)$	$\mathrm{Diag}(M_{(x-a)^3},b)$
$(x-a)^2(x-b)^2$	$(x-a)(x-b)$	$(1,1),(1,1)$	$\mathrm{Diag}(a,a,b,b)$
	$(x-a)^2(x-b)$	$(2),(1,1)$	$\mathrm{Diag}(M_{(x-a)^2},b,b)$
	$(x-a)^2(x-b)^2$	$(2),(2)$	$\mathrm{Diag}(M_{(x-a)^2},M_{(x-b)^2})$
$(x-a)^2(x-b)(x-c)$	$(x-a)(x-b)(x-c)$	$(1,1),(1),(1)$	$\mathrm{Diag}(a,a,b,c)$
	$(x-a)^2(x-b)(x-c)$	$(2),(1),(1)$	$\mathrm{Diag}(M_{(x-a)^2},b,c)$
$\prod_{w=a,b,c,d}(x-w)$	$\prod_{w=a,b,c,d}(x-w)$	$(1),(1),(1),(1)$	$\mathrm{Diag}(a,b,c,d)$

Übung 7.2.3. Gib ein Vertretersystem aller Konjugiertenklassen von Endomorphismen von $\mathbb{F}_2^{3\times 1}$ und von $\mathbb{R}^{3\times 1}$ an.

7.2.c Die JORDAN Normalform

Aus der primären rationalen Form erhält man durch eine etwas andere Basiswahl leicht die JORDAN Normalform. Dazu genügt es, die zyklischen Moduln innerhalb eines Hauptraums zu behandeln.

Satz 7.2.19. *Sei $\alpha \in \mathrm{End}(\mathcal{V})$ mit $\mu_\alpha = \chi_\alpha = p^m$ mit $p \in K[x]$ irreduzibel. Setze $\nu := p(\alpha)$, $d := \mathrm{Grad}(p)$.*
Jedes $V \in \mathcal{V} \setminus \nu(\mathcal{V})$ liefert eine Basis

$$B := (\underbrace{V, \alpha(V), \ldots, \alpha^{d-1}(V)}, \underbrace{\nu(V), \alpha(\nu(V)), \ldots \alpha^{d-1}(\nu(V))}, \ldots, \underbrace{\nu^{m-1}(V), \ldots, \alpha^{d-1}(\nu^{m-1}(V))})$$

von $\mathcal{V}$, so dass die Matrix von α bzgl. B gegeben ist durch

$${}^B\alpha^B = J_m(p) := \begin{pmatrix} M_p & 0 & 0 & \ldots & 0 & 0 \\ N_d & M_p & 0 & \ldots & 0 & 0 \\ 0 & N_d & M_p & \ldots & 0 & 0 \\ 0 & 0 & N_d & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & N_d & M_p \end{pmatrix}$$

wobei $N_d = \begin{pmatrix} 0 & \ldots & 0 & 1 \\ 0 & \ldots & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & \ldots & 0 \end{pmatrix} \in K^{d\times d}$.

Beweis. Nachrechnen. □

Folgerung 7.2.20. *Sei $A \in K^{n\times n}$ mit $\mathrm{PRF}(A) = \mathrm{Diag}(M_{p_1^{a_{11}}}, \ldots, M_{p_\ell^{a_{\ell s_\ell}}})$. Dann ist A ähnlich zu*

$$\mathrm{JNF}(A) = \mathrm{Diag}(J_{a_{11}}(p_1), \ldots, J_{a_{\ell s_\ell}}(p_\ell)).$$

$\mathrm{JNF}(A)$ *heißt die* **Jordan-Normalform** *von A.*

Beispiel 7.2.21. Im Beispiel 7.2.9 ist $\mathrm{RKF}(A) = \mathrm{PRF}(A) = \mathrm{JNF}(A)$. Im Beispiel 7.2.10 erhält man

$$\mathrm{JNF}(A) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & -1 & 0 \\ 0 & 1 & -1 \end{array}\right).$$

7.2.d Transformationsmatrizen

Bemerkung 7.2.22. (ohne Beweis) Sei $S = \mathfrak{g}\mathfrak{X}(A)\mathfrak{h} = \mathrm{Diag}(f_1, \ldots, f_n)$ die SMITH-Form von $\mathfrak{X}(A) \in K[x]^{n\times n}$, wobei die f_i normiert seien. Seien $f_1 = \ldots = f_{s-1} = 1$ und $\mathrm{Grad}(f_i) = d_i \geq 1$ für alle $i \geq s$. Dann gilt $d_s + \ldots + d_n = n$. Für $i \geq s$ und $1 \leq j \leq n$ sei

$$\mathfrak{g}_{ij} \equiv c_{i,j,0} + c_{i,j,1}x + \ldots + c_{i,j,d_i-1}x^{d_i-1} \pmod{f_i}.$$

Setze

$$P_j^{(i)} := \begin{bmatrix} c_{i,j,0} \\ \vdots \\ c_{i,j,d_i-1} \end{bmatrix} \in K^{d_i\times 1}$$

und

$$P_j := \begin{bmatrix} P_j^{(s)} \\ \vdots \\ P_j^{(n)} \end{bmatrix} \in K^{n\times 1}.$$

Dann ist die Matrix $P = [P_1, \ldots, P_n] \in K^{n\times n}$ invertierbar und es gilt $PAP^{-1} = \mathrm{RKF}(A)$, wobei $\mathrm{RKF}(A)$ die rationale kanonische Form von A ist.

Beispiel 7.2.23. Ist A wie in Beispiel 7.2.10, so kann $\mathfrak{g} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ x & x^2 & 1 \end{pmatrix}$ gewählt werden und es ergibt sich gemäß obiger Vorschrift

$$P := \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

und erhält

$$PAP^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} = \mathrm{RKF}(A)$$

wie gewünscht.
Die Berechnung der SMITH-Form von $\mathfrak{X}(A)$ ist zu aufwendig. Eine Transformationsmatrix P erhält man einfacher durch direktes Rechnen mit Matrizen über K: Da $\mu_A = \chi_A$ gilt, ist $\mathbb{Q}^{3\times 1}$ ein zyklischer Modul. Beginnt man mit $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $Ae_1 = e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $A^2e_1 = Ae_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, so erhält man direkt eine Basis $B = (e_1, e_2, Ae_2)$ mit ${}^B\widetilde{A}^B = M_{\mu_A}$, in Matrizen $T := \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$ erfüllt $T^{-1}AT = M_{\mu_A}$.

Beispiel 7.2.24. Sei nun A wie in Beispiel 7.2.9, also $\mathfrak{g}\mathfrak{X}(A)\mathfrak{h} = \mathrm{Diag}(1, 1, x, x^3)$, wobei

$$\mathfrak{g} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1/24x + 1/4 & -3/4 & 0 & 1/4 \\ 1/6x^2 - x - 4 & x + 12 & -4 & x \end{pmatrix}.$$

Dann erhält man

$$P = \begin{pmatrix} 1/4 & -3/4 & 0 & 1/4 \\ -4 & 12 & -4 & 0 \\ -1 & 1 & 0 & 1 \\ 1/6 & 0 & 0 & 0 \end{pmatrix}$$

und berechnet $PAP^{-1} = \mathrm{Diag}(M_x, M_{x^3})$.

Ende Vorl. 13

Diese Methode ist aufwendig, da es i.a. nicht so leicht ist, die SMITH-Form der charakteristischen Matrix zu bestimmen. Dazu werden Rechnungen im Polynomring benötigt. Im folgenden wollen wir uns überlegen, wie wir eine geeignete Basis finden, so dass ${}^B\alpha^B$ in Normalform ist, wobei wir ab jetzt mit der JORDAN-Normalform arbeiten werden.

Beispiel 7.2.25. Seien $K := \mathbb{Q}$ und A wie in Beispiel 7.2.9, also

$$A := \begin{pmatrix} -6 & 6 & 0 & 6 \\ -4 & 6 & -1 & 3 \\ -6 & 12 & -3 & 3 \\ -6 & 12 & -3 & 3 \end{pmatrix} \in \mathbb{Q}^{4\times 4}$$

mit Minimalpolynom $\mu_A(x) = x^3$. Der erste Standardbasisvektor E_1 hat x^3 als Minimalpolynom μ_{A,E_1}:

$$(E_1, AE_1, A^2E_1) = \begin{pmatrix} 1 & -6 & -24 \\ 0 & -4 & -12 \\ 0 & -6 & -12 \\ 0 & -6 & -12 \end{pmatrix},$$

Es ist $\langle E_1\rangle_A = \langle E_1, E_2, E_3 + E_4\rangle$ (etwa mit Spalten-Gauß nachrechnen). Dieser Raum enthält also nicht E_3. Es ist $AE_3 = AE_1 - \frac{1}{4}A^2E_1$ also setzen wir $F := 2(E_1 - \frac{1}{4}AE_1 - E_3) = (5, 2, 1, 3)^{tr}$ damit $AF = 0$ wird. Dann ist (F, E_1, AE_1, A^2E_1) eine Basis von $\mathbb{Q}^{4\times 1}$, bezüglich der $\widetilde{A}$ die Matrix $\mathrm{Diag}(M_x, M_{x^3})$ bekommt.

Unsere Aufgabe ist es, diese Normierung auf den Fall von mehr als zwei Summanden zu übertragen. Der Einfachheit halber nehmen wir an, dass $\mu_\alpha = (x-a)^m$ gilt, also nur ein Hauptraum vorliegt (die Zerlegung in Haupträume haben wir also schon erledigt) und (nur um das Verfahren klarer zu machen) dass der irreduzible Faktor Grad 1 hat.

Bemerkung 7.2.26. Algorithmus zur Bestimmung der Jordan-Normalform, der mit Kernen arbeitet
Sei $\alpha \in \mathrm{End}(\mathcal{V})$ mit $\mu_\alpha = p^m = (x-a)^m$ und setze $\nu := p(\alpha) = \alpha - a\,\mathrm{id}_\mathcal{V}$. Sei $\mathcal{W}_i := \mathrm{Kern}(\nu^i)$. Dann ist

$$\mathcal{W}_0 = \{0\} < \underbrace{\mathcal{W}_1}_{=E_\alpha(a)} < \ldots < \mathcal{W}_{m-1} < \mathcal{W}_m = \mathcal{V}.$$

Vorbemerkungen: Für die Elemente $V \in \mathcal{W}_j \setminus \mathcal{W}_{j-1}$ gilt $\nu^j(V) = 0$, aber $\nu^{j-1}(V) \neq 0$. Außerdem ist $\nu : \mathcal{W}_j/\mathcal{W}_{j-1} \to \mathcal{W}_{j-1}/\mathcal{W}_{j-2}$ injektiv, da $\nu^{-1}(\mathcal{W}_{j-2}) \leq \mathcal{W}_{j-1}$ ist.

(1) Ergänze eine Basis von $\mathcal{W}_{m-1}$ durch $V_1, \ldots, V_k$ zu einer Basis von $\mathcal{W}_m = \mathcal{V}$. Dann bilden die Restklassen $(V_1 + \mathcal{W}_{m-1}, \ldots, V_k + \mathcal{W}_{m-1})$ eine Basis von $\mathcal{W}_m/\mathcal{W}_{m-1}$. Da $\nu : \mathcal{W}_m/\mathcal{W}_{m-1} \to \mathcal{W}_{m-1}/\mathcal{W}_{m-2}$ injektiv ist, sind auch $(\nu(V_1)+\mathcal{W}_{m-2}, \ldots, \nu(V_k)+\mathcal{W}_{m-2})$ linear unabhängig und man erhält induktiv, dass

$$B_m := (V_1, \nu(V_1), \ldots, \nu^{m-1}(V_1), V_2, \nu(V_2), \ldots, \nu^{m-1}(V_2), \ldots, V_k, \nu(V_k), \ldots, \nu^{m-1}(V_k))$$

eine Basis eines α-invarianten Teilraums ist mit ${}^{B_m}\alpha^{B_m} = \mathrm{Diag}(\underbrace{J_m(a), \ldots, J_m(a)}_{k})$.

(2) Ergänze eine Basis von

$$\mathcal{W}_{m-2} \oplus \langle \nu(V_1), \dots, \nu(V_k) \rangle \leq \mathcal{W}_{m-1}$$

durch Vektoren $W_1, \dots, W_\ell$ zu einer Basis von $\mathcal{W}_{m-1}$. Die Zahl ℓ kann 0 sein.

$$B_{m-1} := (W_1, \nu(W_1), \dots, \nu^{m-2}(W_1), W_2, \dots, \nu^{m-2}(W_2), \dots, W_\ell, \dots, \nu^{m-2}(W_\ell))$$

ist dann eine Basis eines α-invarianten Teilraums mit

$${}^{B_m,B_{m-1}}\alpha^{B_m,B_{m-1}} = \mathrm{Diag}(\underbrace{J_m(a), \dots, J_m(a)}_{k}, \underbrace{J_{m-1}(a), \dots, J_{m-1}(a)}_{\ell}).$$

(3) Wiederhole (2) mit $m-1$ anstelle von m, dann mit $m-2$ etc..

Beispiel 7.2.27.

$$A := \begin{pmatrix} 1 & 2 & 0 & 0 \\ 1 & 2 & 0 & 2 \\ 1 & 2 & 2 & 1 \\ 2 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{F}_3^{4\times 4}$$

hat Minimalpolynom $\mu_A = (x-2)^3$. Es gilt

$$(A - 2I_4)^2 = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 2 & 0 & 2 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

Der Rang ist 1, wir werden also den ersten JORDAN-Block aus dem zweiten Standardbasisvektor $V = E_2$ bekommen. Der Eigenvektor $(A-2I_4)^2V$ wird offenbar durch den dritten Standardbasisvektor $W = E_3$ zur Basis von $\mathcal{W}_1 := \mathrm{Kern}(A - 2I_4) =: E_A(2)$ ergänzt. Also ist unsere neue Basis $B = (V, (A-2I_4)V, (A-2I_4)^2V, W)$ und die transformierte Matrix ist

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathbb{F}_3^{4\times 4}.$$

7.2.e Eine Anwendung: lineare Differentialgleichungssysteme.

In diesem Abschnitt wollen wir eine Anwendung der Jordan-Normalform sehen. Im Wesentlichen geht es um die Berechnung der Matrix-Exponentialfunktion.

Definition 7.2.28. Sei $A \in \mathbb{R}^{n\times n}$. Sind $u_i(t)$ reelle Funktionen so setze

$$u(t) := \begin{pmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{pmatrix} \text{ und } u'(t) := \begin{pmatrix} u_1'(t) \\ \vdots \\ u_n'(t) \end{pmatrix}$$

Dann heißt

$$u'(t) = Au(t) \qquad (\star)$$

ein **lineares Differentialgleichungssystem** (mit konstanten Koeffizienten). Die **Lösungsmenge** von $(\star)$ ist

$$L(\star) = \{u(t) \mid u_i(t) \text{ reelle, differenzierbare Funktionen und } u'(t) = Au(t)\}.$$

Satz 7.2.29 (Aus der Analysis). *Sei $A \in \mathbb{R}^{n\times n}$. Dann ist die* **Exponentialreihe**

$$\exp(A) := \sum_{k=0}^{\infty} \frac{1}{k!} A^k$$

konvergent und die Abbildung $\mathbb{R} \to \mathbb{R}^{n\times n}$, $t \mapsto \exp(tA)$ auf jedem beschränkten Intervall gleichmäßig stetig (bzgl. der Maximumsnorm $\|M\| := \max\{|M_{ij}| \mid 1 \le i,j \le n\}$*).*

Bemerkung 7.2.30. Seien $A, B \in \mathbb{R}^{n\times n}$.

(1) Aus $AB = BA$ folgt $\exp(A)\exp(B) = \exp(B)\exp(A) = \exp(A+B)$.

(2) $\exp(0) = I_n$.

(3) $\exp(A)$ ist invertierbar mit $\exp(A)^{-1} = \exp(-A)$.

(4) $\frac{d}{dt}(\exp(tA)) = A\exp(tA) = \exp(tA)A$ für alle $t \in \mathbb{R}$.

Beweis. Wir zeigen (1): Sei $AB = BA$. Dann ist

$$\begin{aligned}
\exp(A+B) &= \sum_{k=0}^{\infty} \frac{1}{k!}(A+B)^k = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^{k} \binom{k}{j} A^j B^{k-j} \\
&= \sum_{j=0}^{\infty} \sum_{k=j}^{\infty} \frac{1}{j!} A^j \frac{1}{(k-j)!} B^{k-j} = \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{1}{j!} A^j \frac{1}{m!} B^m \\
&= \exp(A)\exp(B).
\end{aligned}$$

Behauptung (2) ist klar, (3) folgt direkt aus (1) & (2) und (4) ist durch gliedweises Differenzieren eine leichte Übung. □

Ende Vorl. 14

Satz 7.2.31 (Aus der Analysis). *Sei $A \in \mathbb{R}^{n\times n}$. Das lineare Differentialgleichungssystem $u'(t) = Au(t)$ hat die Lösungsmenge $L = \{u : \mathbb{R} \to \mathbb{R}^n \mid u(t) = \exp(tA)c$ mit $c \in \mathbb{R}^n\}$. Die eindeutig bestimmte Lösung des Anfangswertproblems $u'(t) = Au(t), u(t_0) = u_0 \in \mathbb{R}^n$ ist $u(t) = \exp((t-t_0)A)u_0$.*

Diese Lösungsmenge L ist ein Vektorraum der Dimension n. Ist $(b_1, \dots, b_n)$ eine Basis von $\mathbb{R}^n$, so ist $(\exp(tA)b_1, \dots, \exp(tA)b_n)$ eine Basis von L. Die Jordan-Normalform von A wird benutzt, um eine solche schöne Basis von L zu finden. Wir formulieren dies nur für zyklische Vektorräume mit $\mu_A = \chi_A = (x-\lambda)^n$ ist. Der allgemeine Fall folgt durch einfaches Zusammensetzen.

Satz 7.2.32. *Sei $A \in \mathbb{R}^{n\times n}$, $\chi_A = \mu_A = (x-\lambda)^n$. Sei $b_1 \in \mathbb{R}^n$ ein zyklischer Vektor, also $b_2 := (A-\lambda I_n)b_1, \dots, b_n := (A-\lambda I_n)^{n-1}b_1 \in \mathbb{R}^n \setminus \{0\}$ und $Ab_n = \lambda b_n$. Dann ist $(b_1, \dots, b_n)$ eine Basis von $\mathbb{R}^n$ und $(\exp(tA)b_1, \dots, \exp(tA)b_n)$ eine Basis von L. Es gilt für $1 \le \ell \le n$ und alle $k \in \mathbb{N}$:*

$$A^k b_\ell = \sum_{j=0}^{n-\ell} \lambda^{k-j} \binom{k}{j} b_{j+\ell}$$

und

$$\exp(tA)b_\ell = \exp(\lambda t) \left(\sum_{j=0}^{n-\ell} \frac{t^j}{j!} b_{j+\ell} \right)$$

für alle $t \in \mathbb{R}$.

