Skript zur Vorlesung

Lineare Algebra 2

Sommersemester 2023

Prof. Dr. Annette Werner

Inhaltsverzeichnis

1 Die Jordan'sche Normalform **1**

2 Gruppentheorie **15**

3 Dualräume und Quotientenräume **25**

4 Ringe **29**

5 Bilinearformen **43**

6 Symmetrische Bilinearformen über $\mathbb{R}$ **53**

7 Hermitesche Formen über $\mathbb{C}$ **62**

8 Euklidische und faktorielle Ringe **72**

1 Die Jordan'sche Normalform

Wir knüpfen in dieser Vorlesung unmittelbar an die Ergebnisse der Linearen Algebra 1 an (im folgenden als [LA 1] referenziert). Erinnern wir uns daher zunächst an das letzte Resultat jener Vorlesung, über diagonalisierbare Matrizen.

Sei K ein Körper und V ein endlich dimensionaler K-Vektorraum. Wir haben in der Linearen Algebra 1 (Satz 7.17) gelernt, dass ein Endomorphismus $f : V \to V$ genau dann diagonalisierbar ist, wenn die folgenden beiden Bedingungen erfüllt sind:

i) Das charakteristische Polynom ist von der Form

$$\chi_f(X) = (X - \alpha_1)^{m_1} \dots (X - \alpha_r)^{m_r}$$

für gewisse $\alpha_i \in K$ die paarweise verschieden sind, also $\alpha_i \neq \alpha_j$ falls $i \neq j$.

ii) Für den Eigenraum V_{α_i} zum Eigenwert α_i gilt

$$m_i = \dim V_{\alpha_i}.$$

Den Exponenten m_i nennt man auch die algebraische Vielfachheit von α_i, die Zahl $\dim V_{\alpha_i}$, also die Dimension von V_{α_i} heißt auch geometrische Vielfachheit von α_i.

Falls nur die Bedingung i) gilt, so haben wir zumindest folgendes gezeigt:

Satz 1.1 *(7.21, [LA 1]) Für eine Matrix $A \in K^{n\times n}$ sind äquivalent:*

i) χ_A zerfällt vollständig in Linearfaktoren über K.

ii) A ist ähnlich zu einer oberen Dreiecksmatrix. Zur Erinnerung, das heißt dass es eine invertierbare Matrix S gibt, sodass $S \cdot A \cdot S^{-1}$ eine obere Dreiecksmatrix ist.

Natürlich ist die obere Dreiecksmatrix in ii) im allgemeinen nicht eindeutig bestimmt. Unser Ziel in diesem Kapitel ist es, die Gestalt der oberen Dreiecksmatrix noch genauer bestimmen zu können. Genauer wollen wir eine Form finden, die im wesentlichen eindeutig ist, und die Beschreibung diagonalisierbarer Endomorphismen verallgemeinert. Dafür brauchen wir zunächst einige Vorbereitungen.

Definition 1.2 *Ein Endomorphismus $f : V \to V$ eines n-dimensionalen K-Vektorraums heißt **nilpotent**, wenn es ein $k \in \mathbb{N}$ mit $f^k = \underbrace{f \circ \dots \circ f}_{k-mal} = 0$ gibt.*

Proposition 1.3 *Sei V ein endlich dimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus. Sei $n = \dim V$. Dann sind äquivalent:*

i) f *ist nilpotent*

ii) $\chi_f(X) = X^n$

iii) Es gibt eine Basis B von V, so dass $A(f)_B^B$ eine obere Dreiecksmatrix mit Nullen auf der Hauptdiagonalen ist, d.h. es gilt

$$A(f)_B^B = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$$

iv) $f^n = 0$

Beweis : i) $\Rightarrow$ ii): Angenommen $f^k = 0$. Wir zeigen $\chi_f(X) = X^n$ mit Induktion nach n. Ist $n = 1$, so muss $f = 0$ gelten, also gilt die Behauptung.

Für den Induktionsschluss nehmen wir an, die Behauptung gilt für alle nilpotenten Endomorphismen $f : W \to W$ mit dim $W < n$. Für jede Basis B von V ist wegen $f^k = 0$

$$\left(A(f)_B^B\right)^k = 0,$$

also folgt $\det A(f)_B^B = 0$.

Nach [LA 1], Lemma 7.9 ist 0 also ein Eigenwert von f. Es sei $b_1 \in V$ ein Eigenvektor zum Eigenwert 0. Wir ergänzen b_1 zu einer Basis $C = (b_1, c_2, \ldots, c_n)$ von V. Dann ist

$$A(f)_C^C = \begin{pmatrix} 0 & * & \ldots & * \\ & & A' & \\ 0 & & & \end{pmatrix}.$$

Wie im Beweis von [LA 1], Satz 7.20 sei $W = \langle c_2, \ldots, c_n \rangle$ und $g : W \to W$ der Endomorphismus, dessen Koordinatenmatrix bezüglich der Basis $(c_2, \ldots, c_n)$ die Matrix A' ist. Aus $f^k = 0$ folgt $\left(A(f)_C^C\right)^k = 0$. Nun ist $\left(A(f)_C^C\right)^k = \begin{pmatrix} 0 & * & \ldots & * \\ \vdots & & & \\ & & A'^k & \\ 0 & & & \end{pmatrix}$, also ist

$A'^k = 0$. Daraus folgt $g^k = 0$, d.h. g ist nilpotent. Nach Induktionsvoraussetzung ist also $\chi_g(X) = X^{n-1}$. Wie im Beweis von [LA 1], Satz 7.20 sieht man

$$\chi_f(X) = X\chi_g(X).$$

Also ist $\chi_f(X) = X^n$, damit gilt ii).

ii) $\Rightarrow$ iii). Da $\chi_f(X) = X^n$ vollständig in Linearfaktoren zerfällt und 0 die einzige Nullstelle von $\chi_f(X)$, also der einzige Eigenwert von f ist, folgt die Behauptung aus [LA 1], Satz 7.20.

iii) $\Rightarrow$ iv): Ist $A(f)_B^B = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$, so rechnet man nach, dass $\left(A(f)_B^B\right)^n = 0$ gilt. (ÜA) Somit ist $f^n = 0$.

iv) $\Rightarrow$ i) ist klar. $\square$

Um die obere Dreiecksgestalt der Koordinatenmatrix noch weiter zu vereinfachen, brauchen wir folgendes Lemma:

Lemma 1.4 *Sei $f : V \to V, \dim V = n$, und α ein Eigenwert von f.*

Wir setzen $g = \alpha id - f : V \to V$. Wie immer bezeichne $g^j : V \to V$ die Hintereinanderausführung $\underbrace{g \circ \ldots \circ g}_{j-mal}$. Wir setzen

$$d := \min\{j : Kern(g^j) = Kern(g^{j+1})\}.$$

Für $U = Kern(g^d)$ und $W = Bild(g^d)$ gilt dann:

i) $g(U) \subset U$ und $g(W) \subset W$, d.h. U und W sind g-invariante Unterräume.

ii) Für die Einschränkungen $g|_U : U \to U$ und $g|_W : W \to W$ gilt:

$$(g|_U)^d = 0 \text{ und } Kern(g|_W) = 0$$

iii) $V = U \oplus W$.

Beweis : Kern$(g) \subset$ Kern$(g^2) \subset$ Kern$(g^3) \ldots$ ist eine aufsteigende Kette von Unterräumen von V. (Prüfen Sie das!) Da V endlich-dimensional ist, muss es ein j mit Kern$(g^j) =$ Kern(g^{j+1}) geben. Die Zahl d ist also wohldefiniert.

i) Ist $u \in U = \text{Kern}(g^d)$, so gilt

$$g^d(g(u)) = g^{d+1}(u) = g(g^d(u)) = g(0) = 0$$

Also liegt $g(u)$ in $\text{Kern}(g^d) = U$, d.h. es ist $g(U) \subset U$.

Um einzusehen, dass $g(W) \subseteq W$ gilt, können wir ähnlich vorgehen. Sei $w = g^d(v) \in \text{Bild}(g^d)$, dann gilt

$$g(w) = g(g^d(w)) = g^d(g(w)) \in \text{Bild}(g^d).$$

d.h. es gilt $g(W) \subset W$.

ii) Definitionsgemäß ist

$$(g_{|U})^d = (g^d)_{|_U} = 0.$$

Ist $w \in W$ mit $g(w) = 0$, so gibt es ein $v \in V$ mit $w = g^d(v)$. Also ist $g^{d+1}(v) = 0$. Wegen $\text{Kern}(g^d) = \text{Kern}(g^{d+1})$ folgt $w = g^d(v) = 0$. Also ist $\text{Kern}\,(g|_W) = 0$.

iii) Wir zeigen zunächst mit Induktion nach k, dass $\text{Kern}\,(g^d) = \text{Kern}(g^{d+k})$ für alle $k \geq 1$ gilt. Für $k = 1$ ist die Behauptung klar nach Definition von d. Angenommen, es gilt $\text{Kern}\,(g^d) = \text{Kern}(g^{d+k})$. Wir haben die triviale Inklusion $\text{Kern}(g^{d+k}) \subset \text{Kern}(g^{d+k+1})$ und müssen $\text{Kern}(g^{d+k+1}) \subset \text{Kern}(g^{d+k})$ zeigen. Ist $v \in \text{Kern}(g^{d+k+1})$, so folgt $0 = g^{d+k+1}(v) = g^{d+1}(g^k(v))$, d.h. $g^k(v) \in \text{Kern}(g^{d+1}) = \text{Kern}(g^d)$. Also ist $g^{d+k}(v) = g^d(g^k v) = 0$, was den Beweis abschließt.

Nun zeigen wir $U \oplus W = V$. Sei $u \in U \cap W$, d.h. $g^d(u) = 0$ und $u = g^d(v)$ für ein $v \in V$. Es folgt $g^{2d}(v) = 0$, d.h. $v \in \text{Kern}(g^{2d}) = \text{Kern}(g^d)$. Also ist $u = g^d(v) = 0$.

Nach der Dimensionsformel für lineare Abbildungen gilt außerdem

$$\dim V = \dim\ \text{Kern}(g^d) + \dim\ \text{Bild}(g^d) = \dim U + \dim W = \dim(U \oplus W),$$

also folgt $V = U \oplus W$. □

Wir haben gezeigt, dass für einen diagonalisierbaren Endomorphismus $f : V \to V$ eines endlich-dimensionalen Vektorraums V gilt, dass V direkte Summe aller Eigenräume ist. Wir betrachten nun die allgemeineren Haupträume.

Definition 1.5 *Sei $f : V \to V, \dim V = n$, und α ein Eigenwert von f. Dann heißt*

$$H_\alpha := \{v \in V : (f - \alpha id)^k = 0 \textit{ für ein } k \in \mathbb{N}\}$$

*der **Hauptraum** von f zum Eigenwert α.*

Satz 1.6 *Sei $f : V \to V$, dim $V = n$. Das charakteristische Polynom von f zerfalle vollständig in Linearfaktoren über k, d.h. es ist*

$$\chi_f(X) = (X - \alpha_1)^{m_1} \cdot \ldots \cdot (X - \alpha_r)^{m_r}$$

mit paarweise verschiedenen $\alpha_i \in K$.

Dann gilt:

i) $H_{\alpha_i} = \text{Kern}(f - \alpha_i \text{id})^{m_i}$ für alle i

ii) $f(H_{\alpha_i}) \subset H_{\alpha_i}$ und $\dim H_{\alpha_i} = m_i$ für alle i.

iii) $V = \bigoplus_{i=1}^{r} H_{\alpha_i}$.

Beweis : Wir setzen $g = f - \alpha_1 \text{id} : V \to V$ und wenden Lemma 1.4 auf g an. Es gibt also ein $d \geq 1$ mit

$$\text{Kern}(g) \subsetneq \text{Kern}(g^2) \subsetneq \cdots \subsetneq \text{Kern}(g^d) = \text{Kern}(g^{d+1}) = \text{Kern}(g^{d+2}) = \ldots$$

Definitionsgemäß ist

$$H_{\alpha_1} = \bigcup_{k=1}^{\infty} \text{Kern}(g^k) = \text{Kern}(g^d).$$

Es sei $W = \text{Bild}(g^d)$. Nach Lemma 1.4 sind H_{α_1} und W jeweils g-invariante Unterräume mit $V = H_{\alpha_1} \oplus W$, so dass

$$(g_{|H_{\alpha_1}})^d = 0$$

gilt. Also ist $g_{|H_{\alpha_1}}$ nilpotent. Proposition 1.3 liefert $(g_{|H_{\alpha_1}})^{\dim H_{\alpha_1}} = 0$. Also ist $H_{\alpha_1} = \text{Kern}(g^{\dim H_{\alpha_1}})$. Also folgt $\text{Kern}(g^d) = \text{Kern}(g^{\dim H_{\alpha_1}})$. Ferner sind H_{α_1} und W als g-invariante Unterräume auch f-invariant. Aus $V = H_{\alpha_1} \oplus W$ folgt also $\chi_f(X) = \chi_{f_{|H_{\alpha_1}}}(X)\chi_{f_{|W}}(X)$. Ist A eine Koordinatenmatrix von f, so ist $A - \alpha_1 E_n$ eine Koordinatenmatrix von g. Also folgt

$$\begin{aligned} \chi_f(X) &= \det(XE_n - A) \\ &= \det[(X - \alpha_1)E_n - (A - \alpha_1 E_n)] \\ &= \chi_g(X - \alpha_1). \end{aligned}$$

Eine analoge Gleichung gilt für $f_{|H_{\alpha_1}}$ und $f_{|W}$. Also folgt

$$\begin{aligned} \chi_f(X) &= \chi_{f_{|H_{\alpha_1}}}(X)\chi_{f_{|W}}(X) \\ &= \chi_{g_{|H_{\alpha_1}}}(X - \alpha_1)\chi_{f_{|W}}(X) \\ &\overset{1.3}{=} (X - \alpha_1)^{\dim H_{\alpha_1}}\chi_{f_{|W}}(W). \end{aligned}$$

Nach Lemma 1.4 ist Kern$(g_{|W}) = 0$, also ist $\chi_{g_{|W}}(0) \neq 0$, woraus $\chi_{f_{|W}}(\alpha_1) \neq 0$ folgt. Also muss dim$H_{\alpha_1} = m_1$ sein und $\chi_{f_{|W}}(X) = (X - \alpha_2)^{m_2} \cdot \ldots \cdot (X - \alpha_r)^{m_r}$ gelten. Damit sind i) und ii) für α_1 bewiesen. Genauso zeigt man diese Aussagen für die anderen α_i.

iii) zeigen wir mit Induktion nach $n = \dim V$. Die Behauptung ist klar für $n = 1$. Wir nehmen also an, für jeden Endomorphismus $h : W \to W$ eines Vektorraums W der Dimension $< n$ ist W die direkte Summe der Haupträume, falls $\chi_h(X)$ vollständig in Linearfaktoren zerfällt.

Für das gegebene f mit Eigenwert α_1 haben wir oben $V = H_{\alpha_1} \oplus W$ mit einem f-invarianten Vektorraum W gezeigt, für den $\chi_{f_{|W}}(X) = \prod_{i=2}^{r}(X - \alpha_i)^{m_i}$ gilt. Wenden wir die Induktionsvoraussetzung auf $f_{|W} : W \to W$ an, so erhalten wir $W = \bigoplus_{i=2}^{r} \text{Kern}(f_{|W} - \alpha_i \text{id})^{m_i}$.

Nun ist Kern $(f_{|W} - \alpha_i \text{id}_W)^{m_i} = \text{Kern}(f - \alpha_i \text{id})^{m_i} = V\alpha_i$ für $i \geq 2$, denn offenbar ist Kern $(f_{|W} - \alpha_i \text{id}_{|W})^{m_i} \subset \text{Kern}(f - \alpha_i \text{id})^{m_i}$, und die Dimensionen beider Räume stimmen nach ii) überein. Also folgt tatsächlich $V = \bigoplus_{i=1}^{r} V\alpha_i$. □

Definition 1.7 *Für alle $m \in \mathbb{N}$ definieren wir die $(m \times m)$-**Jordanmatrix** J_m als*

$$J_m = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \in K^{m \times m}$$

Es gilt $J_m^m = 0$ (ÜA).

Satz 1.8 *Es sei $g : V \to V$ nilpotent, $\dim V = n$. Wir setzen $d = \min\{j : g^j = 0\}$. Dann gibt es $s_1, \ldots, s_d \in \mathbb{N}_0$ mit*

$$ds_d + (d-1)s_{d-1} + \cdots + 2s_2 + s_1 = n$$

und eine Basis B von V mit

$$A(g)_B^B = \begin{pmatrix} J_d & & & & & & 0 \\ & \ddots & & & & & \\ & & J_d & & & & \\ & & & \ddots & & & \\ & & & & J_1 & & \\ & & & & & \ddots & \\ 0 & & & & & & J_1 \end{pmatrix} \begin{matrix} \\ \\ \} s_d\text{-mal} \\ \\ \\ \} s_1\text{-mal} \\ \\ \end{matrix}$$

Die Zahlen $s_1, \ldots, s_d$ sind durch diese Bedingungen eindeutig bestimmt.

Beweis : Wir setzen $U_j = \text{Kern}(g^j)$. Dann ist $\{0\} = U_0 \subset U_1 \subset U_2 \subset \ldots \subset U_d = V$ eine aufsteigende Kette von Unterräumen von V. Wir zeigen zunächst $g(U_j) \subset U_{j-1}$. Sei $u \in U_j = \text{Kern}(g^j)$. Dann ist $g^{j-1}(g(u)) = g^j(u) = 0$, also $g(U_j) \subset \text{Kern} g^{j-1} = U_{j-1}$.

Für jeden Unterraum W von V mit $W \cap U_j = 0$ für ein $j \geq 1$ folgt wegen $\text{Kern} g = U_1 \subset U_j$ auch $W \cap \text{Kern} g = 0$, d.h. $\text{Kern} g_{|W} = 0$. Wir wählen nun ein Komplement W_d von U_{d-1} in V, d.h. $V = U_d = U_{d-1} \oplus W_d$. Dann ist $g(W_d) \subset g(U_d) \subset U_{d-1}, g(U_{d-1}) \subset U_{d-2}$ und $\text{Kern}(g_{|W_d}) = 0$.

Ferner gilt $U_{d-2} \cap g(W_d) = 0$, denn ist $v \in U_{d-2} \cap g(W_d)$, so ist $g^{d-2}(v) = 0$ und $v = g(w)$ für ein $w \in W_d$. Dann ist $g^{d-1}(w) = 0$, also $w \in U_{d-1} \cap W_d = 0$, woraus $v = g(w) = 0$ folgt.

Also finden wir einen Unterraum W_{d-1} von U_{d-1} mit

$$g(W_d) \subset W_{d-1} \text{ und } U_{d-1} = U_{d-2} \oplus W_{d-1}.$$

Wie oben folgt $g(U_{d-2}) \subset U_{d-3}$, $g(W_{d-1}) \subset U_{d-2}$ sowie $U_{d-3} \cap g(W_{d-1}) = 0$.

Dieses Verfahren wenden wir solange an, bis alle U_j verbraucht sind. Wir erhalten eine Zerlegung

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

mit $g(W_j) \subset W_{j-1}$ und $\text{Kern}(g_{|W_j}) = 0$ für alle j.