Beweis. Wir zeigen

$$A^k b_1 = \sum_{j=0}^{n-1} \lambda^{k-j} \binom{k}{j} b_{j+1}.$$

Die Aussage für $\ell > 1$ ergibt sich dann analog. Dazu wenden wir Induktion über k an. Für $k = 0$ liest sich die Behauptung als $b_1 = b_1$, da $\binom{k}{j} = 0$ für $j > k$. Der Induktionsschluss ist nicht schwerer:

$$\begin{aligned} A^{k+1} b_1 &= A \left(\sum_{j=0}^{n-1} \lambda^{k-j} \binom{k}{j} b_{j+1} \right) = \sum_{j=0}^{n-1} \lambda^{k-j} \binom{k}{j} \underbrace{(b_{j+2} + \lambda b_{j+1})}_{\text{Def. von } b_{j+2}} \\ &= \sum_{j=0}^{n-1} \lambda^{k-j+1} \binom{k}{j-1} b_{j+1} + \sum_{j=0}^{n-1} \lambda^{k-j+1} \binom{k}{j} b_{j+1} = \sum_{j=0}^{n-1} \lambda^{k-j+1} \left(\binom{k}{j-1} + \binom{k}{j} \right) b_{j+1} \\ &= \sum_{j=0}^{n-1} \lambda^{k+1-j} \binom{k+1}{j} b_{j+1}, \end{aligned}$$

wobei wir der Einfachheit halber $b_{n+1} := 0$ setzen. Mit dieser Formel gilt

$$\begin{aligned} \exp(tA) b_\ell &= \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{j=0}^{n-\ell} \lambda^{k-j} \binom{k}{j} b_{j+\ell} = \sum_{k=0}^{\infty} \sum_{j=0}^{n-\ell} \frac{t^{k-j}}{(k-j)!} \lambda^{k-j} \frac{t^j}{j!} b_{j+\ell} \\ &= \left(\sum_{i=0}^{\infty} \frac{(t\lambda)^i}{i!} \right) \left(\sum_{j=0}^{n-\ell} \frac{t^j}{j!} b_{j+\ell} \right) = \exp(\lambda t) \left(\sum_{j=0}^{n-\ell} \frac{t^j}{j!} b_{j+\ell} \right). \end{aligned}$$

□

Beispiel 7.2.33. Gesucht ist die Lösungsmenge des Differentialgleichungssystems

$$\begin{aligned} u_1' &= -u_1 - u_2 - 3u_2' \\ u_2'' &= -u_2 - 2u_2' \end{aligned}$$

Um daraus ein DGL-System erster Ordnung zu machen, setzen wir $u_2' =: u_3$ und erhalten

$$\begin{aligned} u_1' &= -u_1 - u_2 - 3u_3 \\ u_2' &= u_3 \\ u_3' &= -u_2 - 2u_3 \end{aligned}$$

also $u' = Au$, wobei

$$A = \begin{pmatrix} -1 & -1 & -3 \\ 0 & 0 & 1 \\ 0 & -1 & -2 \end{pmatrix}$$

gilt. Das Minimalpolynom von A ist $(x+1)^3$, also ist -1 der einzige Eigenwert von A und A ist ähnlich zum Jordan-Block

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

Der Kern von $(A+1)^2$ ist hier[1] gleich dem Bild von $A+1$ und

$$\mathrm{Kern}((A+1)^2) = \mathrm{Bild}(A+1) = \langle (1,0,0)^{tr}, (0,1,-1)^{tr} \rangle$$

[1] Da es genau einen Jordan-Block gibt, allgemeiner, da alle Jordan-Blöcke zum gegebenen Eigenwert gleich groß sind.

und enthält nicht den 2. Basisvektor. Wir bilden also

$$b_1 = (0,1,0)^{tr}, b_2 = (A+1)b_1 = (-1,1,-1)^{tr}, b_3 = (A+1)b_2 = (2,0,0)^{tr}$$

und erhalten mit

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

dass $B^{-1}AB = J$ gilt. Die Lösungsmenge des DGL-Systems ergibt sich also als Erzeugnis von

$$\exp(tA)b_1 = \exp(-t) \begin{pmatrix} -t+t^2 \\ 1+t \\ -t \end{pmatrix},$$

$$\exp(tA)b_2 = \exp(-t) \begin{pmatrix} -1+2t \\ 1 \\ -1 \end{pmatrix},$$

$$\exp(tA)b_3 = \exp(-t) \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}.$$

Kapitel 8

Gruppen und Operationen

8.1 Operationen von Gruppen auf Mengen.

8.1.a Wiederholung und erste Beispiele

Eine Gruppe G ist eine Menge G mit einer Verknüpfung $\cdot : G \times G \to G$, die das Assoziativgesetz erfüllt, ein Einselement enthält und für jedes $g \in G$ ein inverses Element.

Die **Ordnung** der Gruppe G ist die Anzahl der Elemente der Menge G, also eine natürliche Zahl, falls G **endlich** ist und ∞ falls G nicht endlich ist.

Die Gruppe heißt **Abelsch**, falls zusätzlich das Kommutativgesetz gilt, also $gh = hg$ für alle $g, h \in G$.

Eine Gruppe G heißt **zyklisch**, falls es ein $g \in G$ gibt mit $G = \{g^z \mid z \in \mathbb{Z}\} =: \langle g \rangle$. Die **Ordnung eines Elementes** $g \in G$ ist die Ordnung der davon erzeugten zyklischen Gruppe: $\mathrm{ord}(g) := |\langle g \rangle|$. Die zyklische Gruppe der Ordnung n bezeichnen wir auch mit $C_n = (\mathbb{Z}/n\mathbb{Z}, +)$. Zyklische Gruppen sind Abelsch und nach dem Hauptsatz über endlich erzeugte Abelsche Gruppen ist jede e.e. Abelsche Gruppe das **direkte Produkt** zyklischer Gruppen.

Weitere Beispiele für Gruppen sind

- die symmetrische Gruppe

$$S_n := S_{\underline{n}} := \{\pi : \underline{n} \to \underline{n} \mid \pi \text{ bijektiv } \}$$

  mit $|S_n| = n!$,

- die volle lineare Gruppe eines Vektorraums $\mathrm{GL}(\mathcal{V})$,

- die Gruppe $\mathrm{GL}_n(R)$ der invertierbaren $n \times n$-Matrizen über einem (kommutativen) Ring R.

Definition 8.1.1 (Operation)**.** Die Gruppe G **operiert** auf der Menge M (von links), falls es eine Abbildung $G \times M \to M, (g, m) \mapsto gm$ gibt mit

- $1m = m$ für alle $m \in M$;

- $(gh)m = g(hm)$ für alle $m \in M$, $g, h \in G$.

Eine Menge M mit einer **Operation** von G nennt man auch G**-Menge**.

Bemerkung 8.1.2. Die Gruppe G operiere auf der Menge M. Die **Bahn** von $m \in M$ unter G ist definiert als die Teilmenge

$$Gm := \{gm \mid g \in G\} \subset M.$$

Die Menge aller **Bahnen** in M unter G bilden eine **Partition**[1]

$$M/G = \{Gm \mid m \in M\}$$

auf M.

Beweis. Erinnerung: Eine Partition $\mathcal{P}$ ist eine Teilmenge $\mathcal{P} \subset \mathrm{Pot}(M)$ der Potenzmenge von M mit den Eigenschaften:

- $\emptyset \notin \mathcal{P}$;
- Für $X, Y \in \mathcal{P}$ gilt entweder $X = Y$ oder $X \cap Y = \emptyset$;
- $M = \bigcup_{X \in \mathcal{P}} X$.

Diese Eigenschaften sind nun für M/G zu überprüfen:

- $Gm \neq \emptyset$, da $m = 1m \in Gm$.
- Seien Gm, $Gn \in G\backslash M$. Angenommen $Gm \cap Gn \neq \emptyset$. Dann gibt es $x \in Gm \cap Gn$, also $x = gm = hn$ für geeignete $g, h \in G$. Dann ist aber

  $$m = g^{-1}x = g^{-1}(hn) = (g^{-1}h)n \in Gn$$

  und daher $Gm \subseteq Gn$, denn jedes $um \in Gm$ ist von der Form $um = (ug^{-1}h)n \in Gn$. Aus Symmetriegründen gilt dann auch $Gn \subseteq Gm$ also sind die beiden Bahnen gleich.

  Zwei Bahnen sind entweder gleich oder disjunkt.

- $M = \bigcup_{m \in M} Gm$, da die rechte Seite eine Teilmenge von M ist und umgekehrt jedes $m \in M$ in seiner Bahn Gm liegt und daher auch in der Vereinigung auf der rechten Seite. □

Folgerung 8.1.3. *G operiere auf M. Dann ist $\sim_G \subset M \times M$ definiert durch $a \sim_G b$ genau dann, wenn a und b in derselben Bahn liegen (also genau dann wenn ein $g \in G$ existiert mit $a = gb$) eine Äquivalenzrelation auf M.*

Definition 8.1.4. Sei G eine Gruppe.

(1) $U \subseteq G$ heißt **Untergruppe** von G, kurz $U \leq G$, falls

  (a) $U \neq \emptyset$,

  (b) $g, h \in U$ impliziert $gh^{-1} \in U$.

(2) G operiere auf der Menge M. Für $m \in M$ heißt

  $$\mathrm{Stab}_G(m) := \{g \in G | gm = m\}$$

  der **Stabilisator** von m in G.

Bemerkung 8.1.5. G operiere auf M.

(1) Für $m \in M$ gilt $\mathrm{Stab}_G(m) \leq G$.

(2) Ist $m \in M$ und $g \in G$, so gilt $\mathrm{Stab}_G(gm) = g\,\mathrm{Stab}_G(m)g^{-1}$.

Beweis.

(1) $1m = m$ also ist $1 \in \mathrm{Stab}_G(m)$ und somit $\mathrm{Stab}_G(m) \neq \emptyset$. Sind $g, h \in \mathrm{Stab}_G(m)$, so gilt $hm = m$ und somit auch $h^{-1}m = h^{-1}(hm) = (h^{-1}h)m = 1m = m$ und ebenso

$$(gh^{-1})m = g(h^{-1}m) = gm = m \text{ also } gh^{-1} \in \mathrm{Stab}_G(m).$$

(2) Es gilt $h \in \mathrm{Stab}_G(gm)$ genau dann, wenn $h(gm) = gm$, also $g^{-1}hgm = m$, d.h. $g^{-1}hg \in \mathrm{Stab}_G(m)$ oder äquivalent $h \in g\,\mathrm{Stab}_G(m)g^{-1}$. □

Ende Vorl. 15

Bemerkung 8.1.6. Sei G eine beliebige Gruppe und $U \leq G$ eine Untergruppe. Dann operiert U auf G durch inverse Rechtsmultiplikation:

$$U \times G \to G, (u, g) \mapsto gu^{-1}$$

Die Bahnen

$$gU = \{gu^{-1} \mid u \in U\} = \{gu \mid u \in U\}$$

heißen auch **Linksnebenklassen** von U in G. Die Menge der Linksnebenklassen von G nach U bezeichnen wir mit G/U. Die Anzahl der Linksnebenklassen von U in G heißt der **Index** $[G : U]$ von U in G.

Folgerung 8.1.7. *(**Lagrange**) Sei G eine endliche Gruppe und $U \leq G$. Dann teilt die Ordnung von U die Ordnung von G:*

$$|U| \mid |G|.$$

Der Quotient $\frac{|G|}{|U|}$ ist gleich dem Index von U in G.

Beweis. Die Abbildung $U \to gU, u \mapsto gu$ ist eine Bijektion. Also haben je zwei Linksnebenklassen aus der Partition G/U von G genau $|U|$ Elemente. Nun ist $G = g_1U \dot\cup g_2U \ldots \dot\cup g_sU$ mit $s = [G : U]$ und somit $|G| = \sum_{i=1}^{s} |g_iU| = s|U|$. □

Definition 8.1.8. Die Operation der Gruppe G auf dem K-Vektorraum $\mathcal{V}$ heißt **linear**, falls für jedes $g \in G$ die Abbildung

$$\widehat{g} : \mathcal{V} \to \mathcal{V} : V \mapsto gV$$

linear ist.

Beispiel 8.1.9. Sei $M := \mathbb{F}_2^3$. Die symmetrische Gruppe $G = S_3$ operiert auf M durch $(\pi, (a_1, a_2, a_3)) \mapsto (a_{\pi^{-1}(1)}, a_{\pi^{-1}(2)}, a_{\pi^{-1}(3)})$.
Bahnen: $\{(0,0,0)\}$, $\{(1,1,1)\}$, $\{(1,0,0),(0,1,0),(0,0,1)\}$, $\{(1,1,0),(0,1,1),(1,0,1)\}$. Stabilisatoren: $\mathrm{Stab}_{S_3}((1,0,0)) = \left\{\mathrm{id}, \pi := \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 1 & 3 & 2 \\ \hline \end{array}\right\}$. $\mathrm{Stab}_{S_3}(1,1,1) = S_3$.

Diese Operation ist linear, z.B. ist bzgl. der Standardbasis S

$${}^S\widehat{\pi}^S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Bemerkung 8.1.10.

(1) Operiert G auf der Menge M, so erhält man einen Gruppenhomomorphismus

$$G \to S_M, g \mapsto (\widehat{g} : m \mapsto gm).$$

Umgekehrt definiert jeder Gruppenhomomorphismus in S_M eine Operation auf M.

[1] Nicht zu verwechseln mit dem Komplementzeichen!

(2) Operiert G linear auf dem Vektorraum $\mathcal{V}$, so liefert dies einen Gruppenhomomorphismus

$$G \to \mathrm{GL}(\mathcal{V}) \leq S_{\mathcal{V}}.$$

Umgekehrt definiert jeder Gruppenhomomorphismus in $\mathrm{GL}(\mathcal{V})$ eine lineare Operation auf M.

Definition 8.1.11. Die Operation von G auf M heißt

(1) **transitiv**, falls M eine Bahn bildet, d.h. $M = Gm$ für ein $m \in M$ (und somit $M = Gm$ für jedes $m \in M$).

(2) **regulär** oder **scharf transitiv**, falls sie nicht leer und transitiv ist und $\mathrm{Stab}_G(m) = \{1\}$ für ein und somit alle $m \in M$.

(3) **treu**, falls $gm = m$ für alle $m \in M$ impliziert $g = 1$.

Satz 8.1.12. *Die Gruppe G operiere auf der Menge M. Folgende Aussagen sind äquivalent:*

(1) G operiert scharf transitiv auf M.

(2) Zu je zwei $m, n \in M$ gibt es genau ein $g \in G$ mit $gm = n$. (Man ist versucht dieses $g \in G$ mit $\overrightarrow{mn}$ zu bezeichnen.)

(3) Für jedes feste $m_0 \in M$ ist die Abbildung

$$G \to M : g \mapsto gm_0$$

bijektiv.

(4) Es existiert ein $m_0 \in M$, so daß die Abbildung

$$G \to M : g \mapsto gm_0$$

bijektiv ist.

Beweis.

(1) $\Rightarrow$ (2): Wegen der Transitivität existiert ein $g \in G$ mit $gm = n$. Angenommen es gibt ein weiteres $h \in G$ mit $hm = n$. Dann ist $h^{-1}g \in \mathrm{Stab}_G(m) = \{1_G\}$, also $h = g$.

(2) $\Rightarrow$ (3): Definiert ist die Abbildung immer. Sie ist surjektiv, da G transitiv operiert. Sie ist injektiv wegen der vorausgesetzten Eindeutigkeit.

(3) $\Rightarrow$ (4): Klar.

(4) $\Rightarrow$ (1): $\mathrm{Stab}_G(m_0) = \{1_G\}$ wegen der vorausgesetzten Injektivität. Wegen der vorausgesetzten Surjektivität ist die Operation auch transitiv. □
Zusatz: Ist $m \in M$ beliebig, so haben wir ein $g \in G$ mit $gm_0 = m$. Also $\mathrm{Stab}_G(m) = \mathrm{Stab}_G(gm_0) = g\,\mathrm{Stab}_G(m_0)g^{-1} = \{1_G\}$

Beispiel 8.1.13.

(1) Die Operation von G auf sich per Linksmultiplikation $G \times G \to G, (g, h) \mapsto gh$ ist scharf transitiv.

(2) Sei $\mathcal{V}$ ein endlich erzeugter K-Vektorraum und $\mathcal{B}(\mathcal{V}) \subset \mathcal{V}^n$ die Menge der Basen von $\mathcal{V}$. Dann ist die Operation

$$\mathrm{GL}(\mathcal{V}) \times \mathcal{B}(\mathcal{V}) \to \mathcal{B}(\mathcal{V}) : (g, B) \mapsto gB := (g(B_1), \ldots, g(B_n))$$

scharf transitiv, da jede lineare Abbildung durch die Bilder einer Basis festgelegt ist.

Beispiel 8.1.14. Sei $L = \{x \in K^n \mid Ax = b\} \neq \emptyset$ die Lösungsmenge eines linearen GLS und $\mathcal{U} = \{x \in K^n \mid Ax = 0\}$ die Lösungsmenge des zugehörigen homogenen Systems. Dann ist $\mathcal{U}$ ein K-Vektorraum, also insbesondere eine Gruppe, die scharf transitiv auf L durch Addition operiert. Hat man eine partikuläre Lösung x_0 des inhomogenen Systems gefunden, so ist $\{x_0 + y \mid y \in \mathcal{U}\} = L$, die Lösungsmenge des inhomogenen Systems. Die Beobachtung wird auch der Ausgangspunkt für die affine Geometrie sein.

Beispiel 8.1.15. Ist $\varphi : \mathcal{V} \to \mathcal{W}$ eine lineare Abbildung von K-Vektorräumen und $W \in \mathrm{Bild}(\varphi)$, so operiert $\mathrm{Kern}(\varphi)$ scharf transitiv auf der Faser $\varphi^{-1}(\{W\})$. Dies ist die koordinatenfreie Version des letzten Beispiels.

8.1.b Die Konjugationsoperation

Definition 8.1.16. Sei G eine Gruppe. Dann operiert G auf sich selbst durch **Konjugation**,

$$G \times G \to G, (g, m) \mapsto \kappa_g(m) := gmg^{-1}.$$

Die Bahnen unter dieser Operation heißen **Konjugiertenklassen**. Der Stabilisator von $m \in G$ wird auch als **Zentralisator** bezeichnet

$$C_G(m) = \{g \in G \mid gmg^{-1} = m\} = \{g \in G \mid gm = mg\}.$$

Bemerkung 8.1.17. Für $g \in G$ ist die Abbildung $\kappa_g : G \to G, m \mapsto gmg^{-1}$ ein bijektiver Gruppenhomomorphismus von G in sich selbst, also ein Gruppenautomorphismus mit $(\kappa_g)^{-1} = \kappa_{g^{-1}}$.

Übung 8.1.1.

(1) Die Menge aller Gruppenautomorphismen von G bildet (zusammen mit der Komposition) eine Gruppe $\mathrm{Aut}(G)$.

(2) Die Abbildung $\kappa : G \to \mathrm{Aut}(G), g \mapsto \kappa_g$ ist ein Gruppenhomomorphismus von G in ihre **Automorphismengruppe**.

(3) Der Kern von κ ist das **Zentrum** $Z(G)$ von G:

$$Z(G) = \{g \in G \mid gh = hg \text{ für alle } h \in G\}.$$

Das Bild von κ wird auch mit $\mathrm{Inn}(G)$ bezeichnet und heißt die Gruppe der **inneren Automorphismen** von G.

Übung 8.1.2. Betrachte die Diedergruppe $D_8 = \left\langle \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle$ der Ordnung 8 als Symmetriegruppe eines Quadrats. Bestimme alle Untergruppen, Konjugierten, Zentrum, etc.

Ende Vorl. 16

8.1.c Parametrisierung aller transitiver G-Mengen.

Definition 8.1.18. (Ähnlichkeit von G-Mengen) Sei G eine Gruppe, M, N zwei G-Mengen. Eine Abbildung $\varphi : M \to N$ heißt **G-äquivariant**, wenn $\varphi(gm) = g\varphi(m)$ für alle $g \in G$ und $m \in M$.

M und N heißen **ähnlich**, falls es eine G-äquivariante Bijektion $\varphi : M \to N$ gibt (in Zeichen $M \cong_G N$). Die Abbildung φ heißt auch eine **Ähnlichkeit** der G-Mengen M und N.

Beispiel 8.1.19. Sind $U \leq S \leq G$ Untergruppen von G, so ist die Abbildung $G/U \to G/S, gU \mapsto gS$ eine G-äquivariante Abblidung.

Satz 8.1.20 (Bahnensatz)**.** *Seien M ein transitive G-Menge M und $m \in M$. Dann sind M und $G/\operatorname{Stab}_G(m)$ als G-Mengen ähnlich, genauer ist die Abbildung*

$$\overline{\varphi}_m : G/\operatorname{Stab}_G(m) \to M : g\operatorname{Stab}_G(m) \mapsto gm$$

eine G-Ähnlichkeit.

Beweis. Offenbar ist $\varphi_m : G \to M : g \mapsto gm$ eine surjektive G-äquivariante Abbildung, wobei G durch Linksmultiplikation auf sich operiert. Die Fasern von φ_m sind gerade die Linksnebenklassen von G nach $\operatorname{Stab}_G(m)$:

$$\varphi_m^{-1}(\{gm\}) = g\operatorname{Stab}_G(m) \qquad \text{für alle } g \in G.$$

Also (nach dem Homomorphiesatz für Mengen) faktorisiert φ_m über $G/\operatorname{Stab}_G(m)$ mit einer Bijektion

$$\overline{\varphi}_m : G/\operatorname{Stab}_G(m) \to M : g\operatorname{Stab}_G(m) \mapsto gm,$$

die offensichtlich G-äquivariant ist. □

In Worten: Alle transitiven Operationen findet man „in der Gruppe“ wieder.

Folgerung 8.1.21. *Die Gruppe G operiere auf der Menge M. Sei $m \in M$ mit $|Gm| < \infty$. Dann gilt: Die* **Länge der Bahn** *ist gleich dem Index des Stabilisators:*

$$|Gm| = [G : \operatorname{Stab}_G(m)] \qquad (:= |G/\operatorname{Stab}_G(m)|).$$

Und falls $|G| < \infty$, dann gilt sogar

$$|Gm| = \frac{|G|}{|\operatorname{Stab}_G(m)|}.$$

Dies gibt uns eine Strategie sowohl Mengenmächtigkeiten als auch Gruppenordnungen zu bestimmen!

Beispiel 8.1.22 (Bestimmung der Ordnung der **vollen linearen Gruppe**)**.** Sei $\mathbb{F}_q$ ein Körper mit $q = p^a$ Elementen und

$$\operatorname{GL}_n(\mathbb{F}_q) = \{X \in \mathbb{F}_q^{n\times n} \mid \det X \neq 0\}.$$

Dann gilt

$$|\operatorname{GL}_n(\mathbb{F}_q)| = q^{\binom{n}{2}}(q^n-1)(q^{n-1}-1)\cdot\ldots\cdot(q-1) = (q^n-1)\cdot(q^n-q)\cdot\ldots\cdot(q^n-q^{n-1}).$$

Beweis. $\operatorname{GL}_n(\mathbb{F}_q)$ operiert transitiv auf $\mathbb{F}_q^{n\times 1}\backslash\{0\}$ mit Bahnlänge $q^n - 1$. Der Stabilisator von $(1, 0, \ldots, 0)^{tr} \in \mathbb{F}_q^{n\times 1}$ in $\operatorname{GL}_n(\mathbb{F}_q)$ ist

$$\operatorname{Stab}_{\operatorname{GL}_n(\mathbb{F}_q)}\left(\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}\right) = \left\{\begin{pmatrix}1 & *\cdots* \\ 0 & \\ \vdots & X \\ 0 & \end{pmatrix} \mid * \in \mathbb{F}_q, X \in \operatorname{GL}_{n-1}(\mathbb{F}_q)\right\}$$

und hat die Ordnung

$$|\operatorname{Stab}_{\mathrm{GL}_n(\mathbb{F}_q)}\left((1,0,\ldots,0)^{tr}\right)| = q^{n-1}\cdot|\operatorname{GL}_{n-1}(\mathbb{F}_q)|.$$

Nach dem Bahnensatz 8.1.20 gilt also

$$|\operatorname{GL}_n(\mathbb{F}_q)| = (q^n-1)\cdot q^{n-1}\cdot|\operatorname{GL}_{n-1}(\mathbb{F}_q)|.$$

Mit $|\operatorname{GL}_1(\mathbb{F}_q)| = q-1$ folgt die Behauptung (etwa durch Induktion). □

Übung 8.1.3. (**Gaußsche Binomialkoeffizienten**) Betrachte die Menge

$$\mathcal{U}(\mathbb{F}_q^{n\times 1}) := \{X \mid X \leq_{\mathbb{F}_q} \mathbb{F}_q^{n\times 1}\}$$

der $\mathbb{F}_q$-Teilräume von $\mathbb{F}_q^{n\times 1}$ bzw. die Teilmenge $\mathcal{U}_k(\mathbb{F}_q^{n\times 1},\mathbb{F}_q)$ der k-dimensionalen Teilräume von $\mathbb{F}_q^{n\times 1}$. Zeige:

$$|\mathcal{U}_k(\mathbb{F}_q^{n\times 1})| = \frac{|\operatorname{GL}_n(\mathbb{F}_q)|}{|\operatorname{GL}_k(\mathbb{F}_q)|\cdot|\operatorname{GL}_{n-k}(\mathbb{F}_q)|\cdot q^{k(n-k)}} =: \begin{bmatrix} n \\ k \end{bmatrix}_q,$$

der GAUSSsche Binomialkoeffizient.

Übung 8.1.4. $\mathrm{GL}_n(\mathbb{F}_q)$ operiert auf $\mathbb{F}_q^{n\times n}$ durch Konjugation. Die Bahnen sind die Ähnlichkeitsklassen von Matrizen, von denen wir eine Parametrisierung im vorherigen Kapitel kennengelernt haben. Der Stabilisator einer Matrix A ist $C_{\mathrm{GL}_n(\mathbb{F}_q)}(A) := C_{\mathbb{F}_q^{n\times n}}(A)^*$, die Einheitengruppe des Zentralisators, auch Zentralisator von A in $\mathrm{GL}_n(\mathbb{F}_q)$ genannt. Ihr Index gibt an, wieviele Matrizen zu A ähnlich sind. Der Fall $n=2, q=2$ kann man wie folgt zusammenfassen:

μ_A	χ_A	Vertreter	$\lvert C_{\mathrm{GL}_n(\mathbb{F}_q)}(A)\rvert$	Anzahl
x	x^2	$\begin{pmatrix}0&0\\0&0\end{pmatrix}$	6	1
$x+1$	$(x+1)^2$	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	6	1
x^2	x^2	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	2	3
$(x+1)^2$	$(x+1)^2$	$\begin{pmatrix}1&0\\1&1\end{pmatrix}$	2	3
$x(x+1)$	$x(x+1)$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	1	6
x^2+x+1	x^2+x+1	$\begin{pmatrix}0&1\\1&1\end{pmatrix}$	3	2

Behandle den Fall $n=3, q=2$.

Ende Vorl. 17

8.1.d Zykel, Zykelschreibweise und Zykelzähler

Wir wollen eine recht nützliche Schreibweise für Permutationen endlicher Mengen M, also für Elemente der S_M, kennenlernen. Sie hat gegenüber der offensichtlichen Schreibweise, wo man die Elemente von M in die erste Zeile einer $2\times|M|$-Matrix schreibt und die Bilder dieser Elemente direkt unter diese in die zweite Zeile, viele Vorteile.

Bemerkung 8.1.23.

(1) Seien M eine Menge und $a_1, \ldots, a_k \in M$ genau k paarweise verschiedene Elemente. Dann heißt

$$(a_1, a_2, \ldots, a_k) : M \to M : a \mapsto \begin{cases} a & a \notin \{a_1, a_2, \ldots, a_k\} \\ a_{i+1} & a = a_i \text{ mit } i < k \\ a_1 & a = a_k \end{cases}$$

ein k-**Zykel** oder kurz **Zykel**. Es gilt $(a_1, a_2, \ldots, a_k) \in S_M$ und

$$(a_1, a_2, \ldots, a_k) = (a_k, a_1, a_2, \ldots, a_{k-1})$$

und

$$(a_1, a_2, \ldots, a_k)^{-1} = (a_k, a_{k-1}, \ldots, a_1) = (a_1, a_k, \ldots, a_2).$$

1-Zykel sind gleich der Identität, 2-Zykel sind Transpositionen.

(2) Disjunkte Zykel **kommutieren** miteinander, d. h.

$$(a_1, a_2, \ldots, a_k) \circ (b_1, b_2, \ldots, b_l) = (b_1, b_2, \ldots, b_l) \circ (a_1, a_2, \ldots, a_k),$$

falls $\{a_1, a_2, \ldots, a_k\} \cap \{b_1, b_2, \ldots, b_l\} = \emptyset$. (In Zukunft lassen wir $\circ$ einfach weg.)

(3) Jedes $f \in S_M$ läßt sich als Produkt disjunkter Zykel schreiben, falls M endlich ist. Diese Schreibweise ist eindeutig bis auf Reihenfolge der Zykel.

Beweis.

(1) & (2) sind klar.

(3) Die Existenz zeigen wir mit einem Algorithmus, der u.A. die Bahnen der zyklischen Untergruppe $\langle f \rangle \leq S_M$ auf M konstruiert:
Eingabe: $f \in S_M$ (Beachte: M ist endlich und nicht leer)
Algorithmus:

1. Setze $N := M$.
2. Solange $N \neq \emptyset$ ist, wähle $a \in N$ und finde kleinstes k mit $f^k(a) = a$. Dann ist $z_a := (a, f(a), \ldots, f^{k-1}(a))$ ein Zykel.
3. Ersetze N durch $N \setminus \{a, f(a), \ldots, f^{k-1}(a)\}$ und springe zu Schritt 2 falls $N \neq \emptyset$.
4. Ausgabe: Menge der berechneten disjunkten Zykel z_a. Ihr Produkt ist gleich f.

Eindeutigkeit: Übung. □

Beispiel 8.1.24. $(1, 2, 3, 4, 5) \circ (2, 3, 4, 5) = (1, 2, 4)(3, 5) \in S_6$.

Bemerkung 8.1.25. Für ein Zykel $\pi = (a_1, \ldots, a_k)$ gilt $\operatorname{sign}(\pi) = (-1)^{k-1}$, da

$$\pi = \underbrace{(a_2, a_3)(a_3, a_4) \cdots (a_{k-1}, a_k)(a_1, a_k)}_{k-1 \text{ Transpositionen}}$$

Wir schauen uns im Folgenden die Konjugationsoperation in der symmetrischen Gruppe genauer an.

Übung 8.1.5. Konjugation erhält den Zykeltyp, genauer

$$\sigma(a_1, \ldots, a_k)\sigma^{-1} = (\sigma(a_1), \ldots, \sigma(a_k))$$

für alle $\sigma \in S_n$.

Satz 8.1.26. *Für $\pi \in S_n$ sei $a_i(\pi)$ die Anzahl der Zykel der Länge i in disjunkter Zykelzerlegung von π und $a(\pi) := (a_1(\pi), \ldots, a_n(\pi))$ der* **Zykelzähler**. *Es gilt: $\pi, \rho \in S_n$ sind genau dann konjugiert in S_n, wenn $a(\pi) = a(\rho)$.*

Mit anderen Worten, der Zykelzähler ist eine trennende Invariante der Operation von S_n auf sich per Konjugation.

Beweis. 1. Konjugieren erhält den Zykelzähler:
Seien $\pi, \sigma \in S_n$ mit

$$\pi = (\alpha_1, \ldots, \alpha_k)(\beta_1, \ldots, \beta_l)\ldots$$

in disjunkter Zykelzerlegung. Dann ist

$$\sigma\pi\sigma^{-1} = \sigma(\alpha_1, \ldots, \alpha_k)\sigma^{-1}\sigma(\beta_1, \ldots, \beta_l)\sigma^{-1}\ldots = (\sigma(\alpha_1), \ldots, \sigma(\alpha_k))(\sigma(\beta_1), \ldots, \sigma(\beta_l))\ldots$$

die disjunkte Zykelzerlegung von $\sigma\pi\sigma^{-1}$. Insbesondere hat sich der Zykelzähler nicht verändert.
2. Umgekehrt: Sei $a(\pi) = a(\rho)$. Ordnet man bei beiden Permutationen die Zykel (aus der disjunkten Zykelzerlegung) der Länge nach, so stehen im folgenden Schema die Zykel gleicher Länge untereinander:

$$\begin{aligned} \pi &= (\alpha_1, \ldots, \alpha_k)(\beta_1, \ldots, \beta_l)\ldots \\ \rho &= (\alpha'_1, \ldots, \alpha'_k)(\beta'_1, \ldots, \beta'_l)\ldots \end{aligned}$$

in disjunkter Zykelzerlegung. Also definiere nach Übung 8.1.5 $\sigma \in S_n$ durch

$$\sigma : \alpha_i \to \alpha'_i; \beta_j \to \beta'_j \text{ etc.}$$

Dann gilt $\sigma\pi\sigma^{-1} = \rho$. □

Übung 8.1.6. Zeige jedes $a \in \mathbb{Z}^n_{\geq 0}$ mit $\sum ia_i = n$ kommt als Zykelzähler eines Elementes von S_n vor. Wieviele Konjugiertenklassen hat S_6? Was haben Zykelzähler mit Partitionen der Zahl n zu tun?

Beispiel 8.1.27 (Konjugation in der S_n)**.**

(1) Gegeben seien $\pi = (1,2)(3,4,5,6,7)(8,9), \rho = (9,4)(1,2,3,7,6)(5,8) \in S_9$. Suche ein $\sigma \in S_9$ mit $\sigma\pi\sigma^{-1} = \rho$.
Lösung:

$$\begin{aligned} \pi &= (1,2)(3,4,5,6,7)(8,9) \\ \rho &= (9,4)(1,2,3,7,6)(5,8) \end{aligned}$$

Also ist $\sigma = (1,9,8,5,3)(2,4)(6,7)$ eine Lösung.
Gibt es andere Lösungen? Ja, z.B. liefert das obige Vorgehen nach dem Umschreiben von ρ

$$\begin{aligned} \pi &= (1,2)(3,4,5,6,7)(8,9) \\ \rho &= (5,8)(2,3,7,6,1)(9,4) \end{aligned}$$

eine weitere Lösung $\sigma = (1,5,7)(2,8,9,4,3)$. Beachte, die Lösungen bilden eine Linksnebenklasse nach $C_{S_9}(\pi)$. Man sieht leicht, die Elemente von $C_{S_9}(\pi)$ entsprechen genau den Möglichkeiten eine disjunkte Zykelzerlegung von π kompatibel unter die gegebene disjunkte Zykelzerlegung von π zu schreiben. Also $|C_{S_9}(\pi)| = 2 \cdot 2^2 \cdot 5$.

(2) Bestimme $C_{S_{10}}((1,2,3))$.
Lösung:

$$\begin{aligned} C_{S_{10}}((1,2,3)) &= \{\pi \in S_{10} | \pi(1,2,3)\pi^{-1} = (1,2,3)\} \\ &= \{\pi \in S_{10} | (\pi(1), \pi(2), \pi(3)) = (1,2,3)\} \end{aligned}$$

Also $\pi = (1,2,3)^i\rho$ mit $\rho \in S_{10}, \rho(i) = i$ für $i = 1,2,3$. Kurz

$$C_{S_{10}}((1,2,3)) = \langle(1,2,3)\rangle \times S_{\underline{10}-\underline{3}}$$

8.1.e Anzahl der Bahnen des Stabilisators

Für unsere geometrischen Anwendungen der Gruppentheorie ist die folgende Bemerkung grundlegend.

Bemerkung 8.1.28. Die Gruppe G operiere auf den Menge M und N.

(1) G operiert auf $M \times N$ durch

$$G \times (M \times N) \to M \times N : (g, (m, n)) \mapsto (gm, gn).$$

Diese Operation heißt **diagonale Operation**.

(2) Ist die Operation von G auf M transitiv, dann ist die Abbildung

$$\begin{aligned}(M \times N)/G &\to N/\operatorname{Stab}_G(m)\\ G(m,n) &\mapsto \operatorname{Stab}_G(m)n\end{aligned}$$

eine Bijektion zwischen der Menge der Bahnen von G auf $M \times N$ und der Menge der $\operatorname{Stab}_G(m)$-Bahnen auf N für jedes (feste) $m \in M$.

Beweis.

(1) Übung.

(2) Wir zeigen, daß diese Abbildung wohldefiniert ist:
Wegen der Transitivität von G auf M, ist jede Bahn von G auf $M \times N$ von der Form $G(m,n) = \{(gm, gn) | g \in G\}$. Gilt $G(m,n) = G(m,n')$ für ein $n' \in N$, so sind offenbar n und n' in derselben Bahn unter $\operatorname{Stab}_G(m)$. Also ist die Abbildung wohldefiniert.
Offenbar ist die Abbildung surjektiv. Wir zeigen die Injektivität:
$\operatorname{Stab}_G(m)n = \operatorname{Stab}_G(m)n'$ impliziert (und ist äquivalent zu) $G(m,n) = G(m,n')$. Also haben wir insgesamt eine Bijektion. □

Ende Vorl. 18

Beispiel 8.1.29. Sei $\mathcal{V}$ ein endlich erzeugter K-Vektorraum. Dann operiert $G := \operatorname{GL}(\mathcal{V})$ auf $\mathcal{V} \setminus \{0\}$ transitiv. Der Stabilisator eines $V \in \mathcal{V} \setminus \{0\}$ hat dann jedes Vielfache $\neq 0$ von V als Bahn, sowie die Menge aller Vektoren, die linear unabhängig von V sind. Die Bahnen von $\operatorname{GL}(\mathcal{V})$ auf $(\mathcal{V} \setminus \{0\}) \times (\mathcal{V} \setminus \{0\})$ sind also gegeben durch $\{(V, aV) \mid V \neq 0\}, a \in K$ und $\{(V,W) \mid (V,W)$ linear unabhängig $\}$.
In Matrizen: $\mathcal{V} = K^{n\times 1}$, $G = \operatorname{GL}_n(K)$, Operation durch Linksmultiplikation. Der Stabilisator des ersten Standardbasisvektors $E_1 := (I_n)_{-,1}$ ist

$$\operatorname{Stab}_G(E_1) := \left\{ \left(\begin{array}{c|c} 1 & a \\ \hline 0 & A \end{array}\right) | a \in K^{1\times(n-1)}, A \in \operatorname{GL}_{n-1}(K) \right\}$$

und hat die folgenden Bahnen auf $\mathcal{V}$:

$$\{aE_1\} \text{ mit } a \in K \setminus \{0\} \text{ und } \mathcal{V} \setminus \{aE_1 | a \in K\}.$$

Beispiel 8.1.30. Sei $\mathcal{V}$ ein endlich erzeugter K-Vektorraum. Dann operiert $G := \operatorname{GL}(\mathcal{V})$ auf dem **Dualraum** $\mathcal{V}^* := \operatorname{Hom}(\mathcal{V}, K)$ linear und treu durch

$$G \times \mathcal{V}^* \to \mathcal{V}^* : (g, \varphi) \mapsto \varphi \circ g^{-1}.$$

Der Stabilisator eines $\varphi \in \mathcal{V}^* \setminus \{0\}$ operiert auf jeder Faser $\varphi^{-1}(\{a\})$ mit $a \in K$, also auf jeder Restklasse nach $\operatorname{Kern}(\varphi)$. Das Studium der Operation von $\operatorname{Stab}_G(\varphi)$ auf $\varphi^{-1}(\{1\})$ heißt

affine Geometrie und wird uns noch ausführlich beschäftigen.
In Matrizen: $\mathcal{V} = K^{n\times 1}$, $G = \mathrm{GL}_n(K)$, die Operation auf $K^{1\times n}$, dem bekanntlich $\mathcal{V}^*$ entspricht, ist gegeben durch

$$G \times K^{1\times n} \to K^{1\times n} : (g, Z) \mapsto Zg^{-1}.$$

Der Stabilisator des letzten Standardbasisvektors $Z_n := (I_n)_{n,-}$ ist

$$\mathrm{Stab}_G(Z_n) := \left\{ \left(\begin{array}{c|c} A & a \\ \hline 0 & 1 \end{array}\right) | a \in K^{(n-1)\times 1}, A \in \mathrm{GL}_{n-1}(K) \right\}.$$

Die Operation dieser Gruppe auf

$$\varphi^{-1}(\{1\}) = \left\{ \begin{pmatrix} S \\ \hline 1 \end{pmatrix} | S \in K^{(n-1)\times 1} \right\}$$

mit $\varphi(\begin{pmatrix} V_1 \\ \vdots \\ V_n \end{pmatrix}) \coloneqq Z_n \begin{pmatrix} V_1 \\ \vdots \\ V_n \end{pmatrix} = V_n$ wird also die affine Geometrie sein. Man beachte:

$$\left(\begin{array}{c|c} A & a \\ \hline 0 & 1 \end{array}\right) \begin{pmatrix} S \\ \hline 1 \end{pmatrix} = \begin{pmatrix} AS + a \\ \hline 1 \end{pmatrix}.$$

8.2 Homomorphismen und Normalteiler

Definition 8.2.1. Seien G und H Gruppen. Eine Abbildung $\varphi : G \to H$ heißt ein **Gruppenhomomorphismus**, wenn $\varphi(g_1g_2) = \varphi(g_1)\varphi(g_2)$ für alle $g_1, g_2 \in G$.

Beispiel 8.2.2. G operiere auf M. Dann ist $\varphi : G \to S_M, g \mapsto (m \mapsto gm)$ ein Gruppenhomomorphismus.
Ist $M = \mathcal{V}$ ein K-Vektorraum, so ist die Operation genau dann linear, wenn das Bild dieses Gruppenhomomorphismus in der linearen Gruppe $\mathrm{GL}(\mathcal{V}) \leq S_\mathcal{V}$ liegt.