Nun wählen wir eine Basis $w_1^{(d)}, \ldots, w_{s_d}^{(d)}$ von W_d. Wegen $g(W_d) \subset W_{d-1}$ und $\text{Kern}(g_{|W_d}) = 0$ sind $g(w_1^{(d)}), \ldots, g(w_{s_d}^{(d)})$ linear unabhängig in W_{d-1}. Wir ergänzen sie zu einer Basis

$$(g(w_1^{(d)}), \ldots, g(w_{s_d}^{(d)}), w_1^{(d-1)}, \ldots, w_{s_{d-1}}^{(d-1)})$$

von W_{d-1}. (Hier ist $s_d = \dim W_d$ und $s_{d-1} = \dim W_{d-1} - \dim W_d$). Analog können wir die Bilder unter g dieser Basisvektoren, die in W_{d-2} liegen, zu einer Basis

$$(g^2(w_1^{(d)}), \ldots, g^2(w_{s_d}^{(d)}), g(w_1^{(d-1)}), \ldots, g(w_{s_{d-1}}^{(d-1)}), w_1^{(d-2)}, \ldots, w_{s_{d-2}}^{(d-2)})$$

von W_{d-2} ergänzen.

Dieses Verfahren führen wir so lange durch, bis wir Basen für alle W_j gefunden haben. Insgesamt erhalten wir eine Basis von $V = W_1 \oplus \cdots \oplus W_d$ der Form

$$\begin{aligned} B \;=\; & (g^{d-1}(w_1^{(d)}), g^{d-2}(w_1^{(d)}), \ldots, g(w_1^{(d)}), w_1^{(d)}, \ldots \\ & g^{d-1}(w_{s_d}^{(d)}), g^{d-2}(w_{s_d}^{(d)}), \ldots, g(w_{s_d}^{(d)}), w_{s_d}^{(d)}, \\ & g^{d-2}(w_1^{(d-1)}), g^{d-3}(w_1^{(d-1)}), \ldots, g(w_1^{(d-1)}), w_1^{(d-1)}), \ldots \\ & g^{d-2}(w_{s_{d-1}}^{(d-1)}), g^{d-3}(w_{s_{d-1}}^{(d-1)}), \ldots, g(w_{s_{d-1}}^{(d-1)}), w_{s_{d-1}}^{(d-1)}, \\ & \ldots, \\ & w_1^{(1)}, \ldots, w_{s_1}^{(1)}) \end{aligned}$$

Die Koordinatenmatrix $A(g)_B^B$ hat dann die gewünschte Form

$$A(g)_B^B = \begin{pmatrix} J_d & & & & & & & & \\ & \ddots & & & & & & & \\ & & J_d & & & & & & \\ & & & J_{d-1} & & & & & \\ & & & & \ddots & & & & \\ & & & & & J_{d-1} & & & \\ & & & & & & \ddots & & \\ & & & & & & & J_1 & \\ & & & & & & & & \ddots \\ \underbrace{\qquad}_{s_d} & & \underbrace{\qquad}_{s_{d-1}} & & & & \underbrace{\qquad}_{s_1} & & J_1 \end{pmatrix}$$

Da $A(g)_B^B$ eine $(n \times n)$-Matrix ist, folgt $n = ds_d + (d-1)s_{d-1} + \cdots + s_1$.

Jetzt wollen wir noch die Eindeutigkeit von $s_1, \ldots, s_d$ zeigen. Es ist

$$J_k^2 = \begin{pmatrix} 0 & & \\ \vdots & J_{k-1} & \\ 0 & \ldots & 0 \end{pmatrix} \text{ und } \dim \text{Kern} J_k = 1.$$

Also folgt

$$\begin{aligned}
\dim \text{Kern}(g) &= \dim \text{ Kern } A_{g,B,B} = s_1 + \cdots + s_d \\
\dim \text{Kern}(g^2) &= \dim \text{ Kern } A^2_{g,B,B} = s_1 + 2s_2 + 2s_3 + \cdots + 2s_d \\
\dim \text{Kern}(g^3) &= \dim \text{ Kern} A^3_{g,B,B} = s_1 + 2s_2 + 3s_3 + \cdots + 3s_d \\
&\cdots \\
\dim \text{Kern}(g^{d-1}) &= \dim \text{Kern} A^{d-1}_{g,B,B} = s_1 + 2s_2 + \cdots + (d-1)s_{d-1} + (d-1)s_d \\
\dim \text{Kern}(g^d) &= n.
\end{aligned}$$

Also folgt für $k = 2, \ldots, d$:

$$\dim \text{ Kern } g^k - \dim \text{Kern} g^{k-1} = s_k + \cdots + s_d.$$

Durch diese Formeln sind offenbar $s_2, \ldots, s_d$ eindeutig durch g bestimmt. Wegen $s_1 + 2s_2 + \cdots + ds_d = n$ ist auch s_1 eindeutig durch g bestimmt. □

Satz 1.8 besagt, dass der nilpotente Endomorphismus bis auf Äquivalenz eindeutig durch die Zahlen $s_1, \ldots, s_d$ festgelegt („klassifiziert“) wird. Aus dem Existenzbeweis in Satz 1.8 folgt insbesondere, dass $s_d = \dim W_d = n - \dim U_{d-1} \geq 1$ ist, da $U_{d-1} \subsetneq V$ gilt.

Für kleine Dimensionen können wir aus Satz 1.8 folgende Schlüsse ziehen: Ist $g : V \to V$ nilpotent mit $\dim V = 2$, so gilt entweder $g = 0$ oder ($g^2 = 0$ und $g \neq 0$). Ist $g^2 = 0$ und $g \neq 0$, so ist $d = 2$. Wegen $s_2 \geq 1$ und $2s_2 + s_1 = 2$ folgt schon $s_2 = 1$ und $s_1 = 0$. Die Matrix aus Satz 1.8 hat dann also die Gestalt

$$J_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Ist $g : V \to V$ nilpotent mit $\dim V = 3$, $g \neq 0$, so ist entweder $g^2 = 0$ oder ($g^3 = 0$ und $g^2 \neq 0$). Im ersten Fall $g^2 = 0$ ist $d = 2$. Aus $s_2 \geq 1$ und $2s_2 + s_1 = 3$ folgt schon $s_2 = 1, s_1 = 1$. Die Matrix aus Satz 1.8 hat also die Gestalt

$$J_3 = \begin{pmatrix} J_2 & \\ & J_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Im zweiten Fall $g^3 = 0, g^2 \neq 0$ ist $d = 3$. Aus $s_3 \geq 1$ und $3s_3 + 2s_2 + s_1 = 3$ folgt schon $s_3 = 1, s_1 = s_2 = 0$. Die Matrix aus Satz 1.8 hat also die Gestalt

$$J_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Satz 1.9 (Jordansche Normalform) *Sei $f : V \to V$ ein Endomorphismus des n-dimensionalen K-Vektorraums V, so dass $\chi_f(X) = (X-\alpha_1)^{m_1} \dot{.} \dot{.} \dot{.} (X-\alpha_r)^{m_r}$ über K vollständig in Linearfaktoren zerfällt. Hier seien die $\alpha_1, \ldots, \alpha_r$ paarweise verschieden. Dann existiert eine Basis B von V, so dass $A(f)_B^B$ folgende Gestalt hat:*

$$A(f)_B^B = \begin{pmatrix} \alpha_1 E_{m_1} + N_1 & & & \\ & \alpha_2 E_{m_2} + N_2 & & 0 \\ & & & \ddots \\ & 0 & & \alpha_r E_{m_r} + N_r \end{pmatrix} \tag{1}$$

mit nilpotenten Matrizen $N_i \in K^{m_i \times m_i}$ der Form

$$N_i = \begin{pmatrix} J_{d_i} & & & & & & 0 \\ & \ddots & & & & & \\ & & J_{d_i} & & & & \\ & & & \ddots & & & \\ & & & & J_1 & & \\ & & & & & \ddots & \\ 0 & & & & & & J_1 \end{pmatrix} \quad \begin{matrix} \} \, s_{d_i}^{(i)}\text{-mal} \\ \\ \\ \\ \} \, s_1^{(i)}\text{-mal,} \end{matrix}$$

wobei $s_1^{(i)}, \ldots, s_{d_i}^{(i)} \in \mathbb{N}_0$, $s_{d_i}^{(i)} \geq 1$ Zahlen mit $d_i s_{d_i}^{(i)} + (d_i - 1) s_{d_{i-1}}^{(i)} + \cdots + 2 s_2^{(i)} + s_1^{(i)} = m_i$ sind.

Die $s_j^{(i)}, j = 1, \ldots, d_i$ sind durch f eindeutig bestimmt. Die (bis auf Umnumerierung der Eigenwerte) eindeutig bestimmte Koordinatenmatrix $A(f)_B^B$ der Form (1) heißt Jordansche Normalform von f.

Beweis : Nach Satz 1.6 ist $V = \bigoplus_{i=1}^{r} H_{\alpha_i}$, wobei H_{α_i} den f-invarianten Hauptraum zu α_i bezeichnet. Auf H_{α_i} ist die lineare Abbildung $g_i = f - \alpha_i \mathrm{id}$ nilpotent. Wir wenden Satz 1.8 auf $g_{i|H_{\alpha_i}}$ an. Für

$$d_i = \min\{k : (f - \alpha_i \mathrm{id})^k_{|H_{\alpha_i}} = 0\}$$

gilt also: Es gibt eine Basis B_i von H_{α_i} mit

$$A(g_i|_{H_{\alpha_i}})_{B_i}^{B_i} = \begin{pmatrix} J_{d_i} & & & & & & 0 \\ & \ddots & & & & & \\ & & J_{d_i} & & & & \\ & & & \ddots & & & \\ & & & & J_1 & & \\ & & & & & \ddots & \\ 0 & & & & & & J_1 \end{pmatrix} \begin{matrix} \} s_{d_i}^{(i)} \\ \\ \\ \\ \} s_1^{(i)} \end{matrix}$$

für eindeutig bestimmte Zahlen $s_1^{(i)}, \ldots, s_{d_i}^{(i)} \in \mathbb{N}_0$ mit $d_i s_{d_i}^{(i)} + \cdots + 2s_2^{(i)} + s_1^{(i)} = \dim H_{\alpha_i} = m_i$. Nun ist

$$A(f|_{H_{\alpha_i}})_{B_i}^{B_i} = \alpha_i E_{m_i} + A_{g_i|H_{\alpha_i}}.$$

Also liefert die Basis $B = (B_1, \ldots, B_r)$ eine Koordinatenmatrix der gewünschten Gestalt. □

Für eine $n \times n$-Matrix A, so dass $\chi_A(X)$ vollständig in Linearfaktoren über K zerfällt, gibt es nach Satz 1.9 eine Matrix der Form

$$\tilde{A} = \begin{pmatrix} \alpha_1 E_{d_1} + J_{d_1} & & & & & & \\ & \ddots & & & & 0 & \\ & & \alpha_1 E_1 + J_1 & & & & \\ & & & \ddots & & & \\ & & & & \alpha_r E_{d_r} + J_{d_r} & & \\ & 0 & & & & \ddots & \\ & & & & & & \alpha_r E_1 + J_1 \end{pmatrix} \begin{matrix} \} s_{d_1}^{(1)}\text{-mal} \\ \\ \} s_1^{(1)}\text{-mal} \\ \\ \} s_{d_r}^{(r)}\text{-mal} \\ \\ \} s_1^{(r)}\text{-mal} \end{matrix}$$

konjugiert zu A. Die Matrix $\tilde{A}$ ist eindeutig bestimmt bis auf die Reihenfolge der $\alpha_1, \ldots, \alpha_r$. $\tilde{A}$ heißt **Jordansche Normalform** zu A. Die Matrizen der Form $\alpha_i E_k + J_k$ auf der Hauptdiagonalen von $\tilde{A}$ heißen **Jordankästchen** zu α_i.

Die Anzahl der Jordankästchen zum Eigenwert α_i stimmt mit der Dimension des Eigenraums zu α_i überein (ÜA).

Beispiel: Es sei $A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$.

Dann ist

$$\chi_A(X) = \det \begin{pmatrix} X-3 & -4 & -3 \\ 1 & X & 1 \\ -1 & -2 & X-3 \end{pmatrix}$$

Also ist $\alpha_1 = 2$ der einzige Eigenwert von A. Wir betrachten

$$A - 2E_3 = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix}$$

und berechnen

$$A - 2E_3 \neq 0, (A-2E_3)^2 \neq 0, (A-2E_3)^3 = 0.$$

Also ist $d_1 = 3$, woraus $s_3^{(1)} = 1, s_2^{(1)} = s_1^{(1)} = 0$ folgt. Somit ist

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

die Jordan'sche Normalform zu A. Sie besteht aus genau einem Jordankästchen.

Wir wollen nun noch die Jordan'sche Normalform benutzen, um das Minimalpolynom eines Endomorphismus zu definieren. Ist $p(X) = a_m X^m + \cdots + a_1 X + a_0$ ein beliebiges Polynom mit Koeffizienten $a_i \in K$, so können wir einen Endomorphismus $f : V \to V$ des K-Vektorraums V in p einsetzen: Es sei $p(f) : V \to V$ der Endomorphismus

$$p(f) = a_m f^m + \cdots + a_1 f + a_0 \cdot \mathrm{id},$$

wobei natürlich wieder $f^k = \underbrace{f \circ \cdots \circ f}_{k-\text{mal}}$ gilt.

Ist $A \in K^{n\times n}$, so definieren wir

$$p(A) = a_m A^m + \cdots + a_1 A + a_0 E_n \in K^{n\times n}.$$

Ist B eine Basis von V und $A = A(f)_B^B$, dann gilt $p(A) = A\big(p(f)\big)_B^B$. (Prüfen Sie das!)

Für ein Polynom $p(X) = a_m X^m + \cdots + a_1 X + a_0$ mit $a_m \neq 0$ heißt $m = \mathrm{grad}(p)$ der Grad von p. Ist $a_m = 1$, so nennt man p normiert. Sind alle $a_i = 0$, d.h. gilt $p(X) = 0$, so setzen wir $\mathrm{grad}(p) = -\infty$.

Satz 1.10 *Es sei $f : V \to V$ ein Endomorphismus V des n-dimensionalen K-Vektorraums V, so dass $\chi_f(X) = (X-\alpha_1)^{m_1} \cdot \ldots \cdot (X-\alpha_r)^{m_r}$ über K vollständig in Linearfaktoren zerfällt, wobei $\alpha_1, \ldots, \alpha_r$ paarweise verschieden sind. Für jedes α_i sei d_i die Größe des größten Jordankästchens zu α_i in der Jordanschen Normalform. Das Polynom*

$$p_f(X) = (X-\alpha_1)^{d_1} \cdot \ldots \cdot (X-\alpha_r)^{d_r}$$

hat dann folgende Eigenschaften:

i) $p_f(f) = 0$

ii) Ist $u(X) \neq 0$ ein Polynom mit $u(f) = 0$, so ist u ein Vielfaches von p_f, d.h. es gilt $u(X) = p_f(X)v(X)$ für ein Polynom $v(X)$. Das Polynom p_f heißt ***Minimalpolynom*** *von f.*

Beweis : Wir zerlegen wie im Beweis von Satz 1.9 $V = \bigoplus_{i=1}^{r} H_{\alpha_i}$. Dann gibt es eine Basis B_i von H_{α_i} mit

$$A(f-\alpha_i \mathrm{id}_{H_{\alpha_i}})_{B_i}^{B_i} = \begin{pmatrix} J_{d_i} & & & & & & \\ & \ddots & & & & & \\ & & J_{d_i} & & & & \\ & & & \ddots & & & \\ & & & & J_1 & & \\ & 0 & & & & \ddots & \\ & & & & & & J_1 \\ \underbrace{\qquad}_{s_{d_i}^{(i)}} & & & & \underbrace{\qquad}_{s_1^{(i)}} & & \end{pmatrix}$$

Für alle Vektoren $w \in H_{\alpha_i}$ gilt somit $(f-\alpha_i \mathrm{id})^{d_i}(w) = 0$, da $J_i^{d_i} = 0$ ist. Somit ist für alle Vektoren $w \in H_{\alpha_i}$ auch

$$p(f)(w) = (f-\alpha_1 \mathrm{id})^{d_1} \circ \cdots \circ (f-\alpha_r \mathrm{id})^{d_r}(w) = 0$$

Da der Endomorphismus $p(f)$ somit auf allen H_{α_i} verschwindet, gilt $p(f) = 0$.

Nun sei $u(X) \neq 0$ ein Polynom mit $v(f) = 0$. Wir wollen zeigen, dass $u(X)$ ein Vielfaches von $p_f(X)$ ist. Dazu betrachten wir zunächst den Spezialfall, dass f nilpotent auf V ist. Wir setzen $d = \min\{j : f^j = 0\}$. Dann gilt für das Polynom $w(X) = X^d$ ebenfalls $w(f) = 0$. Wir zeigen nun, dass $\mathrm{grad}(u) \geq d = \mathrm{grad}(w)$ gilt. Angenommen,

grad$(u) < d$. Dann wählen wir unter allen Polynomen $\tilde{u} \neq 0$ mit $\tilde{u}(f) = 0$ eines mit minimalem Grad aus. Also ist auch grad$(\tilde{u}) < d$. Nun benutzen wir, dass es für Polynome eine Division mit Rest gibt. Division von $w(X) = X^d$ durch $\tilde{u}$ mit Rest liefert Polynome $\tilde{v}(X), \tilde{r}(X)$ mit grad$(\tilde{r}) <$ grad$(\tilde{u})$ und $w(X) = X^d = \tilde{u}(X)\tilde{v}(X) + \tilde{r}(X)$.

Aus $w(f) = 0$ und $\tilde{u}(f) = 0$ folgt $\tilde{r}(f) = 0$. Da $\tilde{u}$ minimal gewählt wurde, muss $\tilde{r} = 0$ sein. Also ist $\tilde{u}(X) = X^{\tilde{d}}$ für ein $\tilde{d} \leq d$, d.h. es gilt $f^{\tilde{d}} = \tilde{u}(f) = 0$. Nach Definition von d folgt daraus $d \leq \tilde{d}$, was $\tilde{d} =$ grad$(u) < d$ widerspricht. Also gilt in der Tat grad$(u) \geq d$. Daher können wir $u(X)$ mit Rest durch $w(X) = X^d$ dividieren und erhalten

$$u(X) = X^d v(X) + r(X)$$

mit Polynomen $v(X), r(X)$, so dass grad $r(X) < d$ gilt.

Aus $u(f) = 0$ und $f^d = 0$ folgt durch Einsetzen von f wieder $r(f) = 0$. Da es kein Polynom $\neq 0$ vom Grad $< d$ mit dieser Eigenschaft gibt, ist $r = 0$. Somit folgt $u(X) = X^d v(X)$, d.h. $u(X)$ ist ein Vielfaches von X^d.

Damit haben wir die Behauptung im Spezialfall f nilpotent bewiesen. Jetzt können wir den allgemeinen Fall mit Induktion nach n zeigen. Für $n = 0$ ist sie klar. Wir nehmen an, sie gilt für alle Vektorräume der Dimension $< n$. Nach Lemma 1.4 ist $V = H_{\alpha_1} \oplus W$ mit einem f-invarianten Unterraum W von f, für den Kern $(f_{|W} - \alpha_1 \mathrm{id}_{|W}) = 0$ gilt. Also ist Bild$(f_{|W} - \alpha_1 \mathrm{id}_{|W}) = W$. Aus $u(f) = 0$ folgt $u(f_{|H_{\alpha_1}}) = 0$ und $u(f_{|W}) = 0$. Wir setzen $\hat{u}(X) = u(X + \alpha_1)$. Dann ist $\hat{u}(f_{|H_{\alpha_1}} - \alpha_1 \mathrm{id}_{|H_{\alpha_1}}) = u(f_{|H_{\alpha_1}}) = 0$. Da $f_{|H_{\alpha_1}} - \alpha_1 \mathrm{id}_{|H_{\alpha_1}}$ nilpotent ist, folgt nach dem schon behandelten Fall der nilpotenten Endomorphismen

$$\hat{u}(X) = X^d v_0(X)$$

für ein Polynom v_0 und die Zahl $d = \min\{j : (f_{|H_{\alpha_1}} - \alpha_1 \mathrm{id}_{|H\alpha_1})^j = 0\}$. Also ist $d = d_1$ die Größe des größten Jordankästchens zu α_1 aus der Jordan'schen Normalform. Es folgt

$$u(X) = \hat{u}(X - \alpha_1) = (X - \alpha_1)^{d_1} v_0(X - \alpha_1) = v_1(X)(X - \alpha_1)^{d_1}$$

für das Polynom $v_1(X) = v_0(X - \alpha_1)$.

Nun ist Bild$(f_{|W} - \alpha_1 \mathrm{id}_{|W}) = W$, also auch Bild$(f_{|W} - \alpha_1 \mathrm{id}_{|W})^{d_1} = W$. Aus $u(f_{|W}) = 0$ folgt $v_1(f_{|W}) \circ (f_{|W} - \alpha_1 \mathrm{id}_{|W})^{d_1} = 0$. Nun können wir jedes $w \in W$ als $w = (f - \alpha_1 \mathrm{id})^{d_1}(w')$ für ein $w' \in W$ schreiben, woraus

$$v_1(f_{|W})(w) = 0$$

folgt. Somit gilt $v_1(f_{|W}) = 0$. Wir können daher die Induktionsvoraussetzung auf $f_{|W} : W \to W$ anwenden und erhalten

$$v_1(X) = (X - \alpha_2)^{d_2} \cdot \ldots \cdot (X - \alpha_r)^{d_r} v(X)$$

für ein Polynom $v(X)$. Daraus folgt die Behauptung. □

Beispiel: Hat $f : \mathbb{R}^5 \to \mathbb{R}^5$ die Jordan'sche Normalform

$$A(f)_B^B = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix},$$

so ist $p_f(X) = (X - 3)^3(X + 1)$.

Definitionsgemäß ist das Minimalpolynom eines Endomorphismus $f : V \to V$, für den $\chi_f(X)$ vollständig in Linearfaktoren zerfällt, ein Teiler des charakteristischen Polynoms $\chi_f(X)$. Zeigen Sie, dass daraus der Satz von Cayley-Hamilton ([LA 1], Satz 7.18) folgt.

Ist der Grundkörper K algebraisch abgeschlossen (etwa $K = \mathbb{C}$), so gelten diese Ergebnisse für beliebige Endomorphismen endlich-dimensionaler K-Vektorräume. In der Algebra lernt man, dass es für jeden Körper K einen algebraisch abgeschlossenen Erweiterungskörper $\overline{K}$ gibt.

2 Gruppentheorie

Wir haben in der Linearen Algebra 1 bereits den Begriff der Gruppe kennengelernt. Das wollen wir jetzt vertiefen. Zunächst definieren wir die richtigen Abbildungen zwischen Gruppen.

Definition 2.1 *Eine Abbildung $\phi : G \to G'$ zwischen zwei Gruppen ist ein Gruppenhomomorphismus, wenn*

$$\phi(ab) = \phi(a)\phi(b) \textit{ für alle } a, b \in G$$

gilt.

Lemma 2.2 *Ist $\phi : G \to G'$ ein Gruppenhomomorphismus, so gilt $\phi(e_G) = e_{G'}$. Ferner gilt für alle $a \in G$:*

$$\phi(a^{-1}) = (\phi(a))^{-1}.$$

Die Abbildung ϕ vertauscht also mit Inversenbildung.

Beweis : Aus $e_G^2 = e_G$ folgt $\phi(e_G)^2 = \phi(e_G)$, und nach Multiplikation mit $\phi(e_G)^{-1}$ folgt $\phi(e_G) = e_{G'}$. Die zweite Behauptung folgt aus $\phi(a)\phi(a^{-1}) = \phi(e_G) = e_{G'}$. □

Ein Gruppenhomomorphismus $\phi : G \to G'$ heißt **Isomorphismus**, falls ϕ ein Inverses besitzt, d.h. falls es einen Gruppenhomomorphismus $\psi : G' \to G$ mit

$$\psi \circ \phi = \mathrm{id}_G \text{ und } \phi \circ \psi = \mathrm{id}_{G'}$$

gibt. Äquivalent dazu ist, dass der Homomorphismus ϕ bijektiv ist (ÜA).

Ein injektiver Gruppenhomomorphismus heißt **Monomorphismus**, ein surjektiver Gruppenhomomorphismus heißt **Epimorphismus**.

Einen Homomorphismus $\phi : G \to G$ von G nach G bezeichnet man auch als **Endomorphismus**, ein Isomorphismus $\phi : G \to G$ von G nach G heißt **Automorphismus**.

Sind $\phi : G \to G'$ und $\psi : G' \to G''$ Gruppenhomomorphismen, so ist auch die Verknüpfung $\psi \circ \phi : G \to G''$ ein Gruppenhomomorphismus.

Definition 2.3 *Sei $\phi : G \to G'$ ein Gruppenhomomorphismus. Dann setzen wir*

$$\mathbf{Kern}\,\phi = \{x \in G : \phi(x) = e'\},$$

wobei e' das neutrale Element von G' ist, und

$$\begin{aligned} \mathbf{Bild}\,\phi &= \phi(G) \\ &= \{y \in G' : \textit{ es gibt ein } x \in G \textit{ mit } \phi(x) = y\} \end{aligned}$$

Kern ϕ ist eine Untergruppe von G und Bild ϕ ist eine Untergruppe von G' (ÜA).

Für jeden Gruppenhomomorphismus $\phi : G \to G'$ und jede Untergruppe H' von G' ist allgemeiner das Urbild $\phi^{-1}(H') = \{g \in G : \phi(g) \in H'\}$ eine Untergruppe von G (ÜA).

Ein Gruppenhomomorphismus $\phi : G \to G'$ ist injektiv genau dann, wenn

$$\text{Kern } \phi = \{e\}$$

ist und surjektiv genau dann, wenn

$$\text{Bild } \phi = G'$$

(ÜA: Zeigen Sie das!).

Wir schreiben hier die Gruppenverknüfpung immer als gh. G könnte aber auch etwa $(\mathbb{Z}, +)$ sein.

Definition 2.4 *Es sei G eine Gruppe und $H \subset G$ eine Untergruppe. Eine **Linksnebenklasse** von H in G ist eine Teilmenge von G der Gestalt*

$$aH := \{ab : b \in H\}.$$

Lemma 2.5 *Für zwei Linksnebenklassen aH und bH von H in G sind äquivalent:*

i) $aH = bH$

ii) $aH \cap bH \neq \emptyset$

iii) $a \in bH$

iv) $b^{-1}a \in H$

Beweis : i) $\Rightarrow$ ii) ist klar, da $a \cdot 1 \in aH$ ist und somit $aH \neq \emptyset$ gilt.

ii) $\Rightarrow$ iii): Es existiere ein $c \in aH \cap bH$, also $c = ah_1 = bh_2$ für $h_1, h_2 \in H$. Daraus folgt $a = bh_2h_1^{-1} \in bH$.

iii) $\Rightarrow$ iv) folgt durch Multiplikation mit b^{-1}

iv) $\Rightarrow$ i): Aus $b^{-1}a \in H$ folgt $a \in bH$, also $aH \subset bH$. Da H eine Untergruppe von G ist, enthält sie mit $b^{-1}a$ auch $(b^{-1}a)^{-1} = a^{-1}b$. Also folgt $b \in aH$ und somit $bH \subset aH$. Insgesamt also $aH = bH$. □

Satz 2.6 *Zwischen je zwei Linksnebenklassen von H in G gibt es eine bijektive Abbildung. (Man sagt auch, sie sind gleichmächtig.)*

Zwei Linksnebenklassen sind entweder disjunkt oder gleich. G ist disjunkte Vereinigung aller Linksnebenklassen.

Beweis : Für $a, b \in G$ vermittelt die Abbildung

$$\begin{aligned} \phi: \quad & G \to G \\ & g \mapsto ba^{-1}g \end{aligned}$$

eine Bijektion $\phi : aH \to bH$.

Die zweite Behauptung folgt aus Lemma 2.5.

Da jedes $a \in G$ in der Linksnebenklasse aH liegt, folgt auch die dritte Behauptung. □

Die Elemente einer Linksnebenklasse aH werden auch als **Vertreter** (oder **Repräsentanten**) dieser Linksnebenklasse bezeichnet. Für jeden Vertreter a' von aH (also jedes $a' \in aH$) gilt $aH = a'H$ nach Lemma 2.5.

Mit G/H bezeichnen wir die Menge der Linksnebenklassen von H in G.

Ganz analog kann man auch **Rechtsnebenklassen** Ha definieren, nämlich als Teilmengen von G der Form

$$Ha = \{ha : h \in H\}.$$

Die bijektive Abbildung $G \to G$, die g auf g^{-1} schickt, induziert eine Bijektion zwischen der Menge der Linksnebenklassen G/H und der Menge der Rechtsnebenklassen, die wir mit $H\backslash G$ bezeichnen (ÜA).

Definition 2.7 *Die Anzahl der Elemente in G/H bezeichnen wir als*

$$G : H,$$

*und nennen diese Zahl den **Index von** H **in** G.*

*Mit $ord(G)$ bezeichnen wir die **Ordnung** von G, d.h. die Anzahl der Elemente in G.*

Korollar 2.8 (Satz von Lagrange) *Sei G eine endliche Gruppe und H eine Untergruppe von G. Dann gilt*

$$ord(G) = ord(H) \cdot (G : H).$$

Beweis : Das folgt aus Satz 2.6. □

Definition 2.9 *Eine Untergruppe $H \subset G$ heißt **Normalteiler** oder normale Untergruppe von G, wenn $aH = Ha$ für alle $a \in G$ gilt, d.h. wenn für jedes $a \in G$ die zugehörige Linksnebenklasse mit der Rechtsnebenklasse übereinstimmt.*

Wir nennen in diesem Fall die Nebenklasse $aH = Ha$ auch Restklasse von a modulo H.

Bemerkung 2.10 *i) H ist Normalteiler in G*
$\Leftrightarrow$ *für alle $a \in G$ ist $aHa^{-1} = H$*
$\Leftrightarrow$ *für alle $a \in G$ ist $aHa^{-1} \subset H$*

ii) In einer abelschen Gruppe ist jede Untergruppe ein Normalteiler

iii) Der Kern eines Gruppenhomomorphismus $\phi : G \to G'$ ist stets ein Normalteiler von G.

Beweis :

i) H Normalteiler, $\Rightarrow aHa^{-1} = H$ für alle $a \in G$, $\Rightarrow aHa^{-1} \subset H$ für alle $a \in G$.
Falls $aHa^{-1} \subset H$ für alle $a \in G$, so gilt auch $a^{-1}Ha \subset H$ für alle $a \in G$. Also ist $aH \subset Ha$ und $Ha \subset aH$, d.h. $aH = Ha$.

ii) klar

iii) Kern ϕ ist eine Untergruppe von G. Nach i) genügt es, für alle $a \in G$ zu zeigen: $a($ Kern $\phi)a^{-1} \subset$ Kern ϕ.
Sei also $x \in$ Kern ϕ. Dann ist

$$\begin{aligned} \phi(axa^{-1}) &\overset{2.2}{=} \phi(a)\phi(x)\phi(a^{-1}) \\ &\overset{2.5}{=} \phi(a)\underbrace{\phi(x)}_{=1}\phi(a)^{-1} \\ &= 1, \end{aligned}$$

also $axa^{-1} \in$ Kern ϕ.

□

Nun wollen wir zeigen, dass es umgekehrt zu jedem Normalteiler $N \subset G$ einen Gruppenhomomorphismus $\phi : G \to G'$ mit Kern $\phi = N$ gibt.

Dazu definieren wir eine Gruppenstruktur auf der Menge G/N der Linksnebenklassen.

Lemma 2.11 *Definiert man die Multiplikation von Teilmengen X und Y von G durch $XY = \{xy : x \in X, y \in Y\} \subset G$, so ist das Produkt der Linksnebenklassen aN und bN die Linksnebenklasse abN. Mit dieser Multiplikation wird G/N eine Gruppe mit neutralem Element $1N$ und inversem Element $a^{-1}N$ zu aN. Wir nennen sie die Faktor- oder Restklassengruppe von G modulo N.*

Die Abbildung

$$\begin{aligned} \pi : \quad & G \to G/N, \\ & a \mapsto aN \end{aligned}$$

ist ein Gruppenhomomorphismus mit Kern $\pi = N$.

Beweis : $(aN)(bN) \overset{\text{Ass.}}{=} a(Nb)N \overset{N \text{ Normalteiler}}{=} a(bN)N \overset{\text{Ass.}}{=} (ab)(NN) \overset{N \text{ Untergr.}}{=} (ab)N.$

Die Gruppenaxiome aus Definition 2.9 lassen sich leicht nachprüfen (ÜA). Außerdem folgt sofort, dass π ein Gruppenhomomorphismus ist. Es ist

$$\begin{aligned} \text{Kern } (\pi) &= \{a \in G : \pi(a) = 1 \text{ in } G/N\} \\ &= \{a \in G : aN = 1N\} \\ &\overset{2.5}{=} N \end{aligned}$$

□

Ist $(G, +)$ eine Gruppe mit der Addition als Verknüpfung und H eine Untergruppe, so schreiben wir die Linksnebenklassen als $a + H$. Ist H ein Normalteiler von G, so ist die Addition auf G/H gegeben durch $(a + H) + (b + H) = (a + b) + H$.

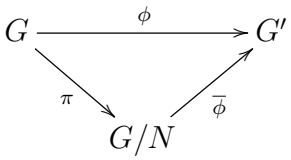
Wie viele algebraische Konstruktionen hat auch der Homomorphismus $\pi : G \to G/N$ eine universelle Eigenschaft:

Satz 2.12 (Homomorphiesatz) *Sei $\phi : G \to G'$ ein Gruppenhomomorphismus und $N \subset G$ ein Normalteiler mit $N \subset$ Kern ϕ.*

Dann existiert genau ein Gruppenhomomorphismus

$$\overline{\phi} : G/N \to G',$$

so dass



kommutiert. Mit anderen Worten, es ist $\phi = \overline{\phi} \circ \pi$.

Es gilt Bild$\overline{\phi} =$ *Bild*ϕ, *Kern*$\overline{\phi} = \pi($*Kern*$\phi)$, *Kern*$\phi = \pi^{-1}($ *Kern* $\overline{\phi})$.

Insbesondere ist $\overline{\phi}$ *genau dann injektiv, wenn* $N =$ *Kern* ϕ *gilt.*

Beweis : Für jede Abbildung $\overline{\phi}$ mit $\phi = \overline{\phi} \circ \pi$ gilt $\overline{\phi}(aN) = \phi(a)$. Daher ist $\overline{\phi}$ eindeutig bestimmt, falls es existiert. Wo liegt hier das Problem? Wir können doch einfach definieren $\overline{\phi}(aN) = \phi(a)$. Das macht aber nur dann Sinn, wenn aus $aN = bN$ schon $\phi(a) = \phi(b)$ folgt. Diese sogenannte Wohldefiniertheit von ϕ wollen wir jetzt zeigen. Aus $aN = bN$ folgt $b^{-1}aN = N$, also $b^{-1}a \in N \subset$ Kern ϕ. Somit gilt $1 = \phi(b^{-1}a) = \phi(b)^{-1}\phi(a)$, also in der Tat $\phi(a) = \phi(b)$.

$\overline{\phi}$ ist also wohldefiniert. Aus $(aN)(bN) = (ab)N$ folgt, dass $\overline{\phi}$ ein Homomorphismus ist. Offenbar ist Bild $(\overline{\phi}) =$ Bild (ϕ) und Kern $\overline{\phi} = \pi($ Kern $\phi)$. Wir zeigen noch

$$\text{Kern } \phi = \pi^{-1}(\text{ Kern } \overline{\phi})$$

„$\subset$ “ : Ist $\phi(a) = 1$, so ist $\overline{\phi}(\pi a) = \overline{\phi}(aN) = 1$.

„$\supset$ “ : Sei $a \in G$ ein Element mit $\pi(a) \in$ Kern $\overline{\phi}$, d.h. $\overline{\phi}(aN) = \overline{\phi}(\pi a) = 1$, so folgt $\phi(a) = 1$.

□

Korollar 2.13 *Ist* $\phi : G \to G'$ *ein Epimorphismus, so ist* $G' \simeq G/$ *Kern* ϕ *mit einem kanonischen („ausgezeichneten“) Isomorphismus.*

Beweis : Nach Satz 2.12 existiert ein eindeutig bestimmter Homomorphismus $\overline{\phi} : G/$ Kern $\phi \to G'$ mit $\phi = \overline{\phi} \circ \pi$. Da Kern $\overline{\phi} = \pi($ Kern $\phi) = 1$ und Bild $\overline{\phi} =$ Bild $\phi = G'$ ist, ist $\overline{\phi}$ bijektiv und somit ein Isomorphismus. □

Nun können wir die Isomorphiesätze für Gruppen beweisen.

Satz 2.14 (1. Isomorphiesatz) *Es sei* G *eine Gruppe,* $H \subset G$ *eine Untergruppe und* $N \subset G$ *ein Normalteiler. Dann ist* HN *eine Untergruppe von* G *mit Normalteiler* N*, und* $H \cap N$ *ist ein Normalteiler von* H*. Die durch* $H \subset HN$ *induzierte Abbildung*

$$H/H \cap N \to HN/N$$

ist ein Ismorphismus

Beweis : HN enthält 1 und ist abgeschlossen unter der Multiplikation, da für $h_1, h_2 \in H$ und $n_1, n_2 \in N$ gilt:

$$(h_1 n_1)(h_2 n_2) = h_1(n_1 h_2)n_2 \in h_1 h_2 N,$$

da $n_1 h_2 \in N h_2 = h_2 N$ ist. Da N ein Normalteiler in G ist, so ist N erst recht ein Normalteiler in HN.

Der Homomorphismus

$$H \hookrightarrow HN \to HN/N$$

ist surjektiv und hat offenbar den Kern $H \cap N$ (ÜA). Nach Bemerkung 2.10 ist $H \cap N$ also ein Normalteiler in H. Ferner folgt aus Korollar 2.13 der Isomorphismus

$$H/H \cap N \to HN/N.$$

□

Satz 2.15 (2. Isomorphiesatz) *Es sei G eine Gruppe und N, H zwei Normalteiler in G mit $N \subset H$. Dann ist N auch Normalteiler in H und man kann H/N als Normalteiler von G/N auffassen. Der kanonische Gruppenhomomorphismus*

$$(G/N)/(H/N) \to G/H$$

ist ein Isomorphismus.