Definition 8.2.3. Eine Untergruppe $U \leq G$ mit $gUg^{-1} = U$ für alle $g \in G$ heißt **Normalteiler** von G. Wir schreiben dann auch $U \trianglelefteq G$.

Satz 8.2.4. *Sei $\varphi : G \to H$ ein Gruppenhomomorphismus. Dann ist* $\mathrm{Bild}(\varphi) = \{\varphi(g) \mid g \in G\}$ *eine Untergruppe von H und* $\mathrm{Kern}(\varphi) := \{g \in G \mid \varphi(g) = 1\}$ *ein Normalteiler von G.*

Beweis. Für $n \in \mathrm{Kern}(\varphi)$, $g \in G$ ist

$$\varphi(gng^{-1}) = \varphi(g)\varphi(n)\varphi(g)^{-1} = \varphi(g)\varphi(g)^{-1} = 1$$

also auch $gng^{-1} \in \mathrm{Kern}(\varphi)$. □

Beispiel 8.2.5 (Konkrete Beispiele für Normalteiler)**.**

(1) Ist G beliebige Gruppe, so sind $\{1\}$ und G Normalteiler von G. Man nennt sie **triviale Normalteiler**.

(2) $\mathrm{sign} : S_n \to (\{1, -1\}, \cdot)$ ist ein Homomorphismus. Der Kern ist also ein Normalteiler vom Index 2 von S_n, die sogenannte **alternierende Gruppe** $A_n := \mathrm{Kern}(\mathrm{sign})$ vom Grad n. Ihre Elemente heißen **gerade Permutationen**.

(3) $\det : \mathrm{GL}_n(K) \to K^* := (K \setminus \{0\}, \cdot)$ ist ein Homomorphismus für jeden Körper K. Also ist sein Kern ein Normalteiler von $\mathrm{GL}_n(K)$. Dieser wird mit $\mathrm{SL}_n(K)$ bezeichnet und heißt **spezielle lineare Gruppe** vom Grad n.

(4) Sei $U \leq G$ eine Untergruppe von G. Dann ist $\mathrm{Core}(U) := \bigcap_{g\in G} gUg^{-1}$ der größte Normalteiler von G, der in U enthalten ist. Es gilt $\mathrm{Core}(U) = \mathrm{Kern}(G \to S_{G/U})$.

(5) Ist G eine abelsche Gruppe (also $gh = hg$ für alle $g, h \in G$), so ist jede Untergruppe von G ein Normalteiler.

Bemerkung 8.2.6. Eine Untergruppe $N \leq G$ ist genau dann ein Normalteiler von G, wenn sie Vereinigung von Konjugiertenklassen von Elementen ist.

Beweis. N ist genau dann ein Normalteiler, wenn N invariant unter der Operation von G auf sich per Konjugation ist, sprich wenn sich die Konjugationsoperation von G auf N einschränken lässt. □

Übung 8.2.1. Mithilfe von Bemerkung 8.2.6 zeige, dass

$$S_4, A_4, V_4, \{1\}.$$

alle Normalteiler von S_4 sind.

Satz 8.2.7. *Sei N ein Normalteiler von G. Dann bildet die Menge der Nebenklassen $G/N = \{gN \mid g \in G\}$ eine Gruppe unter vertreterweiser Multiplikation*

$$G/N \times G/N \to G/N, (gN)(hN) := (gh)N.$$

Beweis. Es ist klar, dass wir eine Gruppe vorliegen haben, sobald die Verknüpfung wohldefiniert ist, da die Rechenregeln dann aus denen von G folgen.
Zur Wohldefiniertheit:
Sei $g' = gn \in gN$ und $h' = hm \in hN$. Dann gilt

$$(g'h')N = (gnhm)N = (gh)\underbrace{(h^{-1}nh)m}_{\in N}N = (gh)N.$$ □

Ende Vorl. 19

Normalteiler sind also genau die Untergruppen N für die die Menge G/N der Linksnebenklassen wieder eine Gruppe ist.

Beispiel 8.2.8. Die $S_3 = \mathrm{Stab}_{S_4}(4)$ ist kein Normalteiler von S_4. Dies können wir auf zwei Arten sehen. Zum einen ist nach $\pi S_3\pi^{-1} = \pi\,\mathrm{Stab}_{S_4}(4)\pi^{-1} = \mathrm{Stab}_{S_4}(\pi(4))$ ist. Und zum anderen bilden die Linksnebenklassen $\{S_3, aS_3, a^2S_3, a^3S_3\}$ mit $a = (1,2,3,4)$ keine Gruppe. Es ist z.B. $(aS_3)^2 = \{ahag \mid g, h \in S_3\} = S_4$.

Bemerkung 8.2.9. Eine Untergruppe N ist genau dann ein Normalteiler von G, wenn $gN = Ng$ ist für alle $g \in G$ (Links- und Rechtsnebenklassen stimmen überein). Insbesondere sind Untergruppen von Index 2 immer Normalteiler, denn es ist $N \cup N = G = N \cup Ng$, also $gN = G \setminus N = Ng$.

Definition 8.2.10. Eine Gruppe $G \neq \{1\}$ heißt **einfach**, falls G und $\{1\}$ die einzigen Normalteiler von G sind.

Übung 8.2.2. Zeige: Eine zyklische Gruppe C_n der Ordnung n ist genau dann einfach, wenn n eine Primzahl ist.
Hinweis: Beachte, $C_n \cong (\mathbb{Z}/n\mathbb{Z}, +)$ (wieso?) und insbesondere abelsch.

Übung 8.2.3. Zeige, dass A_5 eine einfache Gruppe ist.

Die Bausteine aller endlichen Gruppen sind die endlichen **einfachen** Gruppen, das sind die endlichen Gruppen, die nur sich selbst und $\{1\}$ als Normalteiler haben.

Hauptsatz 8.2.11. *(**Homomorphiesatz für Gruppen**)*

(1) *Ist G Gruppe und $N \trianglelefteq G$ ein Normalteiler von G, dann bildet die Menge G/N der Nebenklassen (Links = Rechts) von G nach N eine Gruppe mit vertreterweiser Multiplikation:*

$$gN \cdot hN := ghN \text{ für alle } g, h \in G$$

und der **natürliche Epimorphismus**

$$\nu = \nu_N : G \to G/N : g \mapsto gN$$

ist ein surjektiver Homomorphismus mit Kern N.

(2) *Ist $\varphi : G \to H$ ein Gruppenhomomorphismus, so ist $\operatorname{Kern}\varphi$ ein Normalteiler von G und $\operatorname{Bild}\varphi$ eine Untergruppe von H. Weiter definiert*

$$\widetilde{\varphi} : G/\operatorname{Kern}\varphi \to H : g \operatorname{Kern}\varphi \mapsto \varphi(g)$$

einen Monomorphismus und φ faktorisiert

$$\varphi = \widetilde{\varphi} \circ \nu_{\operatorname{Kern}\varphi},$$

d.h. das Diagramm

$$\begin{array}{ccc} G & \xrightarrow{\varphi} & H \\ {\scriptstyle \nu_{\operatorname{Kern}\varphi}}\searrow & & \nearrow{\scriptstyle \widetilde{\varphi}} \\ & G/\operatorname{Kern}\varphi & \end{array}$$

kommutiert. Insbesondere sind $G/\operatorname{Kern}\varphi$ und $\operatorname{Bild}\varphi \leq H$ isomorph.

Beweis.

(1) Diesen Teil haben wir bereits bewiesen in Satz 8.2.7.

(2) Genauso wie im Homomorphiesatz für Mengen zeigt man, dass $\widetilde{\varphi}$ wohldefiniert und injektiv ist. Es bleibt die Homomorphieeigenschaft von $\widetilde{\varphi}$ zu überprüfen: Setze $N := \operatorname{Kern}\varphi$ und seien $g, h \in G$. Dann gilt:

$$\widetilde{\varphi}(gNhN) = \widetilde{\varphi}(ghN) = \varphi(gh) = \varphi(g)\varphi(h) = \widetilde{\varphi}(gN)\widetilde{\varphi}(hN).$$

Dass $\nu_{\operatorname{Kern}\varphi}$ ein Epimorphismus ist, wissen wir bereits. Die Komposition $\widetilde{\varphi} \circ \nu_{\operatorname{Kern}\varphi} = \varphi$ rechnet man leicht nach. □

Satz 8.2.12. *(**Noetherscher Isomorphiesatz**) Sei N ein Normalteiler von G und $U \leq G$. Dann ist $UN \leq G$, und $N \cap U \trianglelefteq U$ und es gilt*

$$UN/N \cong U/U \cap N.$$

Beweis. Wegen $NU = UN$ folgt $UN \leq G$. Offenbar ist N auch ein Normalteiler von $UN \leq G$ und somit

$$\mu : U \to UN/N : u \mapsto uN$$

ein Homomorphismus mit Kern $U \cap N$ und Bild UN/N. Die Behauptung folgt aus dem Homomorphiesatz. □

Definition 8.2.13. Eine Gruppe G heißt **semidirektes Produkt**, falls es einen Normalteiler $N \trianglelefteq G$ und eine Untergruppe $U \leq G$ gibt mit

(1) $NU = G$ und

(2) $N \cap U = \{1\}$.

In Zeichen $G = N \rtimes U$.

Beispiel 8.2.14. $S_4 = V_4 \rtimes S_3$ wobei es egal ist, welche der 4 Untergruppen S_3 in S_4 gewählt wird. Beachte, dass alle Elemente $\neq 1$ in V_4 keine Fixpunkte auf $\{1, 2, 3, 4\}$ haben, also ist $V_4 \cap S_3 = 1$.

Beispiel 8.2.15.

$$\mathrm{Aff}_n(K) := \left\{ \left(\begin{array}{c|c} A & a \\ \hline 0 & 1 \end{array} \right) | a \in K^{n\times 1}, A \in \mathrm{GL}_n(K) \right\} = K^{n\times 1} \rtimes \mathrm{GL}_n(K)$$

Dabei ist $K^{n\times 1} = \left\{ \left(\begin{array}{c|c} I_n & a \\ \hline 0 & 1 \end{array} \right) \mid a \in K^{n\times 1} \right\}$ der Kern des Gruppenhomomorphismus

$$\mathrm{Aff}_n(K) \to \mathrm{GL}_n(K), \left(\begin{array}{c|c} A & a \\ \hline 0 & 1 \end{array} \right) \mapsto A$$

und $\mathrm{GL}_n(K)$ die Untergruppe $\mathrm{GL}_n(K) = \left(\begin{array}{c|c} A & 0 \\ \hline 0 & 1 \end{array} \right) \leq \mathrm{Aff}_n(K)$.

Bemerkung 8.2.16. In einem semidirekten Produkt $G = N \rtimes U$ hat jedes Element eine eindeutige Darstellung als nu mit $n \in N$, $u \in U$. Es gilt

$$n_1 u_1 n_2 u_2 = n_1 (u_1 n_2 u_1^{-1}) u_1 u_2$$

Kapitel 9

Geometrie

9.1 Affine Geometrie

9.1.a Der affine Raum

In der affinen Geometrie hat man einen Punktraum, dessen Punkte in Bijektion zu einem Vektorraum stehen, welcher in bestimmter Weise auf dem Punktraum (durch Translationen oder Verschiebungen) operiert. Der wesentliche Unterschied zum Vektorraum besteht darin, dass kein Punkt (Nullpunkt) mehr ausgezeichnet ist. Begriffe wie Geraden, Ebenen etc. lassen sich leicht als sogenannte affine Unterräume definieren.

Definition 9.1.1. Sei $\mathcal{V}$ ein K-Vektorraum. Ein **affiner Raum** über $\mathcal{V}$ ist eine nicht leere Menge $\mathscr{A}$, genannt Punktmenge, auf der $\mathcal{V}$ scharf transitiv [1] operiert. Genauer ist ein affiner Raum ein Tripel $(\mathscr{A}, \mathcal{V}, \tau)$, wobei

$$\tau : \mathcal{V} \times \mathscr{A} \to \mathscr{A} : (V, P) \mapsto \tau_V(P)$$

eine scharf transitive Operation des Vektorraumes $\mathcal{V}$ auf dem Punktraum $\mathscr{A}$ ist. Die Abbildung $\tau_V : \mathscr{A} \to \mathscr{A}$ heißt die **Translation** um den Vektor V von $\mathscr{A}$. Der Vektorraum $\mathcal{V}$ wird auch als **Translationsraum** von $\mathscr{A}$ bezeichnet: $\mathcal{T}(\mathscr{A}) := \mathcal{V}$. (Bezeichnung: Oft schreiben wir $V + P$ oder $P + V$ anstatt $\tau_V(P)$. Diese Schreibweise soll nicht implizieren, dass $\mathscr{A} = \mathcal{V}$ ist.)

Jeder Vektorraum $\mathcal{V}$ ist ein affiner Raum mit Translationsraum $\mathcal{V}$. Das Modell $\mathscr{A} = \mathcal{V} = \mathcal{T}(\mathscr{A})$ hat den Nachteil, dass nicht zwischen den Punkten (Elementen von $\mathscr{A}$) und den Vektoren[2] unterschieden wird. Daher bevorzugen wir das folgende Modell:

Beispiel 9.1.2. (Standardbeispiel) Ist $\widetilde{\mathcal{V}}$ ein K-Vektorraum mit nicht verschwindender Linearform $\varphi : \widetilde{\mathcal{V}} \to K$. Setze $\mathcal{V} := \mathrm{Kern}(\varphi)$ und $\mathscr{A}(\varphi) := \varphi^{-1}(\{1\})$. Dann ist $(\mathscr{A}(\varphi), \mathcal{V}, \tau)$ mit

$$\tau : \mathcal{V} \times \mathscr{A}(\varphi) \to \mathscr{A}(\varphi) : (V, P) \mapsto \tau_V(P) := V + P \text{ in } \widetilde{\mathcal{V}} \text{ gerechnet}$$

ein affiner Raum.
Wir setzen speziell für $\widetilde{\mathcal{V}} = K^{(n+1)\times 1}$ und $\varphi \in (K^{(n+1)\times 1})^*$ die Projektion auf die letzte Komponente:

$$\mathscr{A}_n(K) := \mathscr{A}(\varphi) = \left\{ \begin{pmatrix} X \\ 1 \end{pmatrix} | X \in K^{n\times 1} \right\}$$

und nennen ihn den n-dimensionalen affinen Standardraum. Genau genommen ist

$$\mathcal{T}(\mathscr{A}_n(K)) = \left\{ \begin{pmatrix} X \\ 0 \end{pmatrix} | X \in K^{n\times 1} \right\},$$

[1] Da $\mathcal{V}$ eine abelsche Gruppe ist, sind die Begriffe „scharf transitive Operation“ einerseits sowie „treue und transitive Operation“ andererseits äquivalent.

[2] vector (lat.): jemand, der trägt, zieht oder befördert

was wir aber mit dem Vektorraum $K^{n\times 1}$ identifizieren.

Bemerkung 9.1.3. Sei $\mathscr{A}$ ein affiner Raum über dem K-Vektorraum $\mathcal{V}$.

(1) Für jeden Punkt $P_0 \in \mathscr{A}$ ist
$$\mathcal{V} \to \mathscr{A} : V \mapsto \tau_V(P_0)$$
eine Bijektion.

(2) Für jedes Punktepaar $(P, Q) \in \mathscr{A}^2$ gibt es genau einen Vektor $V \in \mathcal{V}$ mit $\tau_V(P) = Q$. Bezeichnung: $V =: \overrightarrow{PQ}$.

Beweis. Spezialfall von Satz 8.1.12 □

Ende Vorl. 20

Übung 9.1.1. Zeige für $P, Q, P', Q' \in \mathscr{A}$ gilt $\overrightarrow{PQ} = \overrightarrow{P'Q'}$ genau dann, wenn $\overrightarrow{PP'} = \overrightarrow{QQ'}$. (Hinweis: Skizze!)

Wir kommen zur Definition affiner Teilräume.

Definition 9.1.4. Sei $(\mathscr{A}, \mathcal{V}, \tau)$ affiner Raum über dem K-Vektorräum $\mathcal{V}$. Eine Teilmenge $\mathscr{A}' \subseteq \mathscr{A}$ heißt **affiner Teilraum** von $\mathscr{A}$, falls ein Teilvektorraum $\mathcal{W} \leq \mathcal{V}$ existiert, so dass $(\mathscr{A}', \mathcal{W}, \tau_{|\mathcal{W}\times\mathscr{A}'})$ ein affiner Raum über $\mathcal{W}$ ist.

Bemerkung 9.1.5. Sei $\mathscr{A}$ affiner Raum über $\mathcal{V} := \mathcal{T}(\mathscr{A})$.

(1) Der Translationsraum eines affinen Teilraums $\mathscr{A}'$ von $\mathscr{A}$ ist eindeutig bestimmt, genauer
$$\mathcal{T}(\mathscr{A}') = \{\overrightarrow{PQ} \mid P, Q \in \mathscr{A}'\}.$$

(2) Zu jedem $\mathcal{W} \leq \mathcal{V}$ und jedem $P \in \mathscr{A}$ gibt es genau einen affinen Teilraum $\mathscr{A}'$ von $\mathscr{A}$ mit $P \in \mathscr{A}'$ und Translationsraum $\mathcal{T}(\mathscr{A}') = \mathcal{W}$, nämlich $P + \mathcal{W} = \mathcal{W} + P := \tau(\mathcal{W} \times \{P\})$, die Bahn von P unter $\mathcal{W}$.

Beweis.

(1) Sofort aus 9.1.3.

(2) Existenz: Verifiziere Eigenschaften für $P + \mathcal{W}$. Eindeutigkeit analog zu 1. □

Bemerkung 9.1.6. Der Schnitt affiner Teilräume eines affinen Raumes $\mathscr{A}$ ist entweder leer oder wieder ein affiner Teilraum.
Sind $\mathscr{A}_i$ $(i \in I)$ affine Teilräume von $\mathscr{A}$ und ist $P \in \cap_{i\in I}\mathscr{A}_i$, so ist
$$\cap_{i\in I}\mathscr{A}_i = P + \cap_{i\in I}\mathcal{T}(\mathscr{A}_i) = \{\tau_V(P) \mid V \in \cap_{i\in I}\mathcal{T}(\mathscr{A}_i)\}.$$

Ist $\emptyset \neq M \subset \mathscr{A}$ eine Teilmenge, so sei das **affine Erzeugnis** von M der kleinste affine Teilraum, der M enthält: $\langle M\rangle_a := \bigcap_{M\subset\mathcal{B}\leq\mathscr{A}} \mathcal{B}$.

Wie sieht das affine Erzeugnis einer zweipunktigen Teilmenge von $\mathscr{A}$ aus?

9.1.b Affine Abbildungen

Nun kommen wir zur Definition affiner Abbildungen. Diese bilden einen ganz wesentlichen Bestandteil der Definition des affinen Raumes, weil wir sonst nicht wissen, wie wir vergleichen können. Es liefert auch eine neue Charakterisierung der affinen Teilräume: Die nicht leeren Fasern affiner Abbildungen werden die affinen Teilräume sein.

Definition 9.1.7. Seien $\mathscr{A}, \mathscr{A}'$ affine Räume über den K-Vektorräumen $\mathcal{V}, \mathcal{V}'$.
Eine Abbildung $f : \mathscr{A} \to \mathscr{A}'$ heißt **affine Abbildung**, falls eine lineare Abbildung $\overline{f} : \mathcal{V} \to \mathcal{V}'$ existiert mit $\overrightarrow{f(P)f(Q)} = \overline{f}(\overrightarrow{PQ})$ für alle $P, Q \in \mathscr{A}$. Die lineare Abbildung $\overline{f}$ heißt auch der **lineare Anteil** von f.

Bemerkung 9.1.8. Seien $\mathscr{A}$, $\mathscr{A}'$ affine Räume mit Translationsvektorraum $\mathcal{V} = \mathcal{T}(\mathscr{A})$ und $\mathcal{V}' = \mathcal{T}(\mathscr{A}')$. Sei $P_0 \in \mathscr{A}$ fest gewählt.

(1) Jede affine Abbildung $f : \mathscr{A} \to \mathscr{A}'$ ist eindeutig festgelegt durch ihren linearen Anteil $\overline{f}$ und $f(P_0)$: Ist nämlich $f(P_0) =: Q_0 \in \mathscr{A}'$ so ist für $P \in \mathscr{A}$ und $V := \overrightarrow{P_0P} \in \mathcal{V}$ (so, dass $P = \tau_V(P_0)$)

$$\overrightarrow{Q_0f(P)} = \overrightarrow{f(P_0)f(P)} = \overline{f}(V) \text{ also } f(P) = \tau_{\overline{f}(V)}(Q_0).$$

(2) Für jeden Punkt $Q_0 \in \mathscr{A}'$ und jede lineare Abbildung $\varphi : \mathcal{V} \to \mathcal{V}'$ gibt es genau eine affine Abbildung $f : \mathscr{A} \to \mathscr{A}'$ mit $f(P_0) = Q_0$ und $\overline{f} = \varphi$.
Es ist f injektiv (surjektiv, bijektiv), genau dann wenn $\overline{f}$ injektiv (surjektiv, bijektiv) ist.

Übung 9.1.2. Translationen sind affine Abbildungen, deren linearer Anteil die Identität des Translationsraumes ist. Sie sind auch die einzigen affinen Abbildungen eines affinen Raumes in sich mit dieser Eigenschaft.

Beispiel 9.1.9 (Affine Abbildungen von $\mathscr{A}_n(K)$)**.** Wähle $P_0 = (0, \ldots, 0|1)^{tr} \in \mathscr{A}_n(K)$. Die affine Abbildung f mit linearem Anteil ${}^S\overline{f}^S = A$ (bzgl. der Standardbasis S von $\mathcal{T}(\mathscr{A}_n(K)) = K^{n\times 1}$) und $f(P_0) = Q_0 = (b_1, \ldots, b_n|1)^{tr}$ ist gegeben durch Matrixmultiplikation mit $\left(\begin{array}{c|c} A & b \\ \hline 0 & 1 \end{array}\right)$.

Satz 9.1.10.

(1) Kompositionen affiner Abbildungen sind affin: Sind $\mathscr{A}, \mathscr{A}', \mathscr{A}''$ affine Räume über K-Vektorräumen mit affinen Abbildungen $f : \mathscr{A} \to \mathscr{A}'$ und $f' : \mathscr{A}' \to \mathscr{A}''$, so ist $f' \circ f : \mathscr{A} \to \mathscr{A}''$ affin mit $\overline{f' \circ f} = \overline{f'} \circ \overline{f}$.

(2) Ist $f : \mathscr{A} \to \mathscr{A}'$ affin und bijektiv, so ist $f^{-1} : \mathscr{A}' \to \mathscr{A}$ ebenfalls affin mit $\overline{f^{-1}} = \overline{f}^{-1}$. (Man sagt, f ist ein **affiner Isomorphismus**.*) Insbesondere ist*

$$\mathrm{Aff}(\mathscr{A}) := \{f : \mathscr{A} \to \mathscr{A} | f \text{ affin und bijektiv } \}$$

eine Gruppe (Untergruppe von $S_{\mathscr{A}}$, der symmetrischen Gruppe von $\mathscr{A}$), genannt die **affine Gruppe** *von $\mathscr{A}$, und*

$$\mathrm{Aff}(\mathscr{A}) \to \mathrm{GL}(\mathcal{T}(\mathscr{A})) : f \mapsto \overline{f}$$

ein Homomorphismus von Gruppen.

(3) Ist $f : \mathscr{A} \to \mathscr{A}'$ affin und $\mathscr{A}''$ ein affiner Teilraum von $\mathscr{A}$, so ist $f(\mathscr{A}'')$ ein affiner Teilraum von $\mathscr{A}'$ mit $\mathcal{T}(f(\mathscr{A}'')) = \overline{f}(\mathcal{T}(\mathscr{A}''))$.

(4) Ist $f : \mathscr{A} \to \mathscr{A}'$ affin und $\mathscr{A}''$ ein affiner Teilraum von $\mathscr{A}'$, so ist $f^{-1}(\mathscr{A}'')$ leer oder ein affiner Teilraum von $\mathscr{A}$ mit $\mathcal{T}(f^{-1}(\mathscr{A}'')) = \overline{f}^{-1}(\mathcal{T}(\mathscr{A}''))$.

Beweis.

(1) Für $P, Q \in \mathscr{A}$ ist

$$\overrightarrow{(f' \circ f)(P)(f' \circ f)(Q)} = \overrightarrow{f'(f(P))f'(f(Q))} =$$
$$\overline{f'}(\overrightarrow{f(P)f(Q)}) = (\overline{f'} \circ \overline{f})(\overrightarrow{PQ}).$$

(2) Wegen der Identifikation von $\mathscr{A}$ mit $\mathcal{T}(\mathscr{A})$ und $\mathscr{A}'$ mit $\mathcal{T}(\mathscr{A}')$ ist klar, dass $\overline{f} : \mathcal{T}(\mathscr{A}) \to \mathcal{T}(\mathscr{A}')$ bijektiv ist. (Genauer Beweis: Übung!). Zeige nur noch

$$\overrightarrow{f^{-1}(P')f^{-1}(Q')} = \overline{f}^{-1}(\overrightarrow{P'Q'})$$

für alle $P', Q' \in \mathscr{A}'$. Dies ist aber äquivalent zu

$$\overline{f}(\overrightarrow{f^{-1}(P')f^{-1}(Q')}) = \overrightarrow{P'Q'}.$$

(3) Übung.

(4) Leicht mit 9.1.5.(2). □

Wir können etwas unscharf sagen, dass affine Geometrie das Studium von Eigenschaften ist, welche unter affinen Isomorphismen erhalten bleiben, oder auch das Studium der Invarianten der affinen Gruppe bei diversen Operationen. Hier ein Anfang: Die Dimension.

Satz 9.1.11. *Zwei affine Räume $\mathscr{A}$ und $\mathscr{A}'$ über demselben Körper K sind genau dann affin isomorph, wenn $\operatorname{Dim}\mathcal{T}(\mathscr{A}) = \operatorname{Dim}\mathcal{T}(\mathscr{A}')$. Man nennt $\operatorname{Dim}\mathscr{A} := \operatorname{Dim}\mathcal{T}(\mathscr{A})$ die* **Dimension** *des affinen Raumes $\mathscr{A}$. Insbesondere ist $\mathscr{A}$ affin isomorph zu $\mathcal{A}_n(K)$ für $n = \operatorname{Dim}\mathscr{A}$. Ein affiner Isomorphismus $\mathscr{A} \to \mathcal{A}_n(K)$ heißt* **affines Koordinatensystem**.

Die Idee des Koordinatensystems geht zurück auf DESCARTES,1596-1650, der hierdurch die Algebra und Analysis als Hilfsmittel der Geometrie zugänglich machte.

Beweis. Ist $f : \mathscr{A} \to \mathscr{A}'$ ein affiner Isomorphismus, so ist $\overline{f} : \mathcal{T}(\mathscr{A}) \to \mathcal{T}(\mathscr{A}')$ ein Vektorraumisomorphismus, also $\operatorname{Dim}\mathcal{T}(\mathscr{A}) = \operatorname{Dim}\mathcal{T}(\mathscr{A}')$. Umgekehrt, sei $\varphi : \mathcal{T}(\mathscr{A}) \to \mathcal{T}(\mathscr{A}')$ ein Vektorraumisomorphismus. Offenbar ist für jedes beliebige, fest gewählte $P_0 \in \mathscr{A}$ die Abbildung $\mathscr{A} \to \mathcal{T}(\mathscr{A}) : P \mapsto \overrightarrow{P_0P}$ ein affiner Isomorphismus (Beweis: Übung). Also erhält man durch Komposition einen affinen Isomorphismus von $\mathscr{A}$ auf $\mathscr{A}'$. (Man zeige als Übungsaufgabe: Dieser Isomorphismus ist gegeben durch $P \mapsto P'_0 + \varphi(\overrightarrow{P_0P})$, wo $P'_0 \in \mathscr{A}'$ beliebig, aber fest gewählt ist.) □

Somit ist die Dimension eine affine Invariante.

Definition 9.1.12. Sei $\mathscr{A}$ ein affiner Raum über $\mathcal{T}(\mathscr{A}) = \mathcal{V}$ mit affinen Teilräumen $\mathscr{A}', \mathscr{A}''$.

(1) Die Teilräume heißen **parallel**, falls $\mathcal{T}(\mathscr{A}') = \mathcal{T}(\mathscr{A}'')$.

(2) Sie heißen **schwach parallel**, falls $\mathcal{T}(\mathscr{A}') \subseteq \mathcal{T}(\mathscr{A}'')$ oder $\mathcal{T}(\mathscr{A}'') \subseteq \mathcal{T}(\mathscr{A}')$.

(3) Sie heißen **windschief**, falls $\mathscr{A}' \cap \mathscr{A}'' = \emptyset$ und $\mathcal{T}(\mathscr{A}'') \cap \mathcal{T}(\mathscr{A}') = \{0\}$.

Übung 9.1.3. Sei $\mathscr{A}$ ein affiner Raum über dem Vektorraum $\mathcal{V}$. Zeige, dass Parallelität eine Äquivalenzrelation auf der Menge aller affinen Teilräume von $\mathscr{A}$ ist. Zeige weiter, dass die Äquivalenzklasse mit zugehörigem Teilraum $\mathcal{W} \leq \mathcal{T}(\mathscr{A})$ wiederum einen affinen Raum $\mathscr{A}/\mathcal{W}$ bildet, und zwar mit Translationsraum $\mathcal{V}/\mathcal{W}$. Man nennt $\mathscr{A}/\mathcal{W}$ auch den Bahnenraum von $\mathscr{A}$ mod $\mathcal{W}$. (Beachte: $\mathcal{V}$ operiert zwar auch transitiv auf $\mathscr{A}/\mathcal{W}$, aber nicht treu, es sei denn $\mathcal{W} = \{0\}$.)

Übung 9.1.4. Ist $\mathscr{A}$ ein affiner Raum über dem K-Vektorraum $\mathcal{V}$, und sind $\mathcal{U}, \mathcal{W} \leq \mathcal{V}$ Teilräume, so gilt für die Abbildung

$$\varphi : \mathscr{A} \to \mathscr{A}/\mathcal{U} \times \mathscr{A}/\mathcal{W} : P \mapsto (P + \mathcal{U}, P + \mathcal{W}),$$

(1) φ ist injektiv genau dann, wenn $\mathcal{U} \cap \mathcal{W} = \{0\}$.

(2) φ ist surjektiv genau dann, wenn $\mathcal{U} + \mathcal{W} = \mathcal{V}$.

Bemerkung 9.1.13. Parallelität und schwache Parallelität von affinen Teilräumen bleiben unter affinen Abbildungen erhalten. Die Eigenschaft, windschief zu sein, bleibt unter injektiven affinen Abbildungen erhalten. Wie steht es mit Urbildern?

Wir wollen uns ansehen, wie in den verschiedenen Modellen für affine Räume, die wir gesehen haben, die affinen Abbildungen aussehen und dargestellt werden.

Beispiel 9.1.14. Seien $\widetilde{\mathcal{V}}, \varphi, \mathcal{V} := \mathrm{Kern}(\varphi), \mathcal{A}(\varphi) := \varphi^{-1}(\{1\})$ wie in Beispiel 9.1.2. Entsprechend nehmen wir einen zweiten affinen Raum mit den Daten $\widetilde{\mathcal{W}}, \psi, \mathcal{W} := \mathrm{Kern}(\psi)$ und $\mathcal{A}(\psi) = \psi^{-1}(\{1\})$. Dann ist eine affine Abbildung $f : \mathcal{A}(\varphi) \to \mathcal{A}(\psi)$ nichts anderes als die Einschränkung einer linearen Abbildung $\alpha : \widetilde{\mathcal{V}} \to \widetilde{\mathcal{W}}$, welche $\mathcal{A}(\varphi)$ in $\mathcal{A}(\psi)$ abbildet, d.h. für die $\psi \circ \alpha = \varphi$. Wir haben also das folgende kommutative Diagramm:

$$\begin{array}{ccc} \mathcal{A}(\varphi) & \xrightarrow{f} & \mathcal{A}(\psi) \\ \downarrow & & \downarrow \\ \widetilde{\mathcal{V}} & \xrightarrow{\alpha} & \widetilde{\mathcal{W}} \\ \varphi \downarrow & & \downarrow \psi \\ K & = & K \end{array}$$

Übung 9.1.5. α legt f eindeutig fest und umgekehrt.
Wichtiger Spezialfall: $\mathrm{Aff}(\mathcal{A}_n(K))$ kann mit der Matrixgruppe

$$\mathrm{Aff}_n(K) := \left\{ \left(\begin{array}{c|c} a & t \\ \hline 0 & 1 \end{array}\right) | a \in \mathrm{GL}_n(K), t \in K^{n\times 1} \right\} \leq \mathrm{GL}_{n+1}(K)$$

identifiziert werden, die durch Linksmultiplikation auf $\mathcal{A}_n(K)$ operiert. Man beachte, dass $\mathrm{Aff}_n(K)$ schon als Stabilisator eines Kovektors[3] als Untergruppe von $\mathrm{GL}_{n+1}(K)$ früher im Beispiel 8.1.30 vorkam.

Ende Vorl. 21

Bemerkung 9.1.15. In $\mathrm{Aff}_n(K)$ gelten:

$$\left(\begin{array}{c|c} a & t \\ \hline 0 & 1 \end{array}\right)\left(\begin{array}{c|c} b & s \\ \hline 0 & 1 \end{array}\right) = \left(\begin{array}{c|c} ab & as+t \\ \hline 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{c|c} a & t \\ \hline 0 & 1 \end{array}\right)^{-1} = \left(\begin{array}{c|c} a^{-1} & -a^{-1}t \\ \hline 0 & 1 \end{array}\right)$$

und

$$\boxed{\left(\begin{array}{c|c} a & t \\ \hline 0 & 1 \end{array}\right)\left(\begin{array}{c|c} I_n & s \\ \hline 0 & 1 \end{array}\right)\left(\begin{array}{c|c} a & t \\ \hline 0 & 1 \end{array}\right)^{-1} = \left(\begin{array}{c|c} I_n & as \\ \hline 0 & 1 \end{array}\right)}$$

Der Homomorphismus "linearen Anteil nehmen" ist gegeben durch

$$\mathrm{Aff}_n(K) \to \mathrm{GL}_n(K) : \left(\begin{array}{c|c} a & t \\ \hline 0 & 1 \end{array}\right) \mapsto a$$

[3]Sprich, eines Vektors im Dualraum, auch Funktional genannt.

9.1.c Das Invarianzprinzip der affinen Geometrie

Bemerkung 9.1.16. $\mathrm{Aff}(\mathscr{A})$ operiert transitiv auf $\mathscr{A}$ und hat genau 2 Bahnen auf $\mathscr{A} \times \mathscr{A}$.

Beweis. $\mathcal{T}(\mathscr{A}) \leq \mathrm{Aff}(\mathscr{A})$ operiert scharf transitiv, also insbesondere transitiv auf $\mathscr{A}$, daher auch $\mathrm{Aff}(\mathscr{A})$. Ist $P_0 \in \mathscr{A}$, so ist der Stabilisator $\mathrm{Stab}_{\mathrm{Aff}(\mathscr{A})}(P_0)$ isomorph zu $\mathrm{GL}(\mathcal{T}(\mathscr{A}))$, vermöge der Abbildung

$$\mathrm{Stab}_{\mathrm{Aff}(\mathscr{A})}(P_0) \ni f \mapsto \left(\overline{f} : \overrightarrow{P_0P} \mapsto \overrightarrow{P_0f(P)}\right) \in \mathrm{GL}(\mathcal{T}(\mathscr{A})).$$

Also hat $\mathrm{Stab}_{\mathrm{Aff}(\mathscr{A})}(P_0)$ zwei Bahnen auf $\mathscr{A}$, nämlich $\{P_0\}$ und $\mathscr{A} \setminus \{P_0\}$. □

In unserem Standardmodell $\mathscr{A}_n(K)$ kann man Œ annehmen, dass $P_0 = (0, ..., 0, 1)^{tr}$. Dann ist

$$\mathrm{Stab}_{\mathrm{Aff}_n(K)}(P_0) = \left\{ \left(\begin{array}{c|c} a & 0 \\ \hline 0 & 1 \end{array} \right) | a \in \mathrm{GL}_n(K) \right\} \cong \mathrm{GL}_n(K)$$

Bei der Operation auf Tripeln bekommen wir die ersten geometrischen Invarianten.

Definition 9.1.17.

(1) $P \in \mathscr{A}^n$ heißt **affin unabhängig**, falls für jeden affinen Raum $\mathscr{A}'$ über K und jedes Tupel $Q \in (\mathscr{A}')^n$ eine affine Abbildung $f : \mathscr{A} \to \mathscr{A}'$ existiert, mit $f \circ P = Q$, d.h. $f(P_i) = Q_i$ für $i = 1, \ldots, n$. Ein maximal affin unabhängiges System P in $\mathscr{A}$ heißt auch **affine Basis** von $\mathscr{A}$.

(2) $P \in \mathscr{A}^n$ heißt **kollinear** bzw. **komplanar**, falls $\mathrm{Dim}\langle P\rangle_a \leq 1$ bzw. ≤ 2 gilt.

Bemerkung 9.1.18. Für $P \in \mathscr{A}^n$ sind folgende Aussagen äquivalent:

(1) P ist affin unabhängig.

(2) $\mathrm{Dim}\langle P\rangle_a = n - 1$.

(3) $(\overrightarrow{P_nP_1}, \ldots, \overrightarrow{P_nP_{n-1}}) \in \mathcal{T}(\mathscr{A})^{n-1}$ ist linear unabhängig.

(4) Die affine Abbildung $f : \mathscr{A}_{n-1}(K) \to \langle P\rangle_a : \widetilde{E_i} := \left(\frac{E_i}{1}\right) \mapsto P_i, \widetilde{E_n} := \left(\frac{0}{1}\right) \mapsto P_n$ definiert einen affinen Isomorphismus.

Beweis. Vorbemerkung: Es gilt $\mathcal{T}(\langle P\rangle_a) = \langle \overrightarrow{P_nP_1}, \ldots, \overrightarrow{P_nP_{n-1}}\rangle \leq \mathcal{T}(\mathscr{A}) =: \mathcal{V}$.

1 ⇒ 4: Es liegt eine affine Abbildung f vor mit $f(\widetilde{E}_n) = P_n$ und

$$\overline{f} : K^{(n-1)\times 1} \to \mathcal{T}(\langle P\rangle_a), E_i \mapsto \overrightarrow{P_nP_i}.$$

Aus der Definition der affinen Unabhängigkeit bekommt man eine affine Abbildung $\mathscr{A} \to \mathscr{A}_{n-1}(K)$, die P_i auf $\widetilde{E_i}$ abbildet. Die Einschränkung dieser Abbildung auf $\langle P\rangle_a$ liefert die Inverse, d.h. es liegt ein affiner Isomorphismus vor.

4 ⇒ 2: Die Dimension ist eine Invariante für affine Isomorphismen.

2 ⇒ 3: Es ist $n - 1 = \mathrm{Dim}\langle P\rangle_a = \mathrm{Dim}\,\mathcal{T}(\langle P\rangle_a) = \mathrm{Dim}\langle \overrightarrow{P_nP_1}, \ldots, \overrightarrow{P_nP_{n-1}}\rangle$.

$3 \Rightarrow 1$: Sei $\mathscr{A}'$ irgendein affiner Raum über dem K-Vektorraum $\mathcal{V}'$ und $Q \in (\mathscr{A}')^n$. Es existiert eine lineare Abbildung $\varphi : \mathcal{V} \to \mathcal{V}'$ mit $\varphi(\overrightarrow{P_nP_i}) = \overrightarrow{Q_nQ_i}$ für $i = 1, \ldots, n-1$. Also gibt es genau eine affine Abbildung $f : \mathscr{A} \to \mathscr{A}'$ mit $\overline{f} = \varphi$ und $f(P_n) = Q_n$. Für diese gilt offenbar $f(P_i) = Q_i$ für $i = 1, \ldots, n$. □

Bemerkung 9.1.19. Ist $\mathscr{A} = \varphi^{-1}(1)$ für $\varphi \in \widetilde{\mathcal{V}}^*$, so ist $P \in \mathscr{A}^n$ affin unabhängig genau dann, wenn $P \in \widetilde{\mathcal{V}}^n$ linear unabhängig ist. Die affinen Basen von $\mathscr{A}$ sind also genau die Basen von $\widetilde{\mathcal{V}}$, die in $\mathscr{A} = \varphi^{-1}(1) \subset \widetilde{\mathcal{V}}$ enthalten sind.