Beweis : N ist normal in G, also auch in H. Die Inklusion $H \hookrightarrow G$ vermittelt den Gruppenhomomorphismus

$$\psi : H \hookrightarrow G \xrightarrow{\pi} G/N.$$

Dieser hat den Kern N. Nach Satz 2.12 gibt es also einen eindeutig bestimmten Gruppenhomomorphismus

$$\overline{\psi} : H/N \to G/N,$$

der

$$\begin{array}{ccc} H & \longrightarrow & G \\ \downarrow & & \downarrow \\ H/N & \xrightarrow{\overline{\psi}} & G/N \end{array}$$

kommutativ macht. Ebenfalls nach Satz 2.12 ist $\overline{\psi}$ injektiv. Mit Hilfe dieses injektiven Homomorphismus $\overline{\psi}$ identifzieren wir H/N mit der Untergruppe Bild $\overline{\psi}$ von G/N.

Ferner induziert der Epimorphismus

$$G \to G/H$$

wegen $N \subset H$ einen Epimorphismus

$$G/N \to G/H,$$

dessen Kern mit dem Bild von H unter der Quotientenabbildung $G \to G/N$, d.h. mit Bild $\overline{\psi}$ übereinstimmt. Dieses haben wir oben mit H/N identifziert. Nach Korollar 2.13 ist also

$$(G/N)/(H/N) \xrightarrow{\sim} G/H$$

□

Definition 2.16 *i) Ist H eine Teilmenge der Gruppe G, so ist die Menge*

$$\langle H \rangle = \{x_1^{\varepsilon_1} \cdots x_n^{\varepsilon_n} : n \in \mathbb{N}, x_1, \ldots, x_n \in H, \varepsilon_1, \ldots, \varepsilon_n \in \{\pm 1\}\}$$

eine Gruppe (ÜA).

*Diese heißt die **von H erzeugte Untergruppe** von G.*

*ii) G heißt **endlich erzeugt**, wenn es eine endliche Teilmenge $H \subset G$ mit $\langle H \rangle = G$ gibt.*

*iii) G heißt **zyklisch**, falls es ein $x \in G$ mit $\langle \{x\} \rangle = G$ gibt. Dafür schreiben wir auch $\langle x \rangle$.*

Offenbar ist jede zyklische Gruppe abelsch. Außerdem gilt

Satz 2.17 *i) Eine Gruppe G ist genau dann zyklisch, wenn es einen surjektiven Gruppenhomomorphismus $\mathbb{Z} \to G$ gibt.*

ii) Für eine zyklische Gruppe G gilt:

$$G \simeq \mathbb{Z}, \text{ falls } G \text{ unendlich viele Elemente enthält.}$$
$$G \simeq \mathbb{Z}/m\mathbb{Z}, \text{ falls } |G| = m \text{ für } m \in \mathbb{N}.$$

Hier ist $m\mathbb{Z} = \langle m \rangle$ die von $m \in \mathbb{N}$ erzeugte zyklische Untergruppe von $\mathbb{Z}$. Da $\mathbb{Z}$ abelsch ist, ist $m\mathbb{Z}$ ein Normalteiler in $\mathbb{Z}$.

Beweis :

i) Ist $G = \langle x \rangle = \{x^m : m \in \mathbb{Z}\}$, so definiert $m \mapsto x^m$ einen surjektiven Gruppenhomomorphismus $\mathbb{Z} \to G$.

Ist umgekehrt $\psi : \mathbb{Z} \to G$ ein surjektiver Gruppenhomomorphismus, so ist $G = \langle x \rangle$ für $x = \psi(1)$ (ÜA).

ii) Es sei $\psi : \mathbb{Z} \to G$ der surjektive Gruppenhomomorphismus aus i). Dann ist nach Korollar 2.13

$$\mathbb{Z}/\text{ Kern } \psi \simeq G.$$

Kern ϕ ist eine Untergruppe von $\mathbb{Z}$. Nach [LA 1], Satz 4.5 gibt es also ein $m \in \mathbb{Z}$ mit Kern $\psi = m\mathbb{Z}$. Ist G eine unendliche Menge, so muss $m = 0$ sein.

□

Satz 2.18 *i) Ist G eine zyklische Gruppe, so ist auch jede Untergruppe $H \subset G$ zyklisch.*

ii) Ist G zyklisch und $\phi : G \to G'$ ein Gruppenhomomorphismus, so sind auch Kern ϕ und Bild ϕ zyklisch.

Beweis : Übungsaufgabe. □

Definition 2.19 *Sei G eine Gruppe und $a \in G$. Die **Ordnung von** a ist definiert als die Ordnung von $\langle a \rangle$, der von a erzeugten zyklischen Untergruppe von G. Wir bezeichnen sie mit ord(a).*

Wir haben schon gesehen, dass

$$\begin{aligned} \psi : \mathbb{Z} &\to \langle a \rangle, \\ k &\mapsto a^k \end{aligned}$$

einen Epimorphismus definiert, und dass gilt

$$\mathbb{Z} \simeq \langle a \rangle, \text{ falls ord}(a) = \infty$$

sowie

$$\mathbb{Z}/m\mathbb{Z} \simeq \langle a \rangle, \text{ falls ord}(a) = m \in \mathbb{N}.$$

Ist ord$(a) = m \in \mathbb{N}$, so folgt aus diesem Isomorphismus:

$$a^k = 1 \Leftrightarrow k \in m\mathbb{Z}.$$

Insbesondere ist $m = \text{ord}(a)$ also die kleinste positive Zahl k mit $a^k = 1$. Ferner gilt

$$\langle a \rangle = \{1, a, a^2, \ldots, a^{m-1}\}.$$

Satz 2.20 (Kleiner Fermat'scher Satz) *Sei G eine endliche Gruppe und $a \in G$. Dann ist $ord(a)$ ein Teiler von $ord(G)$, und es gilt*

$$a^{ord(G)} = 1.$$

Beweis : Aus dem Satz von Lagrange folgt $\text{ord}(G) = \text{ord}(a) \cdot (G : \langle a \rangle)$. Also gilt $\text{ord}(a) \mid \text{ord}(G)$. Daraus folgt $a^{\text{ord}(G)} = 1$. □

Korollar 2.21 *Es sei G eine Gruppe, so dass $ord(G) = p$ eine Primzahl ist. Dann ist G zyklisch, $G \simeq \mathbb{Z}/p\mathbb{Z}$ und für jedes $a \in G$ mit $a \neq 1$ gilt $ord(a) = p$. Jedes $a \neq 1$ aus G ist ein Erzeuger von G.*

Beweis : Es sei $a \neq 1$ ein Element in G. Dann ist $\text{ord}(a) > 1$ und nach Satz 2.20 ist $\text{ord}(a)$ ein Teiler von $\text{ord}(G) = p$. Also ist $\text{ord}(a) = p$. Somit ist $\langle a \rangle$ eine Untergruppe von G mit $p = \text{ord}(G)$ Elementen. Daraus folgt $\langle a \rangle = G$. Somit ist a ein Erzeuger von G, G ist zyklisch und nach Satz 2.17 folgt $G \simeq \mathbb{Z}/p\mathbb{Z}$. □

Lemma 2.22 *Ist G eine abelsche Gruppe und $g, h \in G$ zwei Elemente endlicher Ordnung, so enthält G ein Element der Ordnung $kgV\big(ord(g), ord(h)\big)$.*

Beweis : Übungsaufgabe. □

3 Dualräume und Quotientenräume

Wir wollen nun noch Summen, Dualräume und Quotientenräume von Vektorräumen kennen lernen.

Ist V ein K-Vektorraum mit Unterräumen W_1 und W_2, so ist $W_1 + W_2$ ein Unterraum von V. Ist $W_1 \cap W_2 = 0$, so gilt $W_1 + W_2 = W_1 \oplus W_2$. Man kann aber auch für beliebige Vektorräume, die nicht als Unterräume desselben V gegeben sind, eine direkte Summe definieren.

Definition 3.1 *Es seien zwei K-Vektorräume V_1 und V_2 gegeben. Wir definieren eine Addition und eine skalare Multiplikation auf der Produktmenge $V_1 \times V_2 = \{(v_1, v_2) : v_1 \in V_1, v_2 \in V_2\}$ durch:*

i) $(v_1, v_2) + (w_1, w_2) = (v_1 + w_1, v_2 + w_2)$

ii) $a(v_1, v_2) = (av_1, av_2)$.

Lemma 3.2 *Die Abbildungen aus Definition 3.1 machen $V_1 \times V_2$ zu einem K-Vektorraum, den wir mit $V_1 \oplus V_2$ bezeichnen und **direkte Summe** von V_1 und V_2 nennen.*

Beweis : Offenbar ist $V_1 \times V_2$ mit der komponentenweisen Addition eine abelsche Gruppe mit neutralem Element $(0, 0)$ und Inversem $(-v_1, -v_2)$ zu (v_1, v_2). Es gilt ferner $1(v_1, v_2) = (v_1, v_2)$ und $(ab)(v_1, v_2) = a(b(v_1, v_2))$ sowie die Distributivgesetze. □

Man kann diese Konstruktion auf endlich viele Vektorräume $V_1, \ldots, V_n$ übertragen und so die direkte Summe $V_1 \oplus \cdots \oplus V_n$ definieren. Bei unendlich vielen Summanden muss man etwas aufpassen. Ist I eine beliebige Indexmenge und ist für jedes $i \in I$ ein K-Vektorraum V_i gegeben, so ist

$$\bigoplus_{i \in I} V_i = \{(v_i)_{i \in I} : v_i \in V_i, \text{ fast alle } v_i = 0\}$$

mit der komponentenweisen Addition und skalaren Multiplikation.

Lemma 3.3 *Es seien K-Vektorräume V und W gegeben. Mit $Hom_K(V, W)$ bezeichnen wir die Menge der linearen Abbildungen von V nach W. Dann ist $Hom_K(V, W)$ zusammen mit der Addition*

$$(f + g)(v) = f(v) + g(v)$$

und der skalaren Multiplikation

$$(af)(v) = af(v)$$

ein K-Vektorraum.

Beweis : Das folgt sofort aus den Vektorraumaxiomen für V. □

Definition 3.4 *Der K-Vektorraum $Hom_K(V, K)$ heißt **Dualraum von V**, wir bezeichnen ihn mit $V^\vee$. Die Elemente aus $V^\vee$ heißen auch **Linearformen auf V**.*

Lemma 3.5 *Sind V und W endlich-dimensionale K-Vektorräume, so gilt $\dim_K Hom(V, W) = \dim_K V \dim_K W$. Insbesondere ist $\dim_K V^\vee = \dim_K V$.*

Beweis : Es sei $n = \dim_K V$ und $m = \dim_K W$. Wir wählen Basen $B = (b_1, \ldots, b_n)$ von V und $C = (c_1, \ldots, c_m)$ von W. Wir betrachten die Abbildung

$$\begin{aligned} \Phi : \mathrm{Hom}_K(V, W) &\to K^{m \times n} \\ f &\mapsto A(f)_C^B, \end{aligned}$$

die jeder linearen Abbildung ihre Koordinatenmatrix bezüglich B und C zuordnet. Da $A(f+g)_C^B = A(f)_C^B + A(g)_C^B$ und $A(\alpha f)_C^B = \alpha A(f)_C^B$ gilt, ist Φ eine lineare Abbildung. Ist $A(f)_C^B = 0$, so folgt $f = 0$, d.h. Kern$\Phi = 0$. Ist $A = (a_{ij})$ eine beliebige Matrix in $K^{m \times n}$, so gibt es eine lineare Abbildung $f : V \to W$ mit

$$f(b_j) = \sum_{i=1}^{m} a_{ij} c_i$$

für alle $j = 1, \ldots, n$. Dann ist $A(f)_C^B = A$. Also ist Φ surjektiv und somit ein Isomorphismus von K-Vektorräumen, woraus $\dim_K \mathrm{Hom}_K(V, W) = \dim_K K^{m \times n} = mn$ folgt. □

Nun wollen wir den Quotienten eines Vektorraums nach einem Unterraum definieren.

Es sei V ein K-Vektorraum und $U \subset V$ ein Unterraum.

Da V eine abelsche Gruppe ist, existiert die Faktorgruppe V/U. Wir definieren eine skalare Multiplikation

$$K \times V/U \to V/U$$

durch $(a, v + U) = av + U$. Das hängt nicht von der Wahl des Vertreters v ab (ÜA).

Lemma 3.6 *Die abelsche Gruppe V/U zusammen mit der oben definierten skalaren Multiplikation ist ein Vektorraum.*

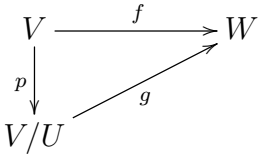
Beweis : Übungsaufgabe. □

Die Abbildung $p : V \to V/U$, definiert durch $p(x) = x + U$, ist linear und surjektiv. Ferner gilt Kern$(p) = U$, denn es ist $p(x) = 0 \Leftrightarrow x + U = U \Leftrightarrow x \in U$. Daraus folgt mit der Dimensionsformel $\dim V/U = \dim V - \dim U$. Wir haben also „$U$ aus V herausgeteilt“. Wir nennen V/U den Quotientenraum von V nach U.

Satz 3.7 (Homomorphiesatz für Vektorräume) *Es sei $f : V \to W$ eine lineare Abbildung mit $U \subset$ Kern(f). Dann existiert genau eine lineare Abbildung*

$$g : V/U \to W,$$

so dass das Diagramm



kommutiert, d.h. so dass $f = g \circ p$ gilt.

Beweis : Nach Satz 2.12 existiert genau ein Gruppenhomomorphismus $g : V/U \to W$, der das Diagramm kommutativ macht. Für jedes $a \in K$ und $v + U \in V/U$ gilt $g(av + U) = f(av) = af(v) = ag(v + U)$, also ist g linear.

Ferner ist g durch die Gleichung $f = g \circ p$ offenbar eindeutig bestimmt. □

Beispiel: Es sei $f : V = \mathbb{R}^3 \to \mathbb{R}^2$ die lineare Abbildung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_3 \\ 2x_3 \end{pmatrix}.$$

und U der von e_1, e_2 erzeugte Unterrraum des $\mathbb{R}^3$.

Dann ist $U \subset$ Kern$f$. Die Abbildung

$$\begin{array}{ccc} g : V/U & \to & \mathbb{R}^2 \\ \left[\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right] & \mapsto & \begin{pmatrix} x_3 \\ 2x_3 \end{pmatrix} \end{array}$$

leistet das Verlangte. g ist die Verknüpfung des Isomorphismus $V/U \to \mathbb{R}$,

$$\left[\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right] \mapsto x_3,$$

mit der linearen Abbildung $\mathbb{R} \to \mathbb{R}^2$, die durch $\lambda \mapsto \begin{pmatrix} \lambda \\ 2\lambda \end{pmatrix}$ gegeben ist.

4 Ringe

Definition 4.1 *Ein **Ring** (manchmal auch Ring mit Eins genannt) ist eine Menge R zusammen mit zwei Abbildungen:*

$$\begin{aligned} + : &\quad R \times R \to R \quad \textit{(Addition) und} \\ \cdot : &\quad R \times R \to R \quad \textit{(Multiplikation),} \end{aligned}$$

so dass gilt

i) $(R, +)$ *ist eine abelsche Gruppe*

ii) Die Multiplikation ist assoziativ und es gibt ein $e \in R$ *mit* $ea = a$ *und* $ae = a$ *für alle* $a \in R$*.*

iii) für alle $a, b, c \in R$ *gilt*

$$\begin{aligned} (a+b)c &= ac + bc \quad \textit{und} \\ c(a+b) &= ca + cb \end{aligned}$$

(Distributivgesetze)

Wie gewohnt, soll immer die Multiplikation vor der Addition ausgeführt werden, d.h. wir schreiben $ac + bc = (ac) + (bc)$*.*

R *heißt **kommutativ**, falls die Multiplikation kommutativ ist.*

Wir werden hier im wesentlichen kommutative Ringe betrachten. Das neutrale Element bezüglich der Addition bezeichnen wir mit 0, das neutrale Element bezüglich der Multiplikation mit 1. Der kleinste Ring ist der Nullring $\{0\}$, den wir auch mit 0 bezeichnen. In diesem gilt $1 = 0$.

Lemma 4.2 *In jedem Ring* R *gilt*

i) $0a = a0 = 0$ *für alle* $a \in R$

ii) $(-a)b = -(ab) = a(-b)$ *für alle* $a, b \in R$

iii) Hat R *mindestens zwei Elemente, so ist* $1 \neq 0$*.*

Beweis :

i) Aus den Distributivgesetzen folgt $0a = (0+0)a = 0a+0a$, also $0a = 0$. Genauso folgt $a0 = 0$.

ii) $(-a)b + (ab) = (-a + a)b = 0b \overset{\text{i)}}{=} 0$. Die zweite Gleichung folgt genauso.

iii) Enthält R ein $a \neq 0$, so gilt $0a \overset{\text{i)}}{=} 0 \neq a$, daher kann 0 nicht das Einselement sein.

□

Im allgemeinen gilt **nicht**, dass aus $ab = ac$ schon $b = c$ folgt! Finden Sie ein Beispiel.

Definition 4.3 *Eine Teilmenge S des Ringes R heißt **Unterring** von R, falls S bezüglich der Addition eine Untergruppe von $(R, +)$ und bezüglich der Multiplikation ein Untermonoid von $(R, \cdot)$ ist.*

Insbesondere ist S mit den eingeschränkten Verknüpfungen selbst ein Ring. Wir nennen $S \subset R$ dann auch eine Ringerweiterung. Für jeden Ring R bezeichnen wir mit

$$R^* = \{a \in R : \text{ es existiert ein } b \in R \text{ mit } ab = ba = 1\}.$$

*die Menge der multiplikativ invertierbaren Elemente. R^* ist eine Gruppe bzgl. der Multiplikation (ÜA). Wir bezeichnen sie auch als **Einheitengruppe**.*

*R heißt **Schiefkörper**, falls $R \neq 0$ und $R^* = R \setminus \{0\}$ gilt. Das heißt, es ist $1 \neq 0$ und jedes Element $\neq 0$ ist multiplikativ invertierbar.*

*Ein Element $a \in R$ heißt **Nullteiler**, falls es ein $b \neq 0$ in R gibt mit $ab = 0$ oder $ba = 0$.*

In Schiefkörpern oder Körpern ist das Nullelement der einzige Nullteiler (ÜA).

*Ein kommutativer Ring $R \neq 0$ heißt **nullteilerfrei** oder **Integritätsring**, wenn R außer 0 keine Nullteiler enthält.*

Ein kommutativer Schiefkörper ist ein Körper (ÜA).

In Integritätsringen gilt die Kürzungsregel, d.h. aus $ab = ac$ für $a \neq 0$ folgt $b = c$ (ÜA).

Beispiel:

i) $\mathbb{Z}$ ist ein Integritätsring, $\mathbb{Z}^* = \{1, -1\}$.

ii) $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ sind Körper.

iii) Für jeden Körper K ist

$$K^{n\times n} = \text{ein Ring mit Einheitengruppe } \mathrm{GL}(n, K)$$

Dieser Ring ist für $n \geq 2$ nicht kommutativ und besitzt Nullteiler $\neq 0$.

iv) Für eine Menge X und einen Ring R ist

$$R^X = \{f : X \to R\}$$

ein Ring, wenn man die Addition und Multiplikation von Funktionen wie gewohnt als $(f + g)(x) = f(x) + g(x)$ und $(fg)(x) = f(x)g(x)$ definiert.

v) Ist I eine Indexmenge und $(R_i)_{i\in I}$ eine Familie von Ringen, so ist

$$\prod_{i\in I} R_i = \{(x_i)_{i\in I} : x_i \in R_i\}$$

ein Ring bezüglich der komponentenweisen Addition und Multiplikation.