Sofort klar, etwa aus Bemerkung 9.1.18.(3), ist nun die folgende Bemerkung:

Bemerkung 9.1.20.

(1) Affine Abhängigkeit von Tupeln bleibt erhalten unter beliebigen affinen Abbildungen.

(2) Affine Unabhängigkeit bleibt unter injektiven affinen Abbildungen erhalten.

Satz 9.1.21. *Sei $\mathscr{A}$ ein affiner Raum über einem endlich erzeugten K-Vektorraum. Dann operiert $\mathrm{Aff}(\mathscr{A})$ scharf transitiv auf der Menge der affinen Basen von $\mathscr{A}$. Letztere bilden eine der Bahnen von $\mathrm{Aff}(\mathscr{A})$ auf $\mathscr{A}^{n+1}$, wo $n = \mathrm{Dim}\,\mathscr{A}$.*

Beweis. Sofort aus Satz 9.1.11 und Bemerkung 9.1.18. □

Satz 9.1.22.

(1) $\mathrm{Aff}(\mathscr{A})$ operiert transitiv auf der Menge $\mathscr{A}^3_{generisch}$ der affin unabhängigen Tripel in $\mathscr{A}^3$ (nicht entartete Dreiecke), falls $\mathrm{Dim}(\mathscr{A}) \geq 2$.

(2) Eine trennende Invariante für die Operation von $\mathrm{Aff}(\mathscr{A})$ auf der Menge $\mathscr{A}^3_{spez} := \{P = (P_1, P_2, P_3) \in \mathscr{A}^3 | P_1 \neq P_2, P \text{ kollinear } \}$ ist das **Teilverhältnis**. *Dabei ist das Teilverhältnis $\mathrm{TV}(P_1, P_2, P_3)$ definiert als das eindeutige $a \in K$ mit $\overrightarrow{P_1P_3} = a\overrightarrow{P_1P_2}$.*

Beweis.

(1) Wir können oBdA in $\mathscr{A} = \mathscr{A}(\varphi) \subset \widetilde{\mathcal{V}}$ arbeiten. Offenbar hat $\widetilde{\mathcal{V}}$ eine Basis $B \in \mathscr{A}^{n+1}$ und jedes affin unabhängige Tripel $P \in \mathscr{A}^3$ kann seinerseits zu einer Basis $\widehat{P} \in \mathscr{A}^{n+1}$ von $\widetilde{\mathcal{V}}$ ergänzt werden. Es genügt nun zu zeigen, dass ein $f \in \mathrm{Aff}(\mathscr{A})$ existiert, mit $f((B_1, B_2, B_3)) = P$. Dies ist aber klar, denn ein solches f wird induziert von der eindeutigen linearen Abbildung von $\widetilde{\mathcal{V}}$, die B auf $\widehat{P}$ abbildet.

(2) Dass eine Invariante vorliegt, ist sofort klar aus der Definition einer affinen Abbildung. Um zu zeigen, dass sie die Bahnen trennt, gehen wir wieder von der Situation des Beweises von (1) aus mit der Basis B. Sei $P \in \mathscr{A}^3_{spez}$ mit Teilverhältnis $a \in K$. Es genügt zu zeigen, dass ein $f \in \mathrm{Aff}(\mathscr{A})$ existiert mit $f((B_1, B_2, B_1 + a(B_2 - B_1))) = P$. Zu diesem Zweck ergänzt man (P_1, P_2) zu einer Basis $\widehat{P} \in \mathscr{A}^{n+1}$ von $\widetilde{\mathcal{V}}$. Die lineare Abbildung, die B auf $\widehat{P}$ abbildet, induziert den gewünschten affinen Automorphismus. □

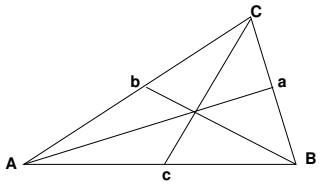
Aus dem letzten Beweis erhalten wir eine Folgerung, die eine sehr anschauliche Vorstellung von der affinen Gruppe liefert.

Folgerung 9.1.23. *Sei $\mathrm{Dim}(\mathscr{A}) = n$. Dann operiert $\mathrm{Aff}(\mathscr{A})$ transitiv auf $\mathscr{A}^k_{generisch} := \{P \in \mathscr{A}^k | P \text{ affin unabhängig}\}$, der Menge der affin unabhängigen k-Tupel ($(k-1)$-Simplizes) für $1 \leq k \leq n+1$. Im Falle $k = n+1$ ist der Stabilisator eines solchen Tupels trivial, d.h. in diesem Falle ist die Operation scharf transitiv, vgl. Satz 9.1.21.*

Wir wollen jetzt als Beispiel einen geometrischen Satz beweisen.

Ende Vorl. 22

Satz 9.1.24. *Sei K ein Körper mit $6 \cdot 1 \neq 0$, $\mathscr{A}$ ein affiner Raum über K, $(A, B, C) \in \mathscr{A}^3_{gen}$. Dann schneiden sich die Seitenhalbierenden des nicht-entarteten Dreiecks (A, B, C) in einem Punkt S, so dass das Teilverhältnis $\mathrm{TV}(A, a, S) = 2/3$ ist, wo a der Mittelpunkt der Seite (C, B) ist.*



Beweis. Zunächst einmal definieren wir Seitenhalbierende:

$$s_a = \langle A, a\rangle_a, s_b = \langle B, b\rangle_a, s_c = \langle C, c\rangle_a$$

wobei $a = B + \frac{1}{2}\overrightarrow{BC}$, $b = A + \frac{1}{2}\overrightarrow{AC}$ und $c = A + \frac{1}{2}\overrightarrow{AB}$ ist. Um zu rechnen wählen wir Koordinaten für die Eckpunkte (A, B, C) des Dreiecks, also einen affinen Isomorphismus $f : \langle A, B, C\rangle_a \to \mathscr{A}_2(K)$ definiert durch

$$f(A) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},\ f(B) = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix},\ f(C) = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$$

Ein solcher Isomorphismus existiert per Definition, da (A, B, C) affin unabhängig ist. Es genügt den Satz für das Bild unter f zu zeigen, da die Aussage invariant unter affinen Isomorphismen ist. Dann ergeben sich die Fußpunkte $f(a) = f(B + \frac{1}{2}\overrightarrow{BC}) = (1, 1, 1)^{tr}$, $f(b) = (0, 1, 1)^{tr}$, $f(c) = (1, 0, 1)^{tr}$. Die Seitenhalbierenden sind dann

$$f(s_a) = \{f(A) + \alpha\overline{f}(\overrightarrow{Aa}) = \begin{pmatrix} \alpha \\ \alpha \\ 1 \end{pmatrix} \mid \alpha \in K\},\ f(s_b) = \{\begin{pmatrix} 2 - 2\beta \\ \beta \\ 1 \end{pmatrix} \mid \beta \in K\},$$

$$f(s_c) = \{\begin{pmatrix} \gamma \\ 2 - 2\gamma \\ 1 \end{pmatrix} \mid \gamma \in K\}$$

Um den Schnittpunkt $\{S\} = s_a \cap s_b \cap s_c$ zu berechnen, suchen wir $\alpha, \beta, \gamma \in K$ mit

$$\begin{pmatrix} \alpha \\ \alpha \\ 1 \end{pmatrix} = \begin{pmatrix} 2 - 2\beta \\ \beta \\ 1 \end{pmatrix} = \begin{pmatrix} \gamma \\ 2 - 2\gamma \\ 1 \end{pmatrix}.$$

und finden als Lösung $\alpha = \beta = \gamma = 2/3$ also $f(S) = \begin{pmatrix} 2/3 \\ 2/3 \\ 1 \end{pmatrix}$. Das Teilverhältnis ergibt sich als $\mathrm{TV}(A, a, S) = \mathrm{TV}(f(A), f(a), f(S)) = 2/3 = \mathrm{TV}(B, b, S) = \mathrm{TV}(C, c, S)$. □

Zum Beweis des nächsten Satzes benötigen wir **Streckungen**, also affine Abbildungen der Form

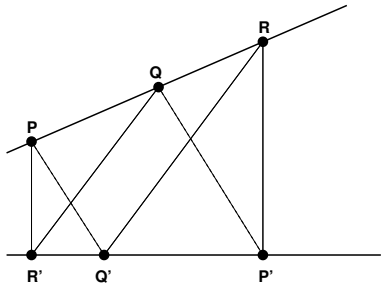
$$\mathscr{A} \to \mathscr{A} : P \mapsto P_0 + a\overrightarrow{P_0P},$$

wobei das feste $P_0 \in \mathscr{A}$ das Streckzentrum ist und das $a \in K^*$ der Streckfaktor.

Übung 9.1.6. Zeige: Je zwei Streckungen von $\mathscr{A}$ sind konjugiert in $\mathrm{Aff}(\mathscr{A})$ genau dann, wenn sie denselben Streckfaktor haben. Die Streckungen zusammen mit den Translationen bilden einen Normalteiler in $\mathrm{Aff}(\mathscr{A})$ isomorph zur Matrixgruppe

$$\left\{ \left(\begin{array}{c|c} aI_n & t \\ \hline 0 & 1 \end{array} \right) | a \in K^*, t \in K^{n\times 1} \right\} \cong K^{n\times 1} \rtimes K^*$$

Satz 9.1.25. *(*PAPPOS*) Seien* $\mathrm{Dim}(\mathscr{A}) = 2$ *und* D, D' *zwei Geraden in* $\mathscr{A}$ *mit sechs verschiedenen Punkten* $P, Q, R \in D$, $P', Q', R' \in D'$, *von denen keiner in* $D \cap D'$ *liegt. Gilt* $\langle P, Q' \rangle_a \parallel \langle Q, P' \rangle_a$ *und* $\langle Q, R' \rangle_a \parallel \langle R, Q' \rangle_a$, *so folgt* $\langle P, R' \rangle_a \parallel \langle R, P' \rangle_a$.



Beweis. Wir betrachten zunächst den Fall, dass D und D' sich schneiden. Dann sei $\{P_0\} = D \cap D'$ und f_1 die Streckung mit Zentrum P_0, die P in $Q = P_0 + a\overrightarrow{P_0P}$ überführt, und f_2 die Streckung mit Zentrum P_0, die Q nach R überführt. Es ist $Q' = \tau_V(P)$, also $f_1(Q') = f_1(\tau_V(P)) = \tau_{aV}(f_1(P)) = \tau_{aV}(Q) = P'$ und ebenso $f_2(R') = Q'$. Dann ist $(f_2 \circ f_1)(P) = R$ und $(f_1 \circ f_2)(R') = P'$. Da der lineare Anteil von $f_2 \circ f_1$ gleich dem von $f_1 \circ f_2$ gleich $b\,\mathrm{id}_\mathcal{V}$ ist (für ein $b \in K^*$), gilt $b\overrightarrow{PR'} = \overrightarrow{RP'}$ und somit $\langle P, R' \rangle_a \parallel \langle R, P' \rangle_a$.
Falls D und D' sich nicht schneiden, arbeitet man mit Translationen, denn dann sind D und D' parallel. □

Den nächsten Satz kann man als weitgehende Verallgemeinerung einer Version des Strahlensatzes auffassen.

Satz 9.1.26. *Sei* $\mathrm{Dim}(\mathscr{A}) = n$ *und* H_i *für* $i = 1, 2, 3$ *Hyperebenen in* $\mathscr{A}$, *also affine Teilräume der Dimension* $n - 1$. *Weiter seien* H_1, H_2, H_3 *parallel und* $H_1 \neq H_2$.

(1) Jede Gerade (= 1-dimensionaler affiner Teilraum von $\mathscr{A}$*), die nicht schwach parallel zu* H_1 *ist, hat genau einen Schnittpunkt mit* H_i *für* $i = 1, 2, 3$. *(Die Schnittpunkte sind offenbar kollinear.)*

(2) Das Teilverhältnis der drei Schnittpunkte aus (1) ist unabhängig von der Wahl der Geraden und legt H_3 *auf Grund der Nebenbedingungen* $\mathrm{Dim}\, H_3 = n - 1, H_3 \parallel H_1$ *eindeutig fest.*

Beweis.

(1) Sei $\mathcal{W} := \mathcal{T}(H_1) = \mathcal{T}(H_2) = \mathcal{T}(H_3)$. Wir betrachten $\mathscr{A}/\mathcal{W} := \{P + \mathcal{W} | P \in \mathscr{A}\}$ als eindimensionalen affinen Raum und beachten, dass $\nu : \mathscr{A} \to \mathscr{A}/\mathcal{W} : P \mapsto P + \mathcal{W}$ eine affine Abbildung ist. Für jede Gerade $G \subset \mathscr{A}$, wie in (1) spezifiziert, ist $\nu_{|G}$ ein affiner Isomorphismus. Klar: Die Schnittpunkte sind $\nu_{|G}^{-1}(H_i)$ und die Behauptung über die Teilverhältnisse folgt auch, da diese bei Anwendung von affinen Isomorphismen fest bleiben.

(2) Sei G eine Gerade wie in (1) angegeben. Dann gilt $\mathcal{T}(\mathscr{A}) = \mathcal{T}(H_1) \oplus \mathcal{T}(G)$. Entsprechend haben wir eine affine Abbildung, genauer eine **Parallelprojektion**, von $\mathscr{A}$ entlang $\mathcal{T}(H_1)$ auf G, nämlich

$$\pi : \mathscr{A} \to \mathscr{A} : P \mapsto P' \text{ mit } \{P'\} = (P + \mathcal{T}(H_1)) \cap G.$$

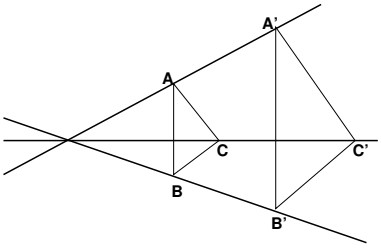
(Man muss nachrechnen, dass dies eine affine Abbildung ist. Die Projektionseigenschaft ist klar.) Jetzt kann der Beweis analog zum ersten Beweis fortgesetzt werden. □

Übung 9.1.7. Definiere Parallelprojektionen allgemein.

Übung 9.1.8. Zeige, dass in der affinen Ebene zwei Geraden sich entweder schneiden oder parallel sind. Genauer: Studiere die Bahnen unter der ebenen affinen Gruppe auf der Paarmenge der Geraden in der affinen Ebene.

Eigentlich ist der Satz von PAPPOS ein Satz, der zur projektiven Geometrie gehört. Ähnlich ist es mit dem Satz von DESARGUES, der sich im affinen Raum abspielt. Der Beweis, den hier gegeben wird, ist vielleicht vom synthetisch-geometrischen Standpunkt aus nicht schön, demonstriert aber die DESCARTESsche Idee, durch Einführung von Koordinaten geometrische Sätze durch algebraische Rechnungen zu beweisen. Für kompliziertere Situationen kann man sogar Computer heranziehen, um derartige Beweise "durchzurechnen".

Satz 9.1.27. *(DESARGUES[4]) Seien $(A, B, C), (A', B', C') \in \mathscr{A}^3_{generisch}$ zwei nicht entartete Dreiecke, die keine Eckpunkte gemeinsam haben und für die $\langle A, B\rangle_a \parallel \langle A', B'\rangle_a$, $\langle B, C\rangle_a \parallel \langle B', C'\rangle_a$, $\langle A, C\rangle_a \parallel \langle A', C'\rangle_a$. Dann schneiden sich die drei Geraden $\langle A, A'\rangle_a$, $\langle B, B'\rangle_a$, $\langle C, C'\rangle_a$ in einem gemeinsamen Punkt oder sind paarweise parallel.*



Beweis. Da die beiden Dreiecke nicht entartet sind, erzeugen sie einen drei- oder zweidimensionalen affinen Raum. Wir behandeln nur den ersten Fall, den zweiten ist als Übung überlassen:
Also sei (A, B, C, A') nicht komplanar. Dann können wir affine Koordinaten $\kappa : \langle A, B, C, A'\rangle_a \to K^{3\times 1}$ so wählen (wir arbeiten hier nicht mit $\mathscr{A}_n(K)$, weil es uns nicht weiter hilft!), dass

$$\kappa(A) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \kappa(B) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \kappa(C) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \kappa(A') = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

ist. Wegen der Parallelität der Seiten folgt durch kurze Rechnung

$$\kappa(B') = \begin{pmatrix} a \\ 0 \\ 1 \end{pmatrix}, \kappa(C') = \begin{pmatrix} 0 \\ a \\ 1 \end{pmatrix}$$

für ein $a \in K$. Da (A', B', C') nicht kollinear ist, folgt $a \neq 0$. Im Falle $a = 1$ sind die drei Geraden $\langle A, A'\rangle_a, \langle B, B'\rangle_a, \langle C, C'\rangle_a$ parallel. Anderenfalls schneiden sie sich in dem Punkt mit den Koordinaten $(0, 0, \frac{1}{1-a})^{tr}$. □

Ende Vorl. 23

[4]Genauer handelt es sich um eine affine Konsequenz der Umkehrung des (projektiven) Satzes von DESARGUES. Die Umkehrung ergibt sich aber durch Dualisieren aus dem ursprünglichen Satz von DESARGUES.

Kapitel 10

Multilineare Algebra

10.1 Tensorprodukte von Moduln

Sei R ein kommutativer Ring und K ein Körper.

Definition 10.1.1. Seien M, N, T Moduln über R.

(1) Eine Abbildung $\Phi : M \times N \to T$ heißt **bilinear**, falls

$$\Phi(aV + bV', W) = a\Phi(V, W) + b\Phi(V', W) \text{ für alle } a, b \in R, V, V' \in M, W \in N$$

und

$$\Phi(V, aW + bW') = a\Phi(V, W) + b\Phi(V, W') \text{ für alle } a, b \in R, V \in M, W, W' \in N.$$

(2) $(\otimes, \mathcal{T})$ heißt **Tensorprodukt** von M und N, falls

(a) $\otimes : M \times N \to \mathcal{T} : (V, W) \mapsto V \otimes W$ bilinear ist und

(b) für jeden R-Modul U und jede bilineare Abbildung $\Phi : M \times N \to U$ *genau* eine lineare Abbildung $\varphi : T \to U$ existiert, so dass das Diagramm

$$\begin{array}{rccc} \otimes : & M \times N & \to & \mathcal{T} \\ & & \Phi \searrow & \downarrow \varphi \\ & & & U \end{array}$$

kommutiert, d.h. $\Phi(V, W) = \varphi(V \otimes W)$.

Man sagt auch kurz: Ein Tensorprodukt ist eine **universelle** bilineare Abbildung .

Bemerkung 10.1.2. Sei $A \in R^{m \times n}, B \in R^{o \times p}$. Das KRONECKER**-Produkt** von A und B ist definiert als

$$A \otimes B := \begin{pmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \vdots & \vdots & & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{pmatrix} \in R^{mo \times np}.$$

Dann ist $\otimes : R^{m \times n} \times R^{o \times p} \to R^{mo \times np}$ sicher bilinear. Frage: Ist es universell, also ein Tensorprodukt?

Bemerkung 10.1.3. Falls ein Tensorprodukt von M und N existiert, ist es eindeutig bis auf Isomorphie bestimmt.

Beweis. Seien $(\otimes, T)$ und $(\otimes', U)$ Tensorprodukte der R-Moduln M und N. Dann haben wir eine lineare Abbildung $\varphi : \mathcal{T} \to \mathcal{U}$ mit $V \otimes' W = \varphi(V \otimes W)$ für alle $V \in M, W \in N$ und eine lineare Abbildung $\psi : \mathcal{U} \to \mathcal{T}$ mit $V \otimes W = \psi(V \otimes' W)$. Also hat die Komposition $\varphi \circ \psi : \mathcal{U} \to \mathcal{U}$ die Eigenschaft $V \otimes' W = (\varphi \circ \psi)(V \otimes' W)$ für alle $V \in M, W \in N$ ebenso wie die Identität $\mathrm{id}_{\mathcal{U}}$ von $\mathcal{U}$. Also ist $\varphi \circ \psi = \mathrm{id}_{\mathcal{U}}$. Analog $\psi \circ \varphi = \mathrm{id}_{\mathcal{T}}$. D.h. φ und ψ sind invers zueinander. □

Bemerkung 10.1.4. $M = \langle m_1, \ldots, m_r \rangle$, $N = \langle n_1, \ldots, n_s \rangle$, so ist jede bilineare Abbildung $\Phi : M \times N \to U$ eindeutig bestimmt durch alle $\Phi(m_i, n_j)$.

Bemerkung 10.1.5. Das KRONECKER-Produkt ist ein Tensorprodukt. Genauer $(\otimes, R^{mo \times np})$ ist ein Tensorprodukt von $R^{m \times n}$ und $R^{o \times p}$.

Beweis. Eine bilineare Abbildung Φ ist festgelegt durch die Bilder von (B, C), wobei B und C die Elemente von je einem freien Erzeugendensystem (=Basis) der beiden Ausgangsmoduln durchlaufen. Im vorliegenden Fall kann man die Standardbasen nehmen. Aber dann bilden die KRONECKER-Produkte $B \otimes C$ eine Basis von $R^{mo \times np}$, nämlich auch wieder die Standardbasis bei geeigneter Anordnung. Also ist durch $\varphi : B \otimes C \mapsto \Phi(B, C)$ eine lineare Abbildung festgelegt, die unser Diagramm zum Kommutieren bringt. Es ist klar, dass man keine andere Wahl für φ hat. □

Damit ist der Existenz- und Eindeutigkeitssatz für Tensorprodukte von freien Moduln vom endlichen Rang (insbesondere von endlich dimensionalen Vektorräumen) bewiesen.

Bemerkung 10.1.6. Sind $\mathcal{V}$ und $\mathcal{W}$ K-VR der Dimension n bzw. m, so hat das Tensorprodukt $\mathcal{V} \otimes \mathcal{W}$ die Dimension nm. Genauer, ist $B \in \mathcal{V}^m$ eine Basis von $\mathcal{V}$ und $C \in \mathcal{W}^n$ eine Basis von $\mathcal{W}$, so ist

$$B \otimes C := (B_1 \otimes C_1, B_1 \otimes C_2, \ldots, B_1 \otimes C_n, B_2 \otimes C_1, \ldots B_m \otimes C_n) \in (\mathcal{V} \otimes \mathcal{W})^{mn}$$

eine Basis von $\mathcal{V} \otimes \mathcal{W}$.

Satz 10.1.7. *Seien M und N R-Moduln. Dann existiert bis auf Isomorphie genau ein Tensorprodukt. Dieses wird mit $(\otimes, M \otimes N)$ bezeichnet.*

Beweis. Die Eindeutigkeit haben wir schon bewiesen. Zur Existenz: Sei F der freie R-Modul mit Basis $M \times N$, d.h.

$$F = \{ \sum_{(m,n) \in M \times N} a_{(m,n)}(m, n) \mid a_{(m,n)} \in R, a_{(m,n)} \neq 0 \text{ nur für endlich viele } (m, n) \}$$

Setze

$$B := \langle (rm_1 + m_2, sn_1 + n_2) - rs(m_1, n_1) - r(m_1, n_2) - s(m_2, n_1) - (m_2, n_2) \mid r, s \in R, m_1, m_2 \in M, n_1, n_2 \in N \rangle$$

Behauptung: Ist $T := F/B$ so ist $\otimes : M \times N \to T, (m, n) \mapsto (m, n) + B =: \pi(m, n)$ ein Tensorprodukt.
$\otimes$ ist bilinear nach Konstruktion, sprich nach Definition von B.
$\otimes$ hat auch die universelle Eigenschaft: Ist U ein R-Modul und $\psi : M \times N \to U$ bilinear, so gibt es genau einen R-Modulhomomorphismus $\varphi : F \to U$ mit $\varphi_{|M \times N} = \psi$, da F frei auf $M \times N$. Aus der Bilinearität von ψ folgt, dass $B \leq \mathrm{Kern}(\varphi)$. Somit induziert φ eine lineare Abbildung $T \to U$ mit der gewünschten Eigenschaft. □

Beispiel 10.1.8.

(1) $R = \mathbb{Z}$, $M = \mathbb{Z}/2\mathbb{Z} = \langle a \rangle$, $N = \mathbb{Z}/3\mathbb{Z} = \langle b \rangle$.
Hier ist $M \otimes N = \{0\}$. Denn jede bilineare Abbildung $\Phi : M \times N \to U$ erfüllt

$$\Phi(a,b) = 3\Phi(a,b) - 2\Phi(a,b) = \Phi(a,3b) - \Phi(2a,b) = 0 - 0 = 0.$$

(2) $R = \mathbb{Z}$, $M = \mathbb{Z}/2\mathbb{Z} = \langle a \rangle$, $N = \mathbb{Z}/4\mathbb{Z} = \langle c \rangle$.
Dann ist $M \otimes N = \mathbb{Z}/2\mathbb{Z}$. Denn jede bilineare Abbildung $\Phi : M \times N \to U$ ist eindeutig bestimmt durch $\Phi(a,c) =: u$. Es ist $2u = \Phi(2a,c) = 0$. Umgekehrt definiert $\Phi : M \times N \to \mathbb{Z}/2\mathbb{Z}$, $\Phi(a,c) := 1 + 2\mathbb{Z}$ eine bilineare Abbildung.

Wir haben noch eine unmittelbare Konsequenz der Definition: Die Existenz von Tensorprodukten von linearen Abbildungen.

Satz 10.1.9. *Seien $\mathcal{V}, \mathcal{V}'$ und $\mathcal{W}, \mathcal{W}'$ K-Vektorräume mit linearen Abbildungen $\alpha : \mathcal{V} \to \mathcal{V}'$ und $\beta : \mathcal{W} \to \mathcal{W}'$.*

(1) Es gibt genau eine lineare Abbildung

$$\alpha \otimes \beta : \mathcal{V} \otimes \mathcal{W} \to \mathcal{V}' \otimes \mathcal{W}'$$

mit $(\alpha \otimes \beta)(V \otimes W) = \alpha(V) \otimes \beta(W)$ für alle $V \in \mathcal{V}, W \in \mathcal{W}$.

(2) Sind B, B', C, C' Basen von $\mathcal{V}, \mathcal{V}', \mathcal{W}, \mathcal{W}'$, so gilt

$$^{B' \otimes C'}(\alpha \otimes \beta)^{B \otimes C} = {}^{B'}\alpha^{B} \otimes {}^{C'}\beta^{C},$$

wobei das zweite $\otimes$ das KRONECKER*-Produkt von Matrizen ist.*

(3)

$$\otimes : \mathrm{Hom}(\mathcal{V}, \mathcal{V}') \times \mathrm{Hom}(\mathcal{W}, \mathcal{W}') \to \mathrm{Hom}(\mathcal{V} \otimes \mathcal{W}, \mathcal{V}' \otimes \mathcal{W}') : (\varphi, \psi) \to \varphi \otimes \psi$$

ist ein Tensorprodukt, insbesondere für endlich dimensionale Vektorräume gilt

$$\mathrm{Hom}(\mathcal{V}, \mathcal{V}') \otimes \mathrm{Hom}(\mathcal{W}, \mathcal{W}') \cong \mathrm{Hom}(\mathcal{V} \otimes \mathcal{W}, \mathcal{V}' \otimes \mathcal{W}').$$

Ende Vorl. 24

Beweis.

(1) Offenbar ist

$$\mathcal{V} \times \mathcal{W} \to \mathcal{V}' \otimes \mathcal{W}' : (V, W) \mapsto \alpha(V) \otimes \beta(W)$$

bilinear. Mit der Definition des Tensorproduktes folgt die Behauptung.

(2) Sei ${}^{B'}\alpha^{B} =: M, {}^{C'}\beta^{C} =: N$ Dann ist

$$(\alpha \otimes \beta)(B_i \otimes C_j) = \alpha(B_i) \otimes \beta(C_j) = (\sum_k M_{ki} B'_k) \otimes (\sum_l N_{lj} C'_l) =$$

$$= \sum_k \sum_l M_{ki} N_{lj} B'_k \otimes C'_l.$$

(3) Die Existenz der Abbildung folgt aus (1) und die Isomorphie aus (2). □

Offenbar kann man $K \otimes \mathcal{W}$ mit $\mathcal{W}$ identifizieren, indem man $k \otimes W$ mit kW für $k \in K, W \in \mathcal{W}$ identifiziert. Dann bekommen wir eine wichtige Folgerung, wobei mit $\mathcal{V}^*$ wie üblich der Dualraum $\mathrm{Hom}(\mathcal{V}, K)$ bezeichnet wird:

Folgerung 10.1.10. $\mathcal{V}^* \otimes \mathcal{W} \cong \mathrm{Hom}(\mathcal{V}, \mathcal{W})$. *Genauer,*

$$\otimes : \mathcal{V}^* \times \mathcal{W} \to \mathrm{Hom}(\mathcal{V}, \mathcal{W}) : (\varphi, W) \mapsto (V \mapsto \varphi(V)W)$$

ist ein Tensorprodukt.

Beweis. $\mathcal{V}^* \otimes \mathcal{W} \cong \mathrm{Hom}(\mathcal{V}, K) \otimes \mathrm{Hom}(K, \mathcal{W}) \cong \mathrm{Hom}(\mathcal{V} \otimes K, K \otimes \mathcal{W}) \cong \mathrm{Hom}(\mathcal{V}, \mathcal{W})$. □

Bemerkung 10.1.11. (Konstantenerweiterung) Sei K Teilkörper eines Körpers F und $\mathcal{V}$ ein K-Vektorraum. Dann ist $\mathcal{V}_F := F \otimes \mathcal{V}$ ein F-Vektorraum, die **Konstantenerweiterung**, mit

$$a(b \otimes V) := (ab) \otimes V \text{ für alle } a, b \in F, V \in \mathcal{V}.$$

(1) Ist $B \in \mathcal{V}^n$ eine K-Basis von $\mathcal{V}$, so ist $1 \otimes B := (1 \otimes B_1, \ldots, 1 \otimes B_n) \in (\mathcal{V}_F)^n$ eine F-Basis von $\mathcal{V}_F$.

(2) Ist $\mathcal{W}$ ein weiterer K-Vektorraum und $\varphi : \mathcal{V} \to \mathcal{W}$ linear (über K), so ist $\mathrm{id}_F \otimes \varphi : \mathcal{V}_F \to \mathcal{W}_F$ linear über F. Ist $C \in \mathcal{W}^m$ eine K-Basis von $\mathcal{W}$, so ist

$${}^{1\otimes C}(\mathrm{id}_F \otimes \varphi)^{1\otimes B} = {}^C\varphi^B.$$

(3) Die beiden F-Vektorräume $\mathrm{Hom}(\mathcal{V}, \mathcal{W})_F$ und $\mathrm{Hom}(\mathcal{V}_F, \mathcal{W}_F)$ werden identifiziert, indem man $a \otimes \varphi$ für $a \in F, \varphi \in \mathrm{Hom}(\mathcal{V}, \mathcal{W})$ mit $\widetilde{a} \otimes \varphi$ identifiziert, wobei $\widetilde{a}$ die Skalierung $\widetilde{a} : F \to F, b \mapsto ab$ bezeichnet. (Hinweis: $F \times \mathrm{Hom}(\mathcal{V}, \mathcal{W}) \to \mathrm{Hom}(\mathcal{V}_F, \mathcal{W}_F) : (a, \varphi) \mapsto \widetilde{a} \otimes \varphi$ ist ein Tensorprodukt.)

Beweis. Übung. □

10.2 Die Tensoralgebra.

Wir haben bislang nur von Tensorprodukten von zwei Vektorräumen gesprochen. Es ist klar, dass man die Konstruktion auch für n Vektorräume durchführen kann. Wir begnügen uns mit dem Fall $n = 3$.

Definition 10.2.1. Seien $\mathcal{V}_i$ für $i = 1, \ldots n$ und $\mathcal{W}, \mathcal{T}$ K-Vektorräume. Eine Abbildung

$$\Psi : \times_{i=1}^n \mathcal{V}_i \to \mathcal{W}$$

heißt multilinear, falls sie in jeder Komponente linear ist.
$(\otimes, \mathcal{T})$ heißt **Tensorprodukt** der $\mathcal{V}_i$'s, falls $\mathcal{T}$ ein K-Vektorraum ist,

$$\otimes : \times_{i=1}^n \mathcal{V}_i \to \mathcal{T} : (V_1, \ldots, V_n) \mapsto V_1 \otimes \ldots \otimes V_n$$

multilinear ist und für jede multilineare Abbildung

$$\Psi : \times_{i=1}^n \mathcal{V}_i \to \mathcal{W}$$

in einen beliebigen K-Vektorraum $\mathcal{W}$ genau eine lineare Abbildung $\varphi : \mathcal{T} \to \mathcal{W}$ existiert mit $\Psi(V_1, \ldots, V_n) = \varphi(V_1 \otimes \ldots \otimes V_n)$ für alle $V_i \in \mathcal{V}_i, i = 1, \ldots, n$.

Satz 10.2.2. *Bis auf Isomorphie gibt es genau ein Tensorprodukt $(\otimes, \mathcal{T})$ von $(\mathcal{V}_1, \ldots, \mathcal{V}_n)$. Dieses wird mit $\otimes_{i=1}^n \mathcal{V}_i$ bezeichnet.*

Beweis. Der Beweis für den Fall $n = 2$ in Bemerkung 10.1.5 (bzw. Satz 10.1.7) überträgt sich, wenn man die offensichtliche Identität von KRONECKER-Produkten von Matrizen beachtet: Für beliebige Matrizen A, B, C über R gilt:

$$A \otimes (B \otimes C) = (A \otimes B) \otimes C.$$ □

Da offensichtlich sowohl $\mathcal{V}_1 \otimes (\mathcal{V}_2 \otimes \mathcal{V}_3)$ als auch $(\mathcal{V}_1 \otimes \mathcal{V}_2) \otimes \mathcal{V}_3$ ein Tensorprodukt von $(\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3)$ ist, haben wir ein Korollar, das uns ermöglicht in Zukunft alle Klammerungen bei Tensorprodukten zu ignorieren.

Folgerung 10.2.3. $(\mathcal{V}_1 \otimes \mathcal{V}_2) \otimes \mathcal{V}_3 \cong \mathcal{V}_1 \otimes (\mathcal{V}_2 \otimes \mathcal{V}_3) \cong \mathcal{V}_1 \otimes \mathcal{V}_2 \otimes \mathcal{V}_3$ *mit* $(V_1 \otimes V_2) \otimes V_3 \leftrightarrow V_1 \otimes (V_2 \otimes V_3) \leftrightarrow V_1 \otimes V_2 \otimes V_3$ *für alle* $V_i \in \mathcal{V}_i, i = 1, 2, 3$.

Definition 10.2.4. Sei $\mathcal{V}$ ein K-Vektorraum. Setze $\mathcal{V}^{\otimes 0} := T^0\mathcal{V} := K1$, ein eindimensionaler K-Vektorraum mit Basis 1 und $\mathcal{V}^{\otimes n} := T^n\mathcal{V} := \otimes^n\mathcal{V}$ für $n > 0$. Die **Tensoralgebra** $T\mathcal{V}$ von $\mathcal{V}$ ist definiert als

$$T\mathcal{V} := \bigoplus_{i \in \mathbb{Z}_{\geq 0}} T^i\mathcal{V}$$

mit der bilinearen Multiplikation

$$(V_1 \otimes \ldots \otimes V_m)(W_1 \otimes \ldots \otimes W_n) := V_1 \otimes \ldots \otimes V_m \otimes W_1 \otimes \ldots \otimes W_n$$

für alle $V \in \mathcal{V}^m, W \in \mathcal{V}^n$ und $1X = X = X1$ für alle $X \in T\mathcal{V}$.

Für die Tensoralgebra können wir nicht nur zeigen, dass sie eine Algebra ist, sondern dass sie sogar eine universelle Eigenschaft hat, ähnlich wie das Tensorprodukt selbst.

Satz 10.2.5.

(1) $T\mathcal{V}$ ist eine assoziative K-Algebra.

(2) Ist $\mathscr{A}$ eine assoziative K-Algebra mit Einselement und $\varphi : \mathcal{V} \to \mathscr{A}$ eine lineare Abbildung, so gibt es genau einen Algebrenhomomorphismus $\overline{\varphi} : T\mathcal{V} \to \mathscr{A}$, der φ fortsetzt.

Beweis.

(1) Zeige, dass die Multiplikation wohldefiniert ist. Klar:

$$T^n\mathcal{V} \times T^m\mathcal{V} \to T^{n+m}\mathcal{V} : (X, Y) \to X \otimes Y$$

ist bilinear, sogar ein Tensorprodukt. Also ist die Multiplikation auf $T^n\mathcal{V} \times T^m\mathcal{V}$ wohldefiniert und bilinear. Diese wird bilinear auf ganz $T\mathcal{V} \times T\mathcal{V} \to T\mathcal{V}$ fortgesetzt, was wegen der Struktur von $T\mathcal{V}$ als direkte Summe wohldefiniert ist. Offenbar ist dieses Produkt auch assoziativ, so dass wir eine K-Algebra mit Eins haben.

(2) Klar: Soll $\overline{\varphi}$ ein Homomorphismus für Algebren mit Eins sein, so muss $\overline{\varphi}(1) = 1 \in \mathscr{A}$ und $\overline{\varphi}(V_1 \otimes \ldots \otimes V_m) = \varphi(V_1) \cdots \varphi(V_m)$ für alle $V \in \mathcal{V}^n$ sein. Offenbar ist aber

$$\mathcal{V}^n \to \mathscr{A} : V \mapsto \varphi(V_1) \cdots \varphi(V_m)$$

multilinear, so dass $\overline{\varphi}_{|T^n\mathcal{V}}$ wohldefiniert ist. Da $T\mathcal{V}$ direkte Summe der $T^m\mathcal{V}$ ist, folgt die Behauptung.

□

Ende Vorl. 25

10.3 Alternierende Tensoren und die Grassmann-Algebra.

Lernziel: Alternierende Multilinearformen, alternierende Multivektoren, geometrische Interpretation, symmetrische Tensoren, Zusammenhang mit Polynomen.

Wir hatten früher Determinanten als alternierende Multilinearformen kennengelernt. Wir wollen jetzt einen breiteren Kontext für Determinanten herstellen. Unsere Vektorräume bleiben in der Regel endlich erzeugt.

Definition 10.3.1. Seien $\mathcal{V}, \mathcal{W}$ K-Vektorräume.

(1) Eine multilineare Abbildung $\Psi : \mathcal{V}^k \to \mathcal{W}$ heißt **alternierend**, falls $\Psi(V) = 0$ für alle $V \in \mathcal{V}^k$, für die verschiedene $i, j \in \underline{k}$ existieren mit $V_i = V_j$.

(2) $(\wedge, \mathcal{T})$ heißt ein **äusseres k-faches Produkt** von $\mathcal{V}$ oder ein **alternierendes k-faches Produkt**, falls $\mathcal{T}$ ein K-Vektorraum ist, $\wedge : \mathcal{V}^k \to \mathcal{T} : V \mapsto V_1 \wedge \ldots \wedge V_k$ eine alternierende multilineare Abbildung ist und für jeden K-Vektorraum $\mathcal{W}$ und jede alternierende multilineare Abbildung $\Psi : \mathcal{V}^k \to \mathcal{W}$ genau eine lineare Abbildung $\psi : \mathcal{T} \to \mathcal{W}$ existiert mit
$$\Psi(V) = \psi(V_1 \wedge \ldots \wedge V_k) \text{ für alle } V \in \mathcal{V}^k.$$

In der Schule gab es schon das erste Beispiele einer alternierenden bilinearen Abbildungen: Das $\times$-Produkt. Die Determinante ist ebenfalls eine alternierende Multilinearform.

Übung 10.3.1. Zeige: Falls ein alternierendes k-faches Produkt von $\mathcal{V}$ existiert, so ist es bis auf *eindeutige* Isomorphie eindeutig.