Von nun an betrachten wir nur noch kommutative Ringe. **Daher ist ab sofort mit „Ring" immer „kommutativer Ring" gemeint.**

Eine Abbildung $\phi : S \to R$ zwischen zwei Ringen heißt **Ringhomomorphismus**, falls ϕ in folgendem Sinn verträglich mit Addition und Multiplikation ist. Für alle a, b in R gilt:

i) $\phi(a + b) = \phi(a) + \phi(b)$

ii) $\phi(1) = 1$

iii) $\phi(ab) = \phi(a)\phi(b)$.

Definition 4.4 (Polynomring) *Es sei R ein Ring. Der zugehörige Polynomring $R[X]$ in einer Unbestimmten X besteht aus allen formalen Summen*

$$\sum_{i=0}^{n} a_i X^i, \quad n \in \mathbb{N}_0, a_i \in R.$$

Addition und Multiplikation werden gegeben durch

$$\sum_{i=0}^{n} a_i X^i + \sum_{i=0}^{m} b_i X^i = \sum_{i=0}^{\max(n,m)} (a_i + b_i) X^i,$$

und

$$\left(\sum_{i=0}^{n} a_i X^i\right)\left(\sum_{i=0}^{m} b_i X^i\right) = \sum_{i=0}^{n+m}\left(\sum_{p+q=i} a_p b_q\right) X^i.$$

Hier setzen wir

$$a_i = 0 \quad \textit{für} \quad i \geq n+1 \textit{ und}$$
$$b_i = 0 \quad \textit{für} \quad i \geq m+1.$$

Das Nullelement ist das Nullpolynom $0 = 0X^0$, das Einselement das Polynom $1 = 1X^0$.

Die natürliche Abbildung

$$\begin{aligned} R &\to R[X] \\ a &\mapsto aX^0 \end{aligned}$$

ist ein injektiver Ringhomomorphismus (ÜA). Sie identifiziert R mit dem Unterring

$$\{aX^0 : a \in R\} \subset R[X]$$

Wir schreiben daher auch einfach $a = aX^0$.

Lemma 4.5 *Für eine Ringerweiterung $R \subset R'$ und ein Element $c \in R'$ ist die Abbildung*

$$\begin{aligned} \phi : R[X] &\to R' \\ f(X) = \sum_{i=0}^{n} a_i X^i &\mapsto f(c) = \sum_{i=0}^{n} a_i c^i \end{aligned}$$

*ein Ringhomomorphismus. Dieser heißt **Einsetzungshomomorphismus**.*

Beweis : Offenbar ist $f(c) + g(c) = (f+g)(c)$ für alle $f, g \in R[X]$. Ferner gilt für $f(X) = \sum_{i=0}^{n} a_i X^i$ und $g(X) = \sum_{j=0}^{m} b_j X^j$:

$$\begin{aligned} f(c)g(c) &= \left(\sum_{i=0}^{n} a_i c^i\right)\left(\sum_{j=0}^{m} b_j c^j\right) \\ &= \sum_{i=0}^{n}\sum_{j=0}^{m} a_i c^i b_j c^j \\ &= \sum_{i=0}^{n}\sum_{j=0}^{m} a_i b_j c^{i+j} \\ &= (fg)(c). \end{aligned}$$

Hier ist es wichtig, dass R' - wie alle unsere Ringe nun - kommutativ ist.

Da außerdem $\phi(1X^0) = 1$ gilt, ist ϕ ein Ringhomomorphismus. □

Für ein Polynom $f(X) = \sum_{i=0}^{n} a_i X^i \in R[X]$ nennen wir a_i den Koeffizienten vom Grad i von f. Der Koeffizient vom Grad 0 heißt auch **Absolutkoeffizient.** Außerdem definieren wir den **Grad** von f als

$$\text{grad } (f) := \max\{i : a_i \neq 0\},$$

falls $f \neq 0$ ist. Ist $f = 0$, so setzen wir $\text{grad}(f) = -\infty$. Ist $f \neq 0$ und $n = \text{grad}(f)$, so heißt a_n höchster Koeffizient oder **Leitkoeffizient** von f. Ist $a_n = 1$, so heißt f **normiert.** Jedes $f \in R[X] \setminus \{0\}$, dessen Leitkoeffizient a_n eine Einheit in R ist, lässt sich durch Multiplikation mit a_n^{-1} normieren.

Bemerkung 4.6 *Für $f, g \in R[X]$ gilt*

$$\begin{aligned} \textit{grad}(f+g) &\leq \max(\textit{grad}(f), \textit{grad}(g)) \\ \textit{grad}(fg) &\leq \textit{grad}(f) + \textit{grad}(g). \end{aligned}$$

Ist R ein Integritätsring, so gilt sogar

$$\textit{grad}(fg) = \textit{grad}(f) + \textit{grad}(g).$$

Hier setzen wir $(-\infty) + (-\infty) = -\infty$ und $(-\infty) + n = -\infty$ für alle $n \in \mathbb{N}_0$.

Beweis : Falls f oder g das Nullpolynom ist, stimmt die Behauptung offenbar. Wir können also annehmen $n = \text{grad}(f) \geq 0$ und $m = \text{grad}(g) \geq 0$. Ist $f(X) = \sum_{i=0}^{n} a_i X^i$ und $g(X) = \sum_{i=0}^{m} b_i X^i$, so ist $a_i + b_i = 0$ für $i > \max(m, n)$, also folgt

$$\text{grad}(f+g) \leq \max(\text{grad}(f), \text{grad}(g)).$$

Außerdem ist $\sum_{p+q=i} a_p b_q = 0$ für $i > m + n$, also $\text{grad}(fg) \leq \text{grad}(f) + \text{grad}(g)$. Der Koeffizient vom Grad $m + n$ in fg ist $a_n b_m$. Ist R ein Integritätsring, so ist $a_n b_m \neq 0$, also

$$\text{grad}(fg) = m + n = \text{grad}(f) + \text{grad}(g).$$

□

Lemma 4.7 *Es sei R ein Integritätsring. Dann ist auch der Polynomring $R[X]$ ein Integritätsring. Ferner gilt $R[X]^* = R^*$.*

Beweis : $f \in R[X]$ sei ein Nullteiler in $R[X]$ mit $fg = 0$ für ein $g \neq 0$. Da R ein Integritätsring ist, folgt aus Lemma 4.7 $-\infty = \text{grad}(0) = \text{grad}(f) + \text{grad}(g)$. Da $\text{grad}(g) \geq 0$ ist, muss hier $\text{grad}(f) = -\infty$, d.h. $f = 0$ sein.

Wir zeigen jetzt $R[X]^* = R^*$. Die Inklusion „$\supset$“ ist klar. Sei also $f \in R[X]^*$, d.h. es existiert ein $g \in R[X]$ mit $fg = 1$. Aus Lemma 4.7 folgt wieder $0 = \text{grad}(fg) = \text{grad}(f) + \text{grad}(g)$, also $f, g \in R$. □

In Polynomringen lässt sich eine **Division mit Rest** durchführen:

Satz 4.8 *Sei R ein Ring und $g = \sum_{i=0}^{d} a_i X^i \in R[X]$ mit $a_d \in R^*$. Dann gibt es zu jedem $f \in R[X]$ eindeutig bestimmte Polynome $q, r \in R[X]$ mit*

$$f = qg + r, \quad \textit{grad}(r) < d.$$

Beweis : Ist q ein Polynom vom Grad n mit Leitkoeffizient $c_n \neq 0$, so ist der Koeffizient vom Grad $n + d$ in qg gerade $a_d c_n$. Da a_d eine Einheit ist, folgt $a_d c_n \neq 0$. Also ist

$$\text{grad}(qg) = \text{grad}(q) + \text{grad}(g).$$

Wir zeigen zunächst die Eindeutigkeit der Division mit Rest. Gilt

$$f = qg + r = q'g + r'$$

mit grad $r < d$ und grad $r' < d$, so folgt

$$0 = (q - q')g + (r - r'),$$

also

$$\begin{aligned} d > \text{grad}(r - r') &= \text{grad}((q - q')g) \\ &\overset{\text{s.o.}}{=} \text{grad}(q - q') + \text{grad}(g) \\ &= \text{grad}(q - q') + d. \end{aligned}$$

Das kann nur stimmen, wenn $q - q' = 0$, also $q = q'$ ist. Daraus folgt $r = r'$ und somit die Eindeutigkeit.

Die Existenz zeigen wir mit Induktion nach $n = \text{grad}(f)$. Für grad $(f) < d$ können wir $q = 0$ und $r = f$ nehmen. Ist $f = \sum_{i=0}^{n} c_n X^i$ mit $c_n \neq 0$ und $n \geq d$, so ist

$$f_1 = f - c_n a_d^{-1} X^{n-d} g$$

ein Polynom mit grad$(f_1) < n$. Dies besitzt nach Induktionsvoraussetzung eine Zerlegung

$$f_1 = q_1 g + r_1$$

mit $q_1, r_1 \in R[X]$, grad$(r_1) < d$. Aus

$$f = (q_1 + c_n a_d^{-1} X^{n-d}) g + r_1$$

folgt die gewünschte Zerlegung von f. □

Der Beweis gibt uns auch ein Verfahren zur Bestimmung von q und r an die Hand, das wir an einem Beispiel vorführen wollen.

Beispiel $f = X^5 + 3X^4 + X^3 - 6X^2 - X + 1$, $g = X^3 + 2X^2 + X - 1$ in $\mathbb{Z}[X]$

$$\begin{array}{rl}
 & (X^5 \quad +3X^4 + \ X^3 - 6X^2 - X \quad +1) : (X^3 + 2X^2 + X - 1) = X^2 + X - 2 = q(X) \\
- & (X^5 \quad +2X^4 + \ X^3 - X^2) \\
\hline
 & \quad X^4 \qquad -5X^2 - X \quad +1 \\
- & \quad (\ X^4 \ +2X^3 + X^2 \ -X) \\
\hline
 & \qquad -2X^3 - 6X^2 \qquad +1 \\
- & \quad (-2X^3 - 4X^2 - 2X \quad +2) \\
\hline
 & \qquad\qquad -2X^2 + 2X \ -1 \quad = r(X)
\end{array}$$

Also ist $f(X) = (X^2 + X - 2)g(X) + (-2X^2 + 2X - 1)$

Satz 4.9 *Ein Polynom $f(X) \in K[X]$ vom Grad $n \geqq 1$ über einem Körper K hat höchstens n verschiedene Nullstellen in K.*

Beweis : (mit Induktion nach n). Hat f den Grad 1, so gilt

$$f(X) = aX + b$$

für $a, b \in K, a \neq 0$. In diesem Fall hat f genau eine Nullstelle, nämlich $-\frac{b}{a}$.

Für den Induktionsschritt nehmen wir an, dass jedes Polynom vom Grad $n-1$ höchstens n verschiedene Nullstellen hat und betrachten ein $f \in K[X]$ vom Grad n. Hat f keine Nullstelle in K, so gilt unsere Behauptung trivialerweise. Ist $a \in K$ eine Nullstelle, so gilt $f(a) = 0$. Division mit Rest von f durch $(X-a)$ liefert ein $g(X) \in K[X]$ und ein konstantes Polynom $r \in K$ mit $f(X) = (X-a)g(X)+r$. Durch Einsetzen von a erhalten wir $r = 0$. Nun hat $g \in K(X)$ den Grad $n-1$ und nach Induktionsvoraussetzung höchstens $n-1$ verschiedene Nullstellen in K. Daraus folgt die Behauptung. □

Wir lernen nun Ideale in Ringen kennen. Diese haben eine ähnliche Funktion wie Normalteiler in Gruppen, man kann nämlich in der Welt der Ringe Quotienten nach Idealen bilden. Allerdings sind Ideale im allgemeinen keine Unterringe.

Definition 4.10 *Es sei R ein Ring (kommutativ, wie immer!). Eine Teilmenge $\mathfrak{a} \subset R$ heißt **Ideal** in R, falls gilt*

i) $\mathfrak{a}$ ist eine additive Untergruppe von R

ii) für alle $r \in R, a \in \mathfrak{a}$ gilt $ra \in \mathfrak{a}$.

Jeder Ring enthält die trivialen Ideale $\{0\}$ (auch mit 0 bezeichnet) und R. Ist R ein Körper, so sind dies die einzigen Ideale (ÜA). Für Ideale $\mathfrak{a}, \mathfrak{b} \subset R$ sind auch folgende Teilmengen von R Ideale (ÜA):

$$\begin{aligned} \mathfrak{a}+\mathfrak{b} &:= \{a+b : a \in \mathfrak{a}, b \in \mathfrak{b}\} \\ \mathfrak{a}\mathfrak{b} &:= \left\{ \sum_{i=1}^{n} a_i b_i : a_i \in \mathfrak{a}, b_i \in \mathfrak{b}, n \in \mathbb{N} \right\} \\ \mathfrak{a} \cap \mathfrak{b} & \end{aligned}$$

Offenbar ist $\mathfrak{a}\mathfrak{b} \subset \mathfrak{a} \cap \mathfrak{b}$ sowie $\mathfrak{a} \subset \mathfrak{a}+\mathfrak{b}$, $\mathfrak{b} \subset \mathfrak{a}+\mathfrak{b}$ (ÜA).

Analog ist das Produkt von endlich vielen Idealen in R wieder ein Ideal. Auch der Schnitt beliebig vieler Ideale $(\mathfrak{a}_i)_{i \in I}$ sowie die Summe beliebig vieler Ideale.

$$\sum_{i \in I} \mathfrak{a}_i = \{ \sum_{i \in I} a_i : a_i \in \mathfrak{a}_i \text{ für alle } i \in I, \text{ fast alle } a_i = 0 \}$$

sind Ideale in R (ÜA). Hier schreiben wir „ fast alle“ für „ alle bis auf endlich viele“.

Für jedes $a \in R$ nennt man

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

das von a erzeugte **Hauptideal**.

Für endlich viele Elemente $a_1, \ldots, a_n \in R$ ist

$$(a_1, \ldots, a_n) := Ra_1 + \cdots + Ra_n = \{r_1a_1 + \cdots + r_na_n : r_1, \ldots, r_n \in R\}$$

ein Ideal in R. Es heißt das von $a_1, \ldots, a_n$ erzeugte Ideal und ist das kleinste Ideal in R, das $a_1, \ldots, a_n$ enthält. Mit anderen Worten, jedes Ideal $\mathfrak{a}$ mit $\{a_1, \ldots, a_n\} \subset \mathfrak{a}$ erfüllt $(a_1, \ldots, a_n) \subset \mathfrak{a}$. (ÜA).

Analog erzeugen beliebig viele Elemente $a_i, i \in I$ aus R ein Ideal:

$$(\{a_i : i \in I\}) := \sum_{i \in I} Ra_i = \left\{ \sum_{i \in I} r_i a_i : r_i \in R, \text{ fast alle } r_i = 0 \right\}.$$

Definition 4.11 *Es sei $\mathfrak{a}$ ein Ideal des Ringes R. Eine Familie $(a_i)_{i \in I}$ von Elementen aus $\mathfrak{a}$ heißt **Erzeugendensystem** von $\mathfrak{a}$, wenn $\mathfrak{a} = \sum_{i \in I} Ra_i$ gilt. $\mathfrak{a}$ heißt **endlich erzeugt**, wenn $\mathfrak{a}$ ein endliches Erzeugendensystem besitzt. Ferner heißt $\mathfrak{a}$ **Hauptideal**, wenn $\mathfrak{a}$ von einem einzigen Element erzeugt ist, d.h. wenn $\mathfrak{a} = (a)$ für ein $a \in R$ gilt.*

*Ist R ein Integritätsring, in dem jedes Ideal ein Hauptideal ist, so nennen wir R einen **Hauptidealring**.*

Beispiel:

i) In jedem Ring sind die trivialen Ideale $0 = (0)$ und $R = (1)$ Hauptideale.

ii) In $\mathbb{Z}$ sind die Untergruppen $m\mathbb{Z}$ auch Ideale, und zwar Hauptideale, da $m\mathbb{Z} = (m)$ ist.

Bemerkung 4.12 *In einem Integritätsring R sind zwei Hauptideale $\mathfrak{a} = (a)$ und $\mathfrak{b} = (b)$ genau dann gleich, wenn es ein $c \in R^*$ mit $b = ca$ gibt.*

Beweis : Es sei $\mathfrak{a} = \mathfrak{b}$. Wir können annehmen $\mathfrak{b} \neq 0$, also $b \neq 0$. Da $b \in \mathfrak{a}$ ist, existiert ein $c \in R$ mit $b = ca$. Analog existiert wegen $a \in \mathfrak{b}$ ein $c' \in R$ mit $a = c'b$. Damit folgt $b = ca = cc'b$, also

$$(1 - cc')b = 0.$$

Da R ein Integritätsring ist, ist $b \neq 0$ kein Nullteiler, also folgt $1 = cc'$. Somit ist $c \in R^*$.

Ist umgekehrt $b = ca$ für $c \in R^*$, so folgt $a = c^{-1} \cdot b$, also $(b) \subset (a)$ und $(a) \subset (b)$. $\square$

Wir nennen zwei Elemente a, b eines Ringes R **assoziiert**, wenn es ein $c \in R^*$ mit $b = ca$ gibt.

Beispiel: Wir betrachten den Polynomring $\mathbb{Z}[X]$. Dieser ist ein Integritätsring nach Lemma 4.7. $\mathbb{Z}[X]$ enthält die Hauptideale

$$(X) = \left\{ \sum_{i=1}^{n} a_i X^i : n \in \mathbb{N}, a_i \in \mathbb{Z} \right\}$$

und

$$(2) = \left\{ \sum_{i=0}^{n} a_i X^i : n \in \mathbb{N}_0, a_i \in 2\mathbb{Z} \right\}.$$

Das Ideal $(2, X)$ erfüllt

$$(2, X) = \left\{ \sum_{i=0}^{n} a_i X^i : n \in \mathbb{N}, a_0 \in 2\mathbb{Z} \right\} \quad (\dot{\text{U}}\text{A}) .$$

Es ist allerdings kein Hauptideal. Sonst gäbe es nämlich ein $f \in \mathbb{Z}[X]$ mit $2 \in (f)$ und $X \in (f)$, d.h.

$$2 = fg \text{ und } X = fh$$

für gewisse $g, h \in \mathbb{Z}[X]$. Aus der ersten Gleichung folgt nach Bemerkung 4.6 aber $f \in \mathbb{Z}$ und $g \in \mathbb{Z}$. Da $(2, X) \neq \mathbb{Z}[X]$, kann f nicht ± 1 sein. Also folgt $f = \pm 2$ im Widerspruch zur zweiten Gleichung $X = fh$.

Somit ist $\mathbb{Z}[X]$ kein Hauptidealring.

Jetzt wollen wir - analog zu Faktorgruppen - auch Faktorringe definieren. Dazu untersuchen wir zunächst Ringhomomorphismen.

Lemma 4.13 *Es sei $\phi : R \to R'$ ein Ringhomomorphismus. Dann gilt*

i) Kern$(\phi) := \{a \in R : \phi(a) = 0\}$ ist ein Ideal in R.

ii) Bild $(\phi) = \phi(R)$ ist ein Unterring von R'.

iii) ϕ induziert einen Gruppenhomomorphismus $R^* \to R'^*$ zwischen den Einheitengruppen

Beweis :

i) Kern (ϕ) ist einfach der Kern der Abbildung ϕ, betrachtet als Gruppenhomomorphismus der zugehörigen abelschen Gruppen Also ist Kern(ϕ) eine Untergruppe von $(R, +)$. Für $r \in R$ und $a \in$ Kern (ϕ) gilt $\phi(ra) = \phi(r)\phi(a) = \phi(r)0 = 0$, also $ra \in$ Kern (ϕ). Somit ist Kern(ϕ) ein Ideal in R.

ii) Wie in i) folgt, dass Bild (ϕ) mit der eingeschränkten Addition eine Untergruppe von $(R', +)$ ist. Wegen $\phi(1) = 1$ liegt $1 \in$ Bild (ϕ). Aus $\phi(a)\phi(b) = \phi(ab) \in$ Bild(ϕ) folgt dann die Behauptung.

iii) Ist $a \in R^*$ mit $ab = 1$, so ist $\phi(a)\phi(b) = 1$, also $\phi(a) \in R'^*$. Somit vermittelt ϕ eine Abbildung $R^* \to R'^*$. Diese ist ein Gruppenhomomorphismus, da $\phi(ab) = \phi(a)\phi(b)$ gilt.

□

Man muss zwischen Idealen und Unterringen unterscheiden. Ein Ideal $\mathfrak{a} \subset R$ ist nur dann ein Unterring, wenn $1 \in \mathfrak{a}$ ist, woraus schon $\mathfrak{a} = R$ folgt.

Ein Unterring hingegen muss nicht abgeschlossen sein unter Multiplikation mit beliebigen Elementen in R. So wird auch das Bild eines Ringhomomorphismus im allgemeinen kein Ideal sein (ÜA: Wann ist es eins?).

Bemerkung 4.14 *Es sei K ein Körper und R ein Ring, $R \neq 0$. Dann ist jeder Ringhomomorphismus*

$$\phi : K \to R$$

injektiv. Insbesondere ist jeder Ringhomomorphismus zwischen Körpern injektiv. Einen Ringhomomorphismus zwischen Körpern nennt man auch ***Körperhomomorphismus***

Beweis : Kern(ϕ) ist ein Ideal in K, das wegen $\phi(1) = 1 \neq 0$ nicht ganz K ist. Da K nur die Ideale 0 und K enthält, folgt also Kern $(\phi) = 0$. □

Beispiel Zu jedem Ring R gibt es den Ringhomomorphismus

$$\phi : \mathbb{Z} \to R$$

mit

$$\phi(n) = n \cdot 1 = (\underbrace{1 + \cdots + 1}_{n\text{-mal}})$$

und

$$\phi(-n) = -\phi(n) \qquad \text{für } n \geq 0.$$

ϕ ist der einzige Ringhomomorphismus $\mathbb{Z} \to R$ (ÜA).

Definition 4.15 *Es sei R ein Ring und $\mathfrak{a} \subset R$ ein Ideal. Da $\mathfrak{a}$ insbesondere eine Untergruppe der abelschen Gruppe $(R, +)$ ist, existiert die **Faktorgruppe** $R/\mathfrak{a}$ der Nebenklassen $x + \mathfrak{a}$, $x \in R$, wobei die Addition definiert ist als*

$$(x + \mathfrak{a}) + (y + \mathfrak{a}) = (x + y) + \mathfrak{a}.$$

Zusätzlich definieren wir eine Multiplikation auf $R/\mathfrak{a}$ durch

$$(x + \mathfrak{a})(y + \mathfrak{a}) = (xy) + \mathfrak{a}.$$

Lemma 4.16 *Mit dieser Multiplikation wird $R/\mathfrak{a}$ zu einem Ring. Die Abbildung*

$$\begin{aligned} \pi : R &\to R/\mathfrak{a} \\ x &\mapsto x + \mathfrak{a} \end{aligned}$$

*ist ein Ringhomomorphismus mit Kern$(\pi) = \mathfrak{a}$. Der Ring $R/\mathfrak{a}$ heißt auch **Faktorring** und π heißt auch die **Projektionsabbildung** von R nach $R/\mathfrak{a}$.*

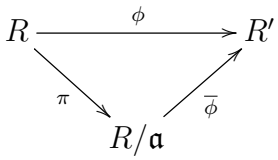
Beweis : Wir wissen aus Lemma 2.11, dass die Addition von Restklassen wohldefiniert ist. Um zu zeigen, dass auch die Multiplikation wohldefiniert ist, sei $x + \mathfrak{a} = x' + \mathfrak{a}$ und $y + \mathfrak{a} = y' + \mathfrak{a}$, d.h. $x' - x = a \in \mathfrak{a}$ und $y' - y = b \in \mathfrak{a}$.

Dann ist

$$x'y' = (x + a)(y + b) = xy + bx + ay + ab \in xy + \mathfrak{a}$$

Nun rechnet man leicht nach, dass $R/\mathfrak{a}$ ein Ring mit Einselement $1 + \mathfrak{a}$ ist (ÜA). Die kanonische Projektion π ist nach Lemma 2.11 ein surjektiver Gruppenhomomorphismus mit Kern $(\pi) = \mathfrak{a}$. Da $\pi(x)\pi(y) = (x + \mathfrak{a})(y + \mathfrak{a}) = xy + \mathfrak{a} = \pi(xy)$ und $\pi(1) = 1 + \mathfrak{a}$ gilt, ist π ein Ringhomomorphismus. □

Satz 4.17 (Homomorphiesatz für Ringe) *Es sei $\phi : R \to R'$ ein Ringhomomorphismus und $\mathfrak{a} \subset R$ ein Ideal mit $\mathfrak{a} \subset$ Kern (ϕ). Dann existiert genau ein Ringhomomorphismus $\overline{\phi} : R/\mathfrak{a} \to R'$, so dass das Diagramm*



kommutiert. Es gilt

$$\begin{aligned} \textit{Bild}\ (\overline{\phi}) &= \textit{Bild}\ (\phi) \\ \textit{Kern}\ (\overline{\phi}) &= \pi(\ \textit{Kern}\ \phi) \\ \textit{Kern}\ (\phi) &= \pi^{-1}(\ \textit{Kern}\ \overline{\phi}). \end{aligned}$$

Insbesondere ist $\overline{\phi}$ genau dann injektiv, wenn $\mathfrak{a} = \textit{Kern}\ \phi$ gilt.

Beweis : Nach Satz 2.12 existiert ein eindeutig bestimmter Gruppenhomomorphismus $\overline{\phi} : R/\mathfrak{a} \to R'$ mit $\phi = \overline{\phi} \circ \pi$, der die Aussagen über Kern $\overline{\phi}$ und Bild $\overline{\phi}$ erfüllt.

Es ist

$$\begin{aligned} \overline{\phi}((x+\mathfrak{a}) \cdot (y+\mathfrak{a})) &= \overline{\phi}(\pi(x) \cdot \pi(y)) \\ &= \overline{\phi}(\pi(xy)) \\ &= \phi(xy) \\ &= \phi(x) \cdot \phi(y) \\ &= \overline{\phi}(\pi(x)) \cdot \overline{\phi}(\pi(y)) \\ &= \overline{\phi}(x+\mathfrak{a}) \cdot \overline{\phi}(y+\mathfrak{a}), \end{aligned}$$

also ist $\overline{\phi}$ ein Ringhomomorphismus. Damit ist alles gezeigt. □

Korollar 4.18 *Ist $\phi : R \to R'$ ein surjektiver Ringhomomorphismus, so ist R' kanonisch isomorph zu $R/\textit{Kern}\phi$.*

Beweis : ϕ ist insbesondere ein surjektiver Gruppenhomomorphismus, also ist nach Korollar 2.13 der eindeutig bestimmte (kanonische) Gruppenhomomorphismus $\overline{\phi} : R/\text{Kern}\phi \to R'$ mit $\phi = \overline{\phi} \circ \pi$ ein Isomorphismus Nach Satz 4.17 ist $\overline{\phi}$ ein Ringhomomorphismus. Als bijektiver Ringhomomorphismus ist $\overline{\phi}$ nach dem folgenden Lemma aber auch ein Isomorphismus von Ringen. □

Lemma 4.19 *Es sei $\phi : R \to R'$ ein bijektiver Ringhomomorphismus. Dann ist ϕ ein **Isomorphismus von Ringen**, d.h. es existiert ein Ringhomomorphismus $\psi : R' \to R$ mit $\psi \circ \phi = id_R$ und $\phi \circ \psi = id_{R'}$.*

Beweis : Wir definieren $\psi(y)$ als das eindeutig bestimmte Urbild von ϕ in R. Dann rechnet man leicht nach, dass ψ mit der Addition und der Mutliplikation verträglich ist und $\phi(1) = 1$ sowie $\psi \circ \phi = \text{id}_R$ und $\phi \circ \psi = \text{id}_{R'}$ erfüllt. □

Die Isomorphiesätze 2.14 und 2.15 lassen sich auch für Ringe formulieren, wenn man überall Untergruppe und Normalteiler durch Ideal ersetzt. Diese Aussagen folgen dann aus den Sätzen 2.14 und 2.15, indem man nachprüft, dass alle Gruppenhomorphismen auch Ringhomomorphismen sind (ÜA).

Wir verallgemeinern jetzt den Begriff einer Primzahl in $\mathbb{Z}$.

Definition 4.20 *Es sei R ein Ring.*

*i) Ein Ideal $\mathfrak{p} \subset R$ heißt **prim** oder **Primideal**, wenn $\mathfrak{p} \neq R$ ist und wenn für alle $a, b \in R$ aus $ab \in \mathfrak{p}$ schon $a \in \mathfrak{p}$ oder $b \in \mathfrak{p}$ folgt.*

*ii) Ein Ideal $\mathfrak{m} \subset R$ heißt **maximal**, wenn $\mathfrak{m} \neq R$ ist und wenn für alle Ideale $\mathfrak{a} \subset R$ mit $\mathfrak{m} \subset \mathfrak{a}$ gilt: $\mathfrak{a} = \mathfrak{m}$ oder $\mathfrak{a} = R$.*

Beispiel: Das Ideal $m\mathbb{Z} \subset \mathbb{Z}$, $m \in \mathbb{N}_0$ ist genau dann ein Primideal, wenn $m = 0$ oder m eine Primzahl ist. $m\mathbb{Z} \subset \mathbb{Z}$ ist ein maximales Ideal genau dann, wenn m eine Primzahl ist. (ÜA)

Satz 4.21 *Es sei R ein Ring.*

i) Ein Ideal $\mathfrak{p} \subset R$ ist prim genau dann, wenn $R/\mathfrak{p}$ ein Integritätsring ist. Insbesondere ist 0 ein Primideal genau dann, wenn R ein Integritätsring ist.

ii) Ein Ideal $\mathfrak{m} \subset R$ ist genau dann ein maximales Ideal, wenn $R/\mathfrak{m}$ ein Körper ist. Insbesondere ist jedes maximale Ideal ein Primideal.

Beweis :

i) „$\Rightarrow$“: Ist $\mathfrak{p} \subset R$ ein Primideal, so ist $\mathfrak{p} \neq R$, d.h. $R/\mathfrak{p} \neq 0$. Gilt für $a, b \in R$, dass $\overline{a} \cdot \overline{b} = 0$ in $R/\mathfrak{p}$ ist, so ist $ab \in \mathfrak{p}$, also $a \in \mathfrak{p}$ oder $b \in \mathfrak{p}$, d.h. $\overline{a} = 0$ oder $\overline{b} = 0$.

„$\Leftarrow$“: Da $R/\mathfrak{p} \neq 0$ ist, ist $\mathfrak{p} \neq R$. Sei $ab \in \mathfrak{p}$, dann ist $\overline{a}\overline{b} = 0$. Da $R/\mathfrak{p}$ nullteilerfrei ist, folgt $\overline{a} = 0$ oder $\overline{b} = 0$, d.h. $a \in \mathfrak{p}$ oder $b \in \mathfrak{p}$.

ii) folgt aus Lemma 4.22.

□

Lemma 4.22 *i) Ein Ideal $\mathfrak{m} \subset R$ ist genau dann maximal, wenn das Nullideal in $R/\mathfrak{m}$ maximal ist.*

ii) Das Nullideal $0 \subset R$ ist maximal genau dann, wenn R ein Körper ist.

Beweis :

i) Für jedes Ideal $\mathfrak{a} \supset \mathfrak{m}$ ist das Bild $\pi(\mathfrak{a})$ unter der Restklassenabbildung $\pi : R \to R/\mathfrak{m}$ ein Ideal in $R/\mathfrak{m}$. Umgekehrt ist für jedes Ideal $\mathfrak{b} \subset R/\mathfrak{m}$ das Urbild $\pi^{-1}(\mathfrak{b})$ ein Ideal in R, das $\mathfrak{m}$ enthält. (ÜA). Diese Abbildungen vermitteln eine Bijektion (ÜA):

$$\begin{array}{rcl} \{ \text{ Ideale } \mathfrak{a} \subset R \text{ mit } \mathfrak{m} \subset \mathfrak{a} \} & \to & \{ \text{ Ideale in } R/\mathfrak{m} \} \\ \mathfrak{a} & \mapsto & \pi(\mathfrak{a}) \\ \pi^{-1}(\mathfrak{b}) & \leftarrow\!\shortmid & \mathfrak{b}. \end{array}$$

Daraus folgt sofort die Behauptung.

ii) Wir haben oben schon gesehen, dass ein Körper R nur die Ideale 0 und R enthält, die außerdem verschieden sind. Also ist 0 ein maximales Ideal in R. Sei umgekehrt R ein Ring, so dass 0 ein maximales Ideal in R ist. Wir müssen zeigen, dass $R \neq 0$ und dass jedes $a \in R \setminus \{0\}$ invertierbar ist. Ersteres folgt, da das maximale Ideal $0 \neq R$ ist. Sei ferner $a \in R \setminus \{0\}$. Dann ist

$$0 \subsetneqq (a) \subset R,$$

wegen der Maximalität von 0 also $(a) = R$. Daher ist $1 \in (a)$, d.h. es existiert ein $b \in R$ mit $ab = 1$.

□

5 Bilinearformen

Wir beginnen gleich mit der Definition:

Definition 5.1 *Es seien U, V, W K-Vektorräume. Eine Abbildung*

$$f : V \times W \to U$$

*heißt **bilinear**, falls folgende Bedingungen erfüllt sind:*

i) $f(v_1 + v_2, w) = f(v_1, w) + f(v_2, w)$ *für alle* $v_1, v_2 \in V$ *und* $w \in W$

ii) $f(v, w_1 + w_2) = f(v, w_1) + f(v, w_2)$ *für alle* $v \in V$ *und* $w_1, w_2 \in W$.

iii) $f(\alpha v, w) = \alpha f(v, w) = f(v, \alpha w)$ *für alle* $\alpha \in K, v \in V$ *und* $w \in W$.

Beispiel:

i) Die Produktabbildung

$$\begin{aligned} K^{l\times m} \times K^{m\times n} &\to K^{l\times n} \\ (A, B) &\mapsto AB \end{aligned}$$

ist bilinear.

ii) Es sei $A \in K^{m\times n}$ *eine beliebige* $(m \times n)$*-Matrix. Dann definiert* A *eine bilineare Abbildung*

$$\beta_A : K^m \times K^n \longrightarrow K$$

durch $\beta_A(v, w) = v^t Aw$.

Ist etwa $A = E_n \in K^{n\times n}$*, so ist die zugehörige bilineare Abbildung gerade*

$$\begin{aligned} \beta_{E_n} : K^n \times K^n &\longrightarrow K \\ \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) &\longmapsto x_1y_1 + x_2y_2 + \cdots + x_ny_n. \end{aligned}$$

Für $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \end{pmatrix} \in \mathbb{Q}^{2\times 3}$ *ergibt sich die bilineare Abbildung*

$$\begin{aligned} \beta_A : \mathbb{Q}^2 \times \mathbb{Q}^3 &\longrightarrow \mathbb{Q} \\ \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \right) &\mapsto (x_1, x_2) \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} &= (x_1 + x_2 \;\; 2x_1 \;\; 3x_1 + x_2) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ &= x_1y_1 + x_2y_1 + 2x_1y_2 + 3x_1y_3 + x_2y_3 \end{aligned}$$

Lemma 5.2 *Seien U, V, W K-Vektorräume und $f : V \times W \to U$ eine Abbildung. Dann ist f genau dann bilinear, falls für alle $v \in V$ die Abbildung*

$$\begin{aligned} f(v,-) : W &\to U \\ w &\mapsto f(v,w) \end{aligned}$$

linear ist und falls für alle $w \in W$ die Abbildung

$$\begin{aligned} f(-,w) : V &\to U \\ v &\mapsto f(v,w) \end{aligned}$$

linear ist.

Beweis : Das folgt sofort aus Definition 5.1. □

Ab jetzt betrachten wir immer den Spezialfall $U = K$. Eine bilineare Abbildung

$$f : V \times W \to K$$

lässt sich folgendermaßen mit einer Matrix beschreiben.

Definition 5.3 *Es sei $f : V \times W \to K$ bilinear mit endlich-dimensionalen Vektorräumen V und W. Ferner sei $B = (b_1, \ldots, b_m)$ eine Basis von V und $C = (c_1, \ldots, c_n)$ eine Basis von W. Dann heißt die $(m \times n)$-Matrix*

$$M_{f,B,C} = (f(b_i, c_j))_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$$

***Koordinatenmatrix** von f bezüglich B und C. Manchmal wird $M_{f,B,C}$ auch **Gram'sche Matrix** genannt. Der Eintrag von $M_{f,B,C}$ an der Stelle (i, j) ist also gerade $f(b_i, c_j)$.*

In [LA 1], Definition 5.28 haben wir die Koordinatenvektoren $\kappa_B(v) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}$ für

$$v = \beta_1 b_1 + \cdots + \beta_m b_m \in V \text{ kennengelernt.}$$

Lemma 5.4 *Es sei $f : V \times W \to K$ bilinear und $B = (b_1, \ldots, b_m)$ eine Basis von V sowie $C = (c_1, \ldots, c_n)$ eine Basis von W. Dann gilt für alle $v \in V$ und $w \in W$:*

$$f(v, w) = \kappa_B(v)^t M_{f,B,C} \kappa_C(w).$$

Beweis : Seien b_i und c_j Vektoren aus den Basen B und C. Dann gilt $\kappa_B(b_i) = e_i$ und $\kappa_C(c_j) = e_i$, wobei $e_i \in K^m$ und $e_j \in K^n$ die entsprechenden Einheitsvektoren sind. Also ist

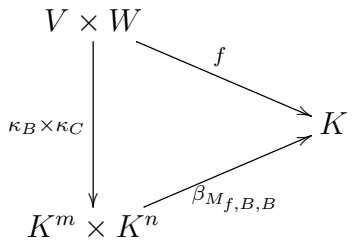
$$\kappa_B(b_i)^t M_{f,B,C} \kappa_C(c_j) = e_i^t M_{f,B,C} e_j$$

gerade der Eintrag von $M_{f,B,C}$ an der Stelle (i,j), d.h. gleich $f(b_i, c_j)$. Sind $v \in V$ und $w \in W$ beliebig, so schreiben wir $v = \sum_{i=1}^{m} \beta_i b_i$ und $w = \sum_{j=1}^{n} \gamma_j c_j$ mit $\beta_i, \gamma_j \in K$. Also ist $\kappa_B(v) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$ und $\kappa_C(w) = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}$. Da f bilinear ist, folgt

$$\begin{aligned}
f(v,w) &= f\left(\sum_{i=1}^{m} \beta_i b_i, \sum_{j=1}^{n} \gamma_j c_j\right) \\
&= \sum_{i=1}^{m} \beta_i f(b_i, \sum_{j=1}^{n} \gamma_j c_j) \\
&= \sum_{i=1}^{m} \beta_i \sum_{j=1}^{n} \gamma_j f(b_i, c_j) \\
&= \sum_{i=1}^{m} \beta_i \sum_{j=1}^{n} \gamma_j (e_i^t M_{f,B,C} e_j) \\
&= \sum_{i=1}^{m} \beta_i (e_i^t M_{f,B,C} \sum_{j=1}^{n} \gamma_j e_j) \\
&= \left(\sum_{i=1}^{m} \beta_i e_i^t\right) M_{f,B,C} \left(\sum_{j=1}^{n} \gamma_j e_j\right) \\
&= \kappa_B(v)^t M_{f,B,C} \kappa_c(w).
\end{aligned}$$

□

Dieses Resultat sagt, dass jede bilineare Abbildung $f : V \times W \to K$ nach Wahl von Basen B von V und C von W durch die Koordinatenmatrix beschrieben werden kann. Wir können Lemma 5.4 auch so ausdrücken: Das Diagramm



ist kommutativ. Hier ist $\beta_{M_{f,B,C}}$ die durch $M_{f,B,C}$ induzierte bilineare Abbildung aus dem obigen Beispiel.