Satz 10.3.2. *Sei $\mathcal{V}$ ein endlich erzeugter K-Vektorraum. Es existiert bis auf eindeutige Isomorphie genau ein äusseres k-faches Produkt $(\wedge, \mathcal{T})$. Bezeichnung: $\bigwedge^k \mathcal{V} := \mathcal{T}$.*

Beweis. Falls $\mathcal{T}$ existiert, sollte es nach Definition der Tensorpotenz ein epimorphes Bild von $T^k\mathcal{V}$ sein. Ist $\varepsilon : T^k\mathcal{V} \to \mathcal{T}$ dieser Epimorphismus, so gilt sicherlich $V_1 \otimes \ldots \otimes V_k \in \operatorname{Kern} \varepsilon$, sobald $V_i = V_j$ ist für ein Paar (i, j) mit $i \neq j$. Dies führt uns zu folgendem Ansatz:
$\mathcal{T} := (T^k\mathcal{V})/\mathcal{U}$, wo $\mathcal{U}$ von allen $V_1 \otimes \ldots \otimes V_k$ mit $V \in \mathcal{V}^k, V_i = V_j$ für ein Paar $(i, j), i \neq j$ erzeugt wird. Es folgt sofort:

$$\wedge : \mathcal{V}^k \to \mathcal{T} : V \mapsto V_1 \wedge \ldots \wedge V_k := V_1 \otimes \ldots \otimes V_k + \mathcal{U}$$

ist multilinear und alternierend. Wir überprüfen die Universalität: Sei $\Psi : \mathcal{V}^k \to \mathcal{W}$ multilinear und alternierend. Nach Definition des Tensorproduktes haben wir eine eindeutige lineare Abbildung $\alpha : T^k\mathcal{V} \to \mathcal{W}$ mit $\Psi(V) = \alpha(V_1 \otimes \ldots \otimes V_k)$ für alle $V \in \mathcal{V}^k$. Da Ψ alternierend ist, folgt $\mathcal{U} \leq \operatorname{Kern} \alpha$. Also faktorisiert α über den natürlichen Epimorphismus $\varepsilon : T^k\mathcal{V} \to \mathcal{T}$, d.h. es gibt eine eindeutige lineare Abbildung $\psi : \mathcal{T} \to \mathcal{W}$ mit $\psi \circ \varepsilon = \alpha$, d.h. mit $\Psi(V) = \psi(V_1 \wedge \ldots \wedge V_k)$. □

Leider sagt uns der Satz nichts über die Dimension von $\bigwedge^k \mathcal{V}$. Wir schauen uns einige Beispiele an:

Beispiel 10.3.3.

(1) Sei $k > \operatorname{Dim} \mathcal{V}$, dann ist $\bigwedge^k \mathcal{V} = \{0\}$. Denn sei $\Psi : \mathcal{V}^k \to \mathcal{W}$ alternierend, so ist jedes $V \in \mathcal{V}^k$ linear abhängig und somit $\Psi(V) = 0$ (Schluss wie bei der Determinante.)

(2) $\bigwedge^1 \mathcal{V} \cong \mathcal{V}$.

(3) Sei $n = \operatorname{Dim}\mathcal{V}$. Dann gilt $\operatorname{Dim}\bigwedge^n \mathcal{V} = 1$. Genauer, ist $B \in \mathcal{V}^n$ eine Basis von $\mathcal{V}$, so ist $(B_1 \wedge \ldots \wedge B_n)$ eine Basis von $\bigwedge^n \mathcal{V}$. (Beweis: Übung.) Andere Sichtweise: Im letzten Semester hatten wir bereits gezeigt, dass der Vektorraum der alternierenden Linearformen eindimensional ist, und zwar von der Determinante erzeugt wird.

Satz 10.3.4. *Sei* $\operatorname{Dim}\mathcal{V} = n$. *Dann gilt* $\operatorname{Dim}\bigwedge^k \mathcal{V} = \binom{n}{k}$. *Genauer: Sei* $B \in \mathcal{V}^n$ *eine Basis von* $\mathcal{V}$. *Für jede* k*-elementige Teilmenge* I *von* $\underline{n}$ *mit Elementen* $i_1 < i_2 < \ldots < i_k$ *sei* $B_I := B_{i_1} \wedge \ldots \wedge B_{i_k}$. *Dann ist* $(B_I | I \subseteq \underline{n}, |I| = k)$ *eine Basis von* $\bigwedge^k \mathcal{V}$.

Beweis.

(1) Die B_I erzeugen $\bigwedge^k \mathcal{V}$: Zur Vorbereitung beachte

$$V_1 \wedge \ldots\ldots \wedge V_k = \operatorname{sign}(\sigma) W_1 \wedge \ldots \wedge W_k$$

für alle $V \in \mathcal{V}^k, \sigma \in S_k, W = V \circ \sigma$. Dies folgt sofort für Transpositionen σ und da diese S_k erzeugen auch für alle $\sigma \in S_k$. Da nun klar ist, dass $\bigwedge^k \mathcal{V}$ von Elementen der Form $B_{a(1)} \wedge \ldots \wedge B_{a(k)}$ erzeugt wird, wo $a : \underline{k} \to \underline{n}$ eine beliebige Abbildung ist, folgt die Erzeugung aus der Multilinearität, der Umsortierungseigenschaft oben und der offensichtlichen Tatsache, dass $B_{a(1)} \wedge \ldots \wedge B_{a(k)} = 0$ ist, falls a nicht injektiv ist.

(2) Lineare Unabhängigkeit: Sei $\sum_I a_I B_I = 0$ für gewisse $a_I \in K, I \subseteq \underline{n}, |I| = k$. Sei nun $J \subseteq \underline{n}, |J| = k$. Um $a_J = 0$ zu zeigen, betrachte $\pi_J : \mathcal{V} \to \langle B_i | i \in J \rangle : \sum a_i B_i \mapsto \sum_{i \in J} a_i B_i$ und sei $\det$ die Determinante von $\langle B_i | i \in J \rangle$ bezüglich $(B_i)_{i \in J}$. Dann ist

$$\Psi_J : \mathcal{V}^k \to K : V \mapsto \det(\pi_J(V_1), \ldots, \pi_J(V_k))$$

multilinear und alternierend. Nach Definition des äusseren Produktes haben wir eine eindeutige lineare Abbildung $\psi_J : \mathcal{V}^k \to K$ mit $\Psi_J(V) = \psi_J(V_1 \wedge \ldots \wedge V_k)$ für alle $V \in \mathcal{V}^k$. Klar: $\psi_J(B_I) = \delta_{IJ} := \begin{cases} 1 & I = J, \\ 0 & I \neq J \end{cases}$. Wenden wir also ψ_J auf $\sum_I a_I B_I = 0$ an, so folgt $a_J = 0$. □

Folgerung 10.3.5. *Ein* k*-Tupel* $V \in \mathcal{V}^k$ *ist genau dann linear abhängig, wenn* $V_1 \wedge \ldots \wedge V_k = 0$ *gilt.*

Bislang war $\wedge$ nur eine alternierende multilineare Abbildung. Wir wollen uns überlegen, dass man $\wedge$ auch als Verknüpfung auffassen kann und dabei alle $\bigwedge^k \mathcal{V}$ zu einem grösseren Ganzen zusammenfassen.

Definition 10.3.6. Sei $\operatorname{Dim}\mathcal{V} = n$ und $\bigwedge^0 \mathcal{V} := K1$ ein eindimensionaler K-Vektorraum mit Basis (1). Man setzt

$$\bigwedge \mathcal{V} := \bigoplus_{k=0}^{n} \bigwedge^k \mathcal{V}$$

und definiert eine Multiplikation auf $\bigwedge \mathcal{V}$, indem man die durch

$$\wedge_{k,l} : \bigwedge^k \mathcal{V} \times \bigwedge^l \mathcal{V} \to \bigwedge^{k+l} \mathcal{V} :$$

$$(V_1 \wedge \ldots \wedge V_k, W_1 \wedge \ldots \wedge W_l) \mapsto V_1 \wedge \ldots \wedge V_k \wedge W_1 \wedge \ldots \wedge W_l$$

definierten bilinearen Abbildungen zu einer bilinearen Abbildung

$$\wedge : \bigwedge \mathcal{V} \times \bigwedge \mathcal{V} \to \bigwedge \mathcal{V}$$

zusammensetzt. $(\bigwedge \mathcal{V}, \wedge)$ heißt die **äussere Algebra** oder GRASSMANN**-Algebra** von $\mathcal{V}$.

Satz 10.3.7. *$\bigwedge \mathcal{V}$ ist eine assoziative Algebra der Dimension $2^{\mathrm{Dim}\,\mathcal{V}}$.*

Ende Vorl. 26

Beweis. Wir müssen zuerst überlegen, dass das Produkt wohldefiniert ist. Dies ist mit der gegebenen Definition etwas umständlich. Hier sind zwei andere Möglichkeiten:

(1) Man wählt eine Basis $B \in \mathcal{V}^n$ und definiert für die Basis $(B_I | I \subseteq \underline{n})$ das Produkt als $(B_I, B_J) \mapsto B_I \wedge B_J := B_{i_1} \wedge \ldots \wedge B_{j_{|J|}}$, was also 0 oder bis aufs Vorzeichen $B_{I \cup J}$ im Falle $I \cap J = \emptyset$ ist. Dann setzt man bilinear fort und rechnet man die Eigenschaft aus der Definition nach.

(2) Man geht von der Tensoralgebra $T\mathcal{V}$ aus und betrachtet den Teilraum $A\mathcal{V} := \bigoplus_{k \geq 0} A^k\mathcal{V}$ mit $A^k\mathcal{V} \leq T^k\mathcal{V}$ erzeugt von allen $V_1 \otimes \ldots \otimes V_k$ für $V \in \mathcal{V}^k$ mit zwei verschiedenen Indizes i, j mit $i \neq j, V_i = V_j$. Wir erinnern uns $\bigwedge^k \mathcal{V} = T^k\mathcal{V}/A^k\mathcal{V}$. Also haben wir einen Vektorraumepimorphismus $\alpha : T\mathcal{V} \to \bigwedge \mathcal{V}$, der bei den Basisvektoren $\otimes$ durch $\wedge$ ersetzt, mit $\mathrm{Kern}\,\alpha = A\mathcal{V}$. Nun hat $A\mathcal{V}$ folgende bemerkenswerte Eigenschaft:

$$X \in A\mathcal{V}, Y \in T\mathcal{V} \text{ impliziert } XY, YX \in A\mathcal{V}$$

(was den Relationen $0y = y0 = 0$ in einem Ring R entspricht). Deswegen gilt für $X, X', Y, Y' \in T\mathcal{V}$:

$$X - X', Y - Y' \in A\mathcal{V} \text{ impliziert } XY - X'Y' \in A\mathcal{V},$$

d.h. wir können für die Restklassen nach $A\mathcal{V}$ widerspruchsfrei eine Multiplikation über die Multiplikation der Vertreter definieren, oder anders ausgedrückt:

$$\alpha(X) \wedge \alpha(Y) := \alpha(X \otimes Y)$$

ist eine wohldefinierte assoziative Multiplikation für $\bigwedge \mathcal{V}$.
Bei der ersten Möglichkeit muss man noch die Assoziativität auf den Basisvektoren überprüfen, bei der zweiten ist dies überflüssig.
Die Dimension ergibt sich aus 10.3.4. Man beachte $B_\emptyset$ ist das Einselement der Algebra. □

10.4 Symmetrische Tensoren.

Lemma 10.4.1. *Die Operation von S_k auf $\mathcal{V}^k$ durch*

$$S_k \times \mathcal{V}^k \to \mathcal{V}^k : (\sigma, V) \mapsto V \circ \sigma^{-1}$$

induziert eine lineare Operation von S_k auf $T^k\mathcal{V}$ gegeben durch

$$\sigma(\sum_V a_V V_1 \otimes \ldots \otimes V_k) := \sum_V a_V V_{\sigma^{-1}(1)} \otimes \ldots \otimes V_{\sigma^{-1}(k)},$$

wobei $\sigma \in S_k$ und die Summe über endlich viele $V \in \mathcal{V}^k$ genommen ist. Diese Operation der S_k ist mit der Operation der $\mathrm{GL}(\mathcal{V})$ vertauschbar. Entsprechend bekommt man eine Operation von S_k auf $T^k\mathcal{V}^*$ mit entsprechenden Eigenschaften.

Beweis. Übung. □

Definition 10.4.2.

(1) Ein Tensor $T \in T^k\mathcal{V}$ oder in $T^k\mathcal{V}^*$ heißt **schiefsymmetrisch**, falls $\sigma(T) = \mathrm{sign}(\sigma)T$ für alle $\sigma \in S_k$.

(2) Sei

$$\mathrm{alt} = \mathrm{alt}_k : T^k\mathcal{V} \to T^k\mathcal{V} : V_1 \otimes \ldots \otimes V_k \mapsto \sum_{\sigma \in S_k} \mathrm{sign}(\sigma) V_{\sigma(1)} \otimes \ldots \otimes V_{\sigma(k)},$$

so dass $\hat{\bigwedge}^k \mathcal{V} = \mathrm{Bild}\,\mathrm{alt}_k$. Ein Tensor $T \in T^k\mathcal{V}$ heißt **alternierend**, falls $T \in \hat{\bigwedge}^k \mathcal{V} = \mathrm{Bild}\,\mathrm{alt}_k$.

Übung 10.4.1. Zeige $T \in T^k\mathcal{V}$ ist genau dann schiefsymmetrisch, wenn $\sigma(T) = -T$ für alle Transpositionen $\sigma \in S_k$. (Hinweis: sign ist ein Gruppenhomomorphismus und die Transpositionen erzeugen S_k als Gruppe.)

Klar: Alternierende Tensoren sind schiefsymmetrisch. Wie steht es mit der Umkehrung?

Bemerkung 10.4.3. Sei $k! \neq 0$ in K. Dann ist $\frac{1}{k!}\,\mathrm{alt}_k \in \mathrm{End}(T^k\mathcal{V})$ eine Projektion. Insbesondere ist ein k-Tensor genau dann schiefsymmetrisch, wenn er alternierend ist.

Beweis. Sei $\overline{\sigma} \in \mathrm{GL}(T^k\mathcal{V})$ der von $\sigma \in S_n$ induzierte Automorphismus. Wegen $\overline{\sigma \circ \tau} = \overline{\sigma} \circ \overline{\tau}$ und der Homomorphieeigenschaft von sign folgt aus $\mathrm{alt} = \sum_{\sigma \in S_k} \mathrm{sign}(\sigma)\overline{\sigma}$, dass $\mathrm{alt}^2 = k!\,\mathrm{alt}$ gilt, woraus die Behauptung folgt. □

Das wahrscheinlich wichtigste Objekt, welches wir in diesem Abschnitt kennengelernt haben, ist die GRASSMANN-Algebra $\bigwedge \mathcal{V}^*$, die als wichtigste Elemente die alternierenden k-Formen auf $\mathcal{V}$ beherbergt.

Wir kommen zu dem Fall symmetrischer Tensoren, bei dem wir uns kurz fassen können, weil wir einerseits schon etwas Erfahrung von den schiefsymmetrischen Tensoren her haben und andererseits bereits über den Polynomring gesprochen haben, zu dem die symmetrische Algebra isomorph ist.

Bemerkung 10.4.4. Sei $\mathcal{V}$ ein K-Vektorraum.

(1) Eine Abbildung $\Psi : \mathcal{V}^k \to \mathcal{W}$ in einen K-Vektorraum $\mathcal{W}$ heißt **symmetrisch**, falls $\Psi(V) = \Psi(V \circ \sigma)$ gilt für alle $\sigma \in S_k, V \in \mathcal{V}^k$. Ein symmetrisches Produkt $(\cdot, S^k\mathcal{V})$ ist eine universelle multilineare symmetrische Abbildung $\cdot : \mathcal{V}^k \to S^k\mathcal{V}$. Wegen der Universalität ist $S^k\mathcal{V}$ bis auf Isomorphie eindeutig bestimmt. Es kann konstruiert werden als

$$S^k\mathcal{V} := \mathcal{V}^{\otimes k} / \underbrace{\langle V_1 \otimes \ldots \otimes V_k - V_{\sigma(1)} \otimes \ldots \otimes V_{\sigma(k)} | V_i \in \mathcal{V}, \sigma \in S_k \rangle}_{D^k\mathcal{V}}.$$

(2) Die **symmetrische Algebra** $S\mathcal{V}$ von $\mathcal{V}$ ist definiert als

$$S\mathcal{V} := \bigoplus_{i \geq 0} S^i\mathcal{V},$$

wobei $S^0\mathcal{V}$ als einsimensionaler K-Vektorraum mit Basis (1) definiert wird. Das Produkt auf $S\mathcal{V}$ ist so definiert, dass der offensichtliche Epimorphismus $T\mathcal{V} \to S\mathcal{V}$ das Produkt der Tensoralgebra $T\mathcal{V}$ auf $S\mathcal{V}$ überträgt. Mit diesem Produkt ist $S\mathcal{V}$ eine assoziative und kommutative K-Algebra, und zwar universell assoziativ-kommutativ, d.h. jede lineare Abbildung $\varphi : \mathcal{V} \to \mathscr{A}$ in eine kommutative, assoziative K-Algebra $\mathscr{A}$ (mit Eins) lässt sich eindeutig zu einem K-Algebrenhomomorphismus $\overline{\varphi} : S\mathcal{V} \to \mathscr{A}$ fortsetzen. (Beachte $S^1\mathcal{V} = \mathcal{V}$.)

Wie beim alternierenden Produkt die GRASSMANN-Algebra das wichtigste Objekt war, so ist es hier die symmetrische Algebra $S\mathcal{V}^*$ auf $\mathcal{V}^*$, weil sie mit der Polynomalgebra identifiziert werden kann und ihre Elemente in natürlicher Weise die Polynomfunktionen induziert. Es folgt eine tabellarische Gegenüberstellung der analogen Begriffe für den alternierenden und den symmetrischen Fall. Man benutze es als Wörterbuch, um die Theorie für den symmetrischen Fall herzuleiten.

Alternierender Fall	Symmetrischer Fall
$A^k\mathcal{V}$ mit Erzeugern $V_1 \otimes \ldots \otimes V_k$ zwei V_i gleich	$D^k\mathcal{V}$ mit Erzeugern $V_1 \otimes \ldots \otimes V_k - V_{\sigma(1)} \otimes \ldots \otimes V_{\sigma(k)}$ mit $\sigma \in S_k$
$\bigwedge^k \mathcal{V}$	$S^k\mathcal{V}$
alternierendes Produkt $\wedge : \mathcal{V}^k \to \bigwedge^k \mathcal{V}$	symmetrisches Produkt $\cdot : \mathcal{V}^k \to S^k\mathcal{V}$
$\operatorname{Dim} \bigwedge^k \mathcal{V} = \binom{n}{k}$	$\operatorname{Dim} S^k\mathcal{V} = \binom{n+k-1}{k}$
GRASSMANN-Algebra $\bigwedge \mathcal{V}$	symmetrische Algebra $S\mathcal{V}$
Gruppenhomomorphismus $S_k \to K^* : \sigma \mapsto \operatorname{sign}(\sigma)$	Gruppenhomomorphismus $S_k \to K^* : \sigma \mapsto 1$
schiefsymmetrische Tensoren in $T^k\mathcal{V}$	symmetrische Tensoren in $T^k\mathcal{V}$
$\operatorname{alt}_k := \sum_{\sigma\in S_k} \operatorname{sign}(\sigma)\overline{\sigma} \in \operatorname{End}(T^k\mathcal{V})$	$\operatorname{sym}_k := \sum_{\sigma\in S_k} \overline{\sigma} \in \operatorname{End}(T^k\mathcal{V})$
GRASSMANN-Algebra $\bigwedge \mathcal{V}^*$ und alternierende Formen auf $\mathcal{V}$	symmetrische Algebra $S\mathcal{V}^*$ und Polynomfunktionen auf $\mathcal{V}$

Übung 10.4.2. Definiere **symmetrische Tensoren** in $\mathcal{T}^k\mathcal{V}$ analog zu schiefsymmetrischen Tensoren.

Übung 10.4.3. Sei $B \in \mathcal{V}^n$ eine Basis von $\mathcal{V}$ und $X := B^* \in (\mathcal{V}^*)^n$ die Dualbasis zu B. Die Elemente von $S\mathcal{V}$ fassen wir auf als Polynome in den B_i und die Elemente von $S\mathcal{V}^*$ fassen wir auf als Polynome in den X_i. Für $g \in \mathrm{GL}(\mathcal{V})$ sei $G := {}^Bg^B$. Zeige:
1.) $gp(B_1, \ldots, B_n) = p(g(B_1), \ldots, g(B_n))$ mit $g(B_i) = BG_{-,i} = \sum_j G_{ji}B_j$.
2.) $g^{-1}p(X_1, \ldots, X_n) = p(g^{-1}X_1, \ldots, g^{-1}X_n)$ mit $g^{-1}X_i = g^{tr}(X_i) = X(G^{tr})_{-,i} = \sum_j G_{ij}X_j$.

Literaturverzeichnis

Index