Jetzt wollen wir untersuchen, wie sich die Koordinatenmatrix einer bilinearen Abbildung ändert, wenn man zu anderen Basen übergeht.

Proposition 5.5 *Es sei $f : V \times W \to K$ bilinear und dim$V = n$, dim $W = m$. Ferner seien Basen B und B' von V sowie C und C' von W gegeben.*

Wir bezeichnen (wie in [LA 1], Definition 5.31) mit $P_{B'}^{B}$ bzw. $P_{C'}^{C}$ die Übergangsmatrix von B nach B' bzw. von C nach C'. Dann gilt

$$M_{f,B',C'} = (P_B^{B'})^t M_{f,B,C} P_C^{C'}$$

Beweis : Es sei $B' = (b'_1, \ldots, b'_m)$ und $C' = (c'_1, \ldots, c'_n)$. Dann hat $P_B^{B'}$ die Spalten $(\kappa_B(b'_1), \ldots, \kappa_B(b'_m))$, die Matrix $P_C^{C'}$ hat die Spalten $(\kappa_C(c'_1), \ldots, \kappa_C(c'_n))$. Definitionsgemäß ist der Eintrag von $M_{f,B',C'}$ an der Stelle (i, j) gerade

$$\begin{aligned} f(b'_i, c'_j) &\overset{5.4}{=} \kappa_B(b'_i)^t \;\; M_{f,B,C} \;\; \kappa_C(c'_j) \\ &= (i\text{-te Zeile von } (P_B^{B'})^t) M_{f,B,C} (j\text{-te-Spalte von } P_C^{C'}) \\ &= \text{ Eintrag an der Stelle } (i,j) \text{ in der Matrix } (P_B^{B'})^t M_{f,B,C} P_C^{C'}. \end{aligned}$$

Also ist in der Tat $M_{f,B',C'} = (P_B^{B'})^t M_{f,B,C} P_C^{C'}$. □

Nach [LA 1], Lemma 6.13 ist jede invertierbare $(m \times m)$-Matrix die Übergangsmatrix von B zu einer anderen Basis von V und jede invertierbare $(n \times n)$-Matrix die Übergangsmatrix von C zu einer anderen Basis von W. Wir können Proposition 5.5 also auch so formulieren: Zwei Matrizen $M, N \in K^{m \times n}$ beschreiben genau dann dieselbe bilineare Abbildung (bezüglich verschiedener Basen), wenn Matrizen $P \in \mathrm{GL}(m, K)$ und $Q \in \mathrm{GL}(n, K)$ existieren mit

$$P^t M Q = N$$

Definition 5.6

*i) Eine bilineare Abbildung $f : V \times W \to K$ heißt **nicht ausgeartet in der ersten Variablen**, falls aus*

$$f(v, w) = 0 \textit{ für alle } w \in W$$

schon $v = 0$ folgt.

ii) $f : V \times W \to K$ *heißt* ***nicht ausgeartet in der zweiten Variablen****, falls aus*

$$f(v, w) = 0 \text{ für alle } v \in V$$

schon $w = 0$ *folgt.*

Beispiel: *Ist* $A = (1 \;\; 2) \in \mathbb{R}^{1\times 2}$*, so ist*

$$\begin{aligned} \beta_A : \mathbb{R} \times \mathbb{R}^2 &\to \mathbb{R} \\ (x, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}) &\mapsto xy_1 + 2xy_2 \end{aligned}$$

nicht ausgeartet in der ersten Variablen, aber ausgeartet in der zweiten Variablen, denn $\beta_A(x, \begin{pmatrix} -2 \\ 1 \end{pmatrix}) = 0$ *für alle* $x \in \mathbb{R}$*.*

Ab jetzt betrachten wir nur noch den Fall $V = W$. Eine bilineare Abbildung

$$f : V \times V \to K$$

nennt man auch **Bilinearform auf *V***.

Definition 5.7 *Sei* $f : V \times V \to K$ *eine Bilinearform.*

i) f *heißt* ***symmetrisch****, falls für alle* $v, w \in V$ *gilt* $f(v, w) = f(w, v)$*.*

ii) f *heißt* ***schiefsymmetrisch*** *oder alternierend, falls für alle* $v \in V$ *gilt* $f(v, v) = 0$*.*

Im Fall char$(K) \neq 2$ *lassen sich schiefsymmetrische Bilinearformen auch anders charakterisieren.*

Lemma 5.8 *Sei char*$K \neq 2$*. Eine Bilinearform* $f : V \times V \to K$ *ist genau dann schiefsymmetrisch, wenn für alle* $v, w \in V$

$$f(v, w) = -f(w, v)$$

gilt.

Beweis : Ist f schiefsymmetrisch, so gilt für alle $v, w \in V$:

$$0 = f(v + w, v + w) = f(v, v) + f(v, w) + f(w, v) + f(w, w) = f(v, w) + f(w, v),$$

also $f(v, w) = -f(w, v)$. (Dieses Argument funktioniert auch für char$(K) = 2$.)

Ist umgekehrt $f(v, w) = -f(w, v)$ für alle $v, w \in V$, so folgt $f(v, v) = -f(v, v)$, also $2f(v, v) = 0$. Wegen char$(K) \neq 2$ folgt $f(v, v) = 0$. □

Definition 5.9 *Sei $A = (a_{ij})_{i,j=1,\ldots,n} \in K^{n\times n}$.*

*i) A heißt **symmetrisch**, falls $a_{ij} = a_{ji}$ für alle i, j gilt.*

*ii) A heißt **schiefsymmetrisch**, falls $a_{ii} = 0$ für alle i und $a_{ij} = -a_{ji}$ für alle $i \neq j$ gilt.*

Proposition 5.10 *Sei $f : V \times V \to K$ bilinear und $B = (b_1, \ldots, b_n)$ eine beliebige Basis von V. Dann ist f genau dann symmetrisch bzw. schiefsymmetrisch, wenn die Koordinatenmatrix $M_{f,B,B}$ symmetrisch bzw. schiefsymmetrisch ist.*

Beweis : Ist f symmetrisch, so gilt

$$(M_{f,B,B})_{i,j} = f(b_i, b_j) = f(b_j, b_i) = (M_{f,B,B})_{j,i},$$

also ist $M_{f,B,B}$ symmetrisch. Ist umgekehrt $M_{f,B,B}$ symmetrisch, so folgt analog $f(b_i, b_j) = f(b_j, b_i)$ für alle i, j. Für $b = \sum_{i=1}^{n} \beta_i b_i$ und $c = \sum_{j=1}^{n} \gamma_j b_j$ in V folgt

$$\begin{aligned} f(b,c) &= \sum_{i=1}^{n}\sum_{j=1}^{n} \beta_i\gamma_j f(b_i, b_j) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \beta_i\gamma_j f(b_j, b_i) = f(c,b). \end{aligned}$$

Für schiefsymmetrische Formen bzw. Matrizen argumentiert man analog. □

Definition 5.11 *Es sei $f : V \times V \to K$ eine symmetrische Bilinearform auf V.*

*i) Dann heißen $v, w \in V$ **orthogonal** bezüglich f (wir schreiben $v \perp w$), falls $f(v, w) = 0$ ist.*

ii) Ist $x \subset V$ eine beliebige Teilmenge, so heißt die Menge

$$X^{\perp} = \{v \in V : f(x, v) = 0 \text{ für alle } x \in X\}$$

***orthogonales Komplement** von X bezüglich f. $X^{\perp}$ ist ein Unterraum von V.*

*iii) Ist $v \neq 0$ orthogonal zu sich selbst bezüglich f, d.h. gilt $f(v, v) = 0$, so heißt v **isotrop** bezüglich f.*

Beispiel: *Es sei*

$$\begin{array}{rcl} f : \mathbb{R}^4 \times \mathbb{R}^4 & \to & \mathbb{R} \\ \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \right) & \mapsto & x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4 \end{array}$$

die sogenannte Lorentzform, d.h. $f = \beta_A$ *für* $A = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & -1 \end{pmatrix}$. *Dann ist* f *nicht ausgeartet in beiden Variablen. Der Vektor* $v = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ *erfüllt* $f(v, v) = 0$, *d.h.* v *ist isotrop.*

Besitzt f einen isotropen Vektor v, so ist das orthogonale Komplement $\langle V \rangle^\perp$ kein Komplement im Sinne der Unterräume von V, da

$$v \in \langle v \rangle \cap \langle v \rangle^\perp \neq 0$$

gilt.

Proposition 5.12 *Es sei* $\mathrm{char} K \neq 2$. *Eine symmetrische Bilinearform* $f : V \times V \to K$ *mit* $f \neq 0$ *besitzt einen anisotropen Vektor, d.h. es existiert ein* $v \neq 0$ *in* V *mit* $f(v, v) \neq 0$.

Beweis : Ist $f \neq 0$, so gibt es $v, w \in V$ mit $f(v, w) \neq 0$. Falls $f(v, v) \neq 0$ oder $f(w, w) \neq 0$ sind, so folgt die Behauptung. Im Fall $f(v, v) = 0 = f(w, w)$ gilt

$$f(v + w, v + w) = f(v, v) + f(w, w) + 2f(v, w) = 2f(v, w) \neq 0,$$

da $\mathrm{char}(K) \neq 2$. □

Jetzt können wir einen Spezialfall angeben, in dem das orthogonale Komplement ein Komplement im üblichen Sinn ist.

Proposition 5.13 *Es sei* $f : V \times V \to K$ *eine symmetrische Bilinearform und* $v \in V$ *anisotrop bezüglich* f. *Dann gilt* $V = \langle v \rangle \oplus \langle v \rangle^\perp$.

Beweis : Wir zeigen zunächst $\langle v\rangle \cap \langle v\rangle^\perp = 0$. Ist für $a \in K$ der Vektor $av \in \langle v\rangle \cap \langle v\rangle^\perp$, so folgt $0 = f(v, av) = af(v, v)$. Da v anisotrop ist, gilt $f(v, v) \neq 0$, also $a = 0$, und somit $av = 0$.

Nun zeigen wir $\langle v\rangle + \langle v\rangle^\perp = V$. Sei $w \in V$. Wir setzen $a := \frac{f(v,w)}{f(v,v)}$ und schreiben $w = (w - av) + av$. Es ist $f(x - av, v) = f(w, v) - af(v, v) = 0$, woraus $w - av \in \langle v\rangle^\perp$ folgt. Also liegt $w = (w - av) + av \in \langle v\rangle^\perp + \langle v\rangle$. □

Offenbar ist eine symmetrische Bilinearform f auf V genau dann nicht ausgeartet in der ersten Variablen, wenn sie nicht-ausgeartet in der zweiten Variablen ist. Wir nennen sie dann einfach nicht ausgeartet. Definitionsgemäß ist f genau dann nicht ausgeartet, wenn $V^\perp = 0$ gilt. Wir können diese Eigenschaft wie folgt an der Koordinatenmatrix von f ablesen:

Satz 5.14 *Es sei $f : V \times V \to K$ eine symmetrische Bilinearform auf V und $B = (b_1, \ldots, b_n)$ eine Basis von V.*

i) Dann gilt

$$V^\perp = \{v \in V : M_{f,B,B}\kappa_B(v) = 0\}.$$

Im orthogonalen Komplement von V sind also genau die Vektoren in V, deren Koordinatenvektor im Kern der durch Multiplikation mit $M_{f,B,B}$ gegebenen linearen Abbildung liegen. Wir können $V^\perp$ daher durch Lösen eines linearen Gleichungssystems bestimmen.

ii) f ist genau dann nicht ausgeartet, wenn $M_{f,B,B}$ invertierbar ist.

Beweis :

i) Wir zeigen $V^\perp = \{v \in V : M_{f,B,B}\kappa_B(v) = 0\}$.

„$\supset$“: Ist $M_{f,B,B}\kappa_B(v) = 0$, so gilt nach Lemma 5.4 für alle $w \in V$:

$$f(w, v) = \kappa_B(w)^t M_{f,B,B}\kappa_B(v) = 0,$$

also ist $v \in V^\perp$.

„$\subset$“: Angenommen, $v \in V^\perp$. Dann gilt für alle $w \in V$

$$0 = f(w, v) \overset{5.4}{=} \kappa_B(w)^t M_{f,B,B}\kappa_B(w).$$

Der Vektor $x = M_{f,B,B}\kappa_B(v) = K^n$ hat also die Eigenschaft, dass für alle $y \in K^n$

$$y^t x = 0$$

gilt. Indem wir für y die Einheitsvektoren $e_1, \ldots, e_n$ einsetzen, sehen wir, dass alle Koeffizienten von $x = M_{f,B,B}\kappa(v)$ verschwinden. Also folgt $M_{f,B,B}\kappa_B(v) = 0$.

ii) f ist genau dann nicht ausgeartet, wenn $V^\perp = 0$ ist, nach i) also genau dann, wenn aus $M_{f,B,B}\kappa_B(v) = 0$ schon $v = 0$ folgt. Da $\kappa_B : V \to K^n$ ein Isomorphismus ist, ist das äquivalent dazu, dass der Kern der Abbildung

$$\begin{aligned} K^n &\to K^n \\ x &\mapsto M_{f,B,B}x \end{aligned}$$

gleich 0 ist, also zu der Tatsache, dass $M_{f,B,B}$ invertierbar ist.

□

Beispiel:

i) Die Lorentzform $f = \beta_A, A = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & -1 \end{pmatrix}$ ist nicht-ausgeartet.

ii) Die symmetrische Bilinearform

$$\begin{aligned} f : \mathbb{C}^2 \times \mathbb{C}^2 &\to \mathbb{C} \\ \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) &\mapsto x_1y_1 + ix_2y_1 + ix_1y_2 - x_2y_2 \end{aligned}$$

ist ausgeartet, da

$$\det \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix} = 0$$

gilt.

Definition 5.15 *Es sei V ein endlich-dimensionaler K-Vektorraum und $f : V \times V \to K$ eine symmetrische Bilinearform. Eine Basis $B = (b_1, \ldots, b_n)$ heißt* ***Orthogonalbasis von V*** *bezüglich f, falls für alle $i \neq j$ $b_i \perp b_j$ (d.h. $f(b_i, b_j) = 0$) gilt. Also ist B genau dann eine Orthogonalbasis bezüglich f, wenn $M_{f,B,B}$ eine Diagonalmatrix ist.*

Ist B eine Orthogonalbasis, so können wir leicht prüfen, ob f nicht ausgeartet ist. Nach Satz 5.14 ist f nämlich genau dann nicht ausgeartet, wenn die Diagonalmatrix $M_{f,B,B}$ invertierbar ist, d.h. wenn alle Diagonalelemente $f(b_i, b_i)(i = 1, \ldots, n)$ dieser Matrix ungleich Null sind.

Satz 5.16 *Es sei $\mathrm{char}K \neq 2$. Jede symmetrische Bilinearform $f : V \times V \to K$ auf einem endlich-dimensionalen Vektorraum V besitzt eine Orthogonalbasis.*

Beweis : Ist $f = 0$, so ist jede beliebige Basis von V eine Orthogonalbasis. Wir können also $f \neq 0$ annehmen. Wir zeigen die Behauptung mit Induktion nach $\dim V$. Für $\dim V = 1$ ist jede Basis von V eine Orthogonalbasis. Also nehmen wir an, die Behauptung gilt für alle symmetrischen Bilinearformen auf Vektorräumen der Dimension $< \dim V$. Nach Proposition 5.12 besitzt f einen anisotropen Vektor v, und nach Proposition 5.13 ist $V = \langle v \rangle \oplus \langle v \rangle^{\perp}$. Nach Induktionsvoraussetzung besitzt der Vektorraum $\langle v \rangle^{\perp}$ zusammen mit der Einschränkung von f auf $\langle v \rangle^{\perp} \times \langle v \rangle^{\perp}$ eine Orthogonalbasis $(b_2, \ldots, b_n)$. Da $f(v, b_i) = 0$ für alle $i = 2, \ldots, n$ gilt, ist $(v, b_2, \ldots, b_n)$ eine Orthogonalbasis von V bezüglich f. □

Korollar 5.17 *Wir nehmen an, dass $\mathrm{char}K \neq 2$. Sei $A \in K^{n \times n}$ eine symmetrische Matrix. Dann existiert ein $P \in GL_n(K)$, so dass $P^t AP$ eine Diagonalmatrix ist.*

Beweis : Ist A symmetrisch, so ist $\beta_A : K^n \times K^n \to K$ eine symmetrische Bilinearform. Nach Satz 5.16 gibt es eine Orthogonalbasis B von K^n bezüglich β_A, d.h. $M_{f,B,B}$ ist eine Diagonalmatrix. Ist C die kanonische Basis von K^n, so ist $M_{f,C,C} = A$. Für die Übergangsmatrix $P = P_C^B \in \mathrm{GL}_n(K)$ gilt daher nach Proposition 5.5 $M_{f,B,B} = P^t AP$, woraus die Behauptung folgt. □

6 Symmetrische Bilinearformen über $\mathbb{R}$

In diesem Kapitel werden wir symmetrische Bilinearformen auf endlich-dimensionalen $\mathbb{R}$-Vektorräumen untersuchen. Wir beginnen gleich mit einem Satz, der die Existenz besonders einfacher Orthogonalbasen garantiert.

Satz 6.1 (Trägheitssatz von Sylvester) *Es sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und $f : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Dann existiert*

eine Basis B von V, für die gilt:

$$M_{f,B,B} = \begin{pmatrix} E_r & & \\ & -E_s & \\ & & 0 \end{pmatrix} = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & -1 & & \\ & & & & & & 0 & \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\&\ddots\\&&1\end{matrix}}\right\} r \\ \left.\vphantom{\begin{matrix}-1\\&\ddots\\&&-1\end{matrix}}\right\} s \\ \left.\vphantom{\begin{matrix}0\\&\ddots\\&&0\end{matrix}}\right\} n-r-s \end{matrix}$$

*Hier sind $r, s \geq 0$ mit $r + s \leq n$. Das Zahlenpaar (r, s) ist durch f eindeutig bestimmt, es heißt **Signatur** von f.*

Beweis : Nach Satz 5.16 existiert eine Orthogonalbasis $c_1, \ldots, c_n$ von V bezüglich f. Wir nummerieren diese so, dass

$$\begin{aligned} f(c_i, c_i) > 0 \qquad & \text{für } 1 \leq i \leq r \\ f(c_i, c_i) < 0 \qquad & \text{für } r+1 \leq i \leq r+s \\ \text{und } f(c_i, c_i) = 0 \qquad & \text{für } r+s+1 \leq i \leq n \end{aligned}$$

gilt. Für $i = 1, \ldots, r$ setzen wir $\alpha_i = \sqrt{f(c_i, c_i)}$ und definieren $b_i = \frac{1}{\alpha_i} c_i$.

Dann ist $f(b_i, b_i) = \frac{1}{\alpha_i^2} f(c_i, c_i) = 1$. Für $i = r+1, \ldots, r+s$ setzen wir $\alpha_i := \sqrt{-f(c_i, c_i)}$ und definieren $b_i = \frac{1}{\alpha_i} c_i$. Dann ist $f(b_i, b_i) = \frac{1}{\alpha_i^2} f(c_i, c_i) = -1$.

Für $i = r + s + 1, \ldots, n$ setzen wir $b_i = c_i$.

Dann ist $B = (b_1, \ldots, b_n)$ eine Orthogonalbasis von V bezüglich f, für die

$$M_{f,B,B} = \begin{pmatrix} E_r & & \\ & -E_s & \\ & & 0 \end{pmatrix}$$

gilt.

Jetzt müssen wir noch zeigen, dass (r, s) durch f eindeutig bestimmt sind. Wir nehmen an, B' ist ebenfalls eine Orthogonalbasis mit

$$M_{f,B',B'} = \begin{pmatrix} E'_r & & \\ & -E'_s & \\ & & 0 \end{pmatrix}.$$

Dann müssen wir $r = r'$ und $s = s'$ zeigen. Offenbar ist $r + s = \text{rang}(M_{f,B,B})$ und $r' + s' = \text{rang}(M_{f,B',B'})$. Ist $P = P_{B'}^{B} \in \text{GL}_n(\mathbb{R})$ die Übergangsmatrix von B nach B', so gilt nach Proposition 5.5 $P^t M_{f,B',B'} P = M_{f,B,B}$. Da P invertierbar ist, folgt $\text{rang}(M_{f,B',B'}) = \text{rang}(M_{f,B,B})$, also $r + s = r' + s'$.

Als nächsten Schritt zeigen wir, dass die $r + (n - r')$ Vektoren $b_1, \ldots, b_r, b'_{r'+1}, \ldots, b'_n$ linear unabhängig sind. Haben wir das gezeigt, so folgt

$$r + (n - r') \leq \dim V = n,$$

also $r \leq r'$.

Es sei also $\sum_{i=1}^{r} \alpha_i b_i + \sum_{i=r'+1}^{n} \beta_i b'_i = 0$ mit $\alpha_i, \beta_i \in K$ eine Linearkombination der Null. Wir setzen $v = \sum_{i=1}^{r} \alpha_i b_i = -\sum_{i=r'+1}^{n} \beta_i b'_i$. Dann ist

$$f(v,v) = f\left(\sum_{i=1}^{r} \alpha_i b_i, \sum_{i=1}^{r} \alpha_i b_i\right) = \sum_{i=1}^{r} \alpha_i f(b_i, \sum_{i=1}^{r} \alpha_i b_i) = \sum_{i=1}^{r} \alpha_i^2 f(b_i, b_i) = \sum_{i=1}^{r} \alpha_i^2 \geq 0,$$

da $B = (b_1, \ldots, b_n)$ eine Orthogonalbasis mit $f(b_i, b_i) > 0$ für $1 \leq i \leq r$ ist.

Analog rechnen wir

$$f(v,v) = f(-\sum_{i=r'+1}^{n} \beta_i b'_i, -\sum_{i=r'+1}^{n} \beta_i b'_i) = \sum_{i=r'+1}^{n} \beta_i^2 f(b'_i, b'_i) = -\sum_{i=r'+1}^{r'+s'} \beta_i^2 \leq 0,$$

da $B' = (b'_1, \ldots, b'_n)$ eine Orthogonalbasis mit $f(b'_i, b'_i) < 0$ für $r' + 1 \leq i \leq r' + s'$ und $f(b'_i, b'_i) = 0$ für $r' + s' + 1 \leq i \leq n$.

Also folgt $\sum_{i=1}^{r'} \alpha_i^2 = -\sum_{i=r'+1}^{r'+s'} \beta_i^2 = 0$, d.h. $\alpha_1 = \cdots = \alpha_r = 0$ und $\beta_{r'+1} = \cdots = \beta_{r'+s'} = 0$. Dann ist $\sum_{i=r'+s'+1}^{n} \beta_i b'_i = 0$, also auch $\beta_{r'+s'+1} = \cdots = \beta_n = 0$. Somit sind $b_1, \ldots, b_r, b'_{r'+1}, \ldots, b'_n$ in der Tat linear unabhängig und es folgt $r \leq r'$. Dasselbe Argument mit vertauschten Rollen liefert $r' \leq r$, also $r = r'$. Da $r + s = r' + s'$ gilt, folgt auch $s = s'$. Die Signatur (r, s) ist also tatsächlich eindeutig durch f bestimmt. □

Wir lernen jetzt eine weitere wichtige Eigenschaft von symmetrischen Bilinearformen über $\mathbb{R}$ kennen.

Definition 6.2

*i) Eine symmetrische Bilinearform $f : V \times V \to \mathbb{R}$ auf einem endlich-dimensionalen $\mathbb{R}$-Vektorraum V heißt **positiv definit**, wenn für alle $v \neq 0$ aus V die Ungleichung $f(v, v) > 0$ gilt. Eine positiv definite symmetrische Bilinearform auf einem endlich-dimensionalen $\mathbb{R}$-Vektorraum V nennt man auch Skalarprodukt auf V.*

*ii) Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ heißt **positiv definit**, falls $\beta_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ positiv definit ist, d.h. falls für alle $x \neq 0$ aus $\mathbb{R}^n$ die Ungleichung $x^t Ax > 0$ gilt.*

***Beispiel:** Eine Diagonalmatrix $A = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \in \mathbb{R}^{n\times n}$ ist genau dann positiv definit, wenn alle $d_i > 0$ sind. Für $A = E_n$ heißt die positiv definite symmetrische Bilinearform*

$$\begin{array}{rcl} \langle , \rangle : \mathbb{R}^n \times \mathbb{R}^n & \to & \mathbb{R}^n \\ \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) & \mapsto & x_1 y_1 + \cdots + x_n y_n \end{array}$$

***kanonisches Skalarprodukt** auf $\mathbb{R}^n$.*

Lemma 6.3 *Eine symmetrische Bilinearform f auf dem endlich-dimensionalen $\mathbb{R}$-Vektorraum V ist genau dann positiv definit, wenn für eine Basis B von V die Koordinatenmatrix $M_{f,B,B}$ positiv definit ist. Dann ist für jede beliebige Basis B von V die Koordinatenmatrix $M_{f,B,B}$ positiv definit.*

Beweis : Es sei B eine beliebige Basis von V. Es ist $f(v, v) = \kappa_B(v)^t M_{f,B,B} \kappa_B(v)$ nach Lemma 5.4. Da $\kappa_B : V \to \mathbb{R}^n$ ein Isomorphismus ist, ist f genau positiv definit, wenn $M_{f,B,B}$ positiv definit ist. □

Definition 6.4 *Es sei $f : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform auf einem endlich-dimensionalen $\mathbb{R}$-Vektorraum. Eine Basis $B = (b_1, \ldots, b_n)$ von V heißt **Orthonormalbasis** von V bezüglich f, falls*

$$f(b_i, b_j) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

gilt.

Eine Basis B ist also genau dann eine Orthonormalbasis von V bezüglich f, wenn $M_{f,B,B} = E_n$ gilt. In diesem Fall ist die Signatur von f gleich $(n, 0)$. Existiert eine Orthonormalbasis von V bezüglich f, dann ist f positiv definit. Wir zeigen nun, dass auch die Umkehrung gilt.

Satz 6.5 *Ist f eine positiv definite symmetrische Bilinearform auf dem endlichdimensionalen $\mathbb{R}$-Vektorraum V, so gibt es eine Orthonormalbasis von V bezüglich f.*

Beweis : (mit Hilfe des sogenannten **Gram-Schmidt-Orthonormalisierungsverfahrens**) Wir wählen eine beliebige Basis $C = (c_1, \ldots, c_n)$ von V und konstruieren daraus eine Orthonormalbasis. Da f positiv definit ist, ist $f(c_1, c_1) > 0$, also gilt für $b_1 = \frac{1}{\sqrt{f(c_1,c_1)}} c_1$ die Gleichung $f(b_1, b_1) = 1$.

Als nächsten Schritt suchen wir einen Vektor $b_2' \in \langle b_1, c_2\rangle$ mit $b_1 \perp b_2'$. Der Ansatz $b_2' = \alpha_1 b_1 + c_2$ liefert $f(b_1, b_2') = \alpha_1 f(b_1, b_1) + f(b_1, c_2) = \alpha_1 + f(b_1, c_2)$. Also gilt für

$$\alpha_1 = -f(b_1, c_2) \in \mathbb{R},$$

dass der Vektor $b_2' = \alpha_1 b_1 + c_2$ orthogonal zu b_1 ist. Offenbar gilt $\langle b_1, b_2'\rangle = \langle b_1, c_2\rangle = \langle c_1, c_2\rangle$. Wir setzen wieder $b_2 = \frac{1}{\sqrt{f(b_2', b_2')}} b_2$ und erhalten so ein Paar(b_1, b_2) von Vektoren in V mit $\langle b_1, b_2\rangle = \langle c_1, c_2\rangle$ und $f(b_i, b_j) = \delta_{ij}$ für $i = 1, 2$. So fahren wir fort. Sind Vektoren $(b_1, \ldots, b_k)$ mit $\langle b_1, \ldots, b_k\rangle = \langle c_1, \ldots, c_k\rangle$ und $f(b_i, b_j) = \delta_{ij}$ gefunden, so sei

$$\alpha_1 = -f(b_1, c_{k+1}), \ldots, \alpha_k = -f(b_k, c_{k+1}).$$

Der Vektor $b_{k+1}' = \alpha_1 b_1 + \cdots + \alpha_k b_k + c_{k+1}$ erfüllt dann

$$\begin{aligned} f(b_i, b_{k+1}) &= f(b_i, \sum_{j=1}^{k} \alpha_j b_j) + f(b_i, c_{k+1}) \\ &= \alpha_i + f(b_i, c_{k+1}) = 0. \end{aligned}$$

Mit $b_{k+1} = \frac{1}{\sqrt{f(b_{k+1}', b_{k+1}')}} b_{k+1}'$ erhalten wir ein System $(b_1, \ldots, b_{k+1})$ von Vektoren, so dass $\langle b_1, \ldots, b_{k+1}\rangle = \langle c_1, \ldots, c_{k+1}\rangle)$ und $f(b_i, b_j) = \delta_{ij}$ gilt. Sind alle c_i verbraucht, so haben wir eine Orthonormalbasis $B = (b_1, \ldots, b_n)$ gefunden. □

Wir hätten Satz 6.5 auch ganz einfach mit dem Sylvesterschen Trägheitsgesetz beweisen können. Der hier angegebene Beweis liefert allerdings ein effektives Verfahren, um eine Orthonormalbasis zu konstruieren. Dieses Resultat wollen wir noch für Matrizen umformulieren.

Satz 6.6 *Sei $A \in \mathbb{R}^{n\times n}$. Dann sind äquivalent.*

i) Es gibt eine Basis B von $\mathbb{R}^n$ mit $A = M_{\langle,\rangle,B,B}$, wobei $\langle,\rangle$ das kanonische Skalarprodukt auf dem $\mathbb{R}^n$ ist.

ii) Es gibt ein $P \in GL(n,\mathbb{R})$ mit $A = P^tP$.

iii) A ist symmetrisch und positiv definit.

Beweis : i) $\Rightarrow$ iii) ist klar mit Proposition 5.10 und Lemma 6.3

iii) $\Rightarrow$ ii) Wir wenden Satz 6.5 auf die positiv definite symmetrische Bilinearform $\beta_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ an. Es gibt also eine Orthonormalbasis B von $\mathbb{R}^n$ bezüglich β_A, d.h. $M_{\beta_A,B,B} = E_n$. Es sei C die kanonische Basis des $\mathbb{R}^n$ und $P = P_B^C \in \mathrm{GL}(n,\mathbb{R})$. Dann gilt nach Proposition 5.5: $A = M_{\beta_A,C,C} = P^tE_nP = P^tP$.

ii) $\Rightarrow$ i): Es sei $C = (e_1,\ldots,e_n)$ die kanonische Basis des $\mathbb{R}^n$. Nach Lemma 6.13 gibt es eine Basis B von $\mathbb{R}^n$ mit $P_C^B = P$. Dann ist

$$\begin{aligned} M_{\langle,\rangle,B,B} &\overset{5.5}{=} P^tM_{\langle,\rangle,C,C}P \\ &= P^tP = A. \end{aligned}$$

□

Wir wollen jetzt noch ein Kriterium zeigen, mit dem man feststellen kann, ob eine Matrix positiv definit ist.

Definition 6.7 *Sei $A = (a_{ij})_{i,j=1,\ldots,n} \in \mathbb{R}^{n\times n}$. Es sei*

$$A_k = \begin{pmatrix} a_{11} & \ldots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \ldots & a_{kk} \end{pmatrix} \in \mathbb{R}^{k\times k}$$

*die obere linke $k \times k$-Untermatrix von A. Die Werte $\det A_1, \det A_2, \ldots, \det A_n = \det A$ heißen **Hauptminoren von A**.*

***Beispiel:** Für $A = \begin{pmatrix} -1 & 1 \\ 1 & -2 \end{pmatrix}$ sind $\det A_1 = -1, \det A_2 = \det A = 1$ die Hauptminoren.*

Satz 6.8 *Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann positiv definit, wenn alle Hauptminoren positiv sind, d.h. wenn $\det A_k > 0$ für $k = 1,\ldots,n$ gilt.*

Beweis : Angenommen, A ist positiv definit. Dann gibt es nach Satz 6.6 ein $P \in \mathrm{GL}(n, \mathbb{R})$ mit $A = P^t P$. Daher ist $\det A = \det(P^t P) = (\det P)^2 > 0$.

Für alle $k \geq 1$ sei V_k der von den ersten k Einheitsvektoren $e_1, \ldots, e_k$ aufgespannte Unterraum von $\mathbb{R}^n$. Die Matrix A definiert eine positiv definite symmetrische Bilinearform β_A auf $\mathbb{R}^n$. Auch ihre Einschränkung $\beta_{A|_{V_k}}$ auf V_k ist positiv definit. Die Koordinatenmatrix von $\beta_{A|_{V_k}}$ bezüglich $e_1, \ldots, e_k$ ist gerade A_k. Also gilt $\det A_k > 0$.

Wir nehmen umgekehrt an, alle Hauptminoren von A seien positiv und zeigen mit Induktion nach n, dass A dann positiv definit ist. Für $n = 1$ ist das klar. Also können wir annehmen, dass A_{n-1}, deren Hauptminoren gerade $A_1, \ldots, A_{n-1}$ sind, positiv definit ist. Nach Satz 6.6 gibt es dann ein $Q \in \mathrm{GL}(n-1, \mathbb{R})$mit $Q\ A_{n-1} Q^t = E_{n-1}$. Es sei $\tilde{P} = \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}(n, \mathbb{R})$. Es ist $\tilde{P}\ A \tilde{P}^t = \begin{pmatrix} E_{n-1} & * \\ * & * \end{pmatrix}$. Durch $(n-1)$ elementare Zeilenumformungen, also Multiplikation mit Elementarmatrizen $L_{n-1}, \ldots, L_1$ von links, kann man die ersten $(n-1)$-Einträge der letzten Zeile eliminieren. Da $\tilde{P} A \tilde{P}^t$ symmetrisch ist, kann man mit den analogen Spaltenoperationen, also durch Multiplikation mit $L_1^t, \ldots, L_{n-1}^t$ von rechts, die ersten $(n-1)$ Einträge der letzte Spalte eliminieren. Für $P = L_{n-1} \cdots L_1 \tilde{P}$ gilt also $PAP^t = \begin{pmatrix} E_{n-1} & 0 \\ 0 & c \end{pmatrix}$ für ein $c \in \mathbb{R}$. Aus $\det A > 0$ folgt $c > 0$. Somit ist PAP^t positiv definit, woraus mit Lemma 6.3 auch A positiv definit folgt. □

Definition 6.9 *Ein endlich-dimensionaler reeller Vektorraum V mit einer positiv definiten, symmetrischen Bilinearform $\langle , \rangle : V \times V \to \mathbb{R}$ heißt auch* ***euklidischer Vektorraum****.*

Ist $(V, \langle , \rangle)$ ein euklidischer Vektorraum, so können wir für alle $v \in V$ eine Länge $||v|| = \sqrt{\langle v, v \rangle}$ definieren. Es gilt dann

Lemma 6.10

i) $||v|| = 0$ *genau dann, wenn* $v = 0$*.*

ii) Für alle $v, w \in V$ *ist* $||v + w|| \leq ||v|| + ||w||$ ***(Dreiecksungleichung)***

iii) Für alle $v, w \in V$ *ist* $|\langle v, w \rangle| \leq ||v||\ ||w||$ ***(Cauchy-Schwarz'sche Ungleichung)****.*

Beweis :

i) $||v|| = 0 \Leftrightarrow \langle v, v \rangle = 0 \Leftrightarrow v = 0$, da $\langle , \rangle$ positiv definit ist.

ii) folgt aus iii), denn es gilt

$$\begin{aligned} ||v+w||^2 &= \langle v+w, v+w\rangle = \langle v,v\rangle + 2\langle v,w\rangle + \langle w,w\rangle \\ &= ||v||^2 + 2\langle v,w\rangle + ||w||^2 \\ &\leq ||v||^2 + 2|\langle v,w\rangle| + ||w||^2 \\ &\overset{\text{(iii)}}{\leq} ||v||^2 + 2||v||\,||w|| + ||w||^2 \leq (||v|| + ||w||)^2. \end{aligned}$$

iii) Es sei $W = \langle v, w\rangle \subset V$. Die Einschränkung von $\langle , \rangle$ auf W ist ebenfalls eine positiv definite, symmetrische Bilinearform (also ein Skalarprodukt) $\langle , \rangle_W$ auf W.

Ist $\dim W = 1$, so können wir $v = \lambda w$ für ein $\lambda \in K$ annehmen. Dann gilt

$$|\langle v, w\rangle| = |\lambda|\,||w||^2 = ||v||\,||w||.$$

Ist $\dim W = 2$, so ist (v, w) eine Basis von W. Die Koordinatenmatrix der Bilinearform $\langle , \rangle$ bezüglich dieser Basis ist

$$A = \begin{pmatrix} \langle v,v\rangle & \langle v,w\rangle \\ \langle v,w\rangle & \langle w,w\rangle \end{pmatrix}.$$

Nach Lemma 6.3 ist A positiv definit, also ist nach Satz 6.8

$$\det A = \langle v,v\rangle\langle w,w\rangle - \langle v,w\rangle^2 > 0.$$

Es folgt also $||v||^2\,||w||^2 > \langle v,w\rangle^2$, und daher $||v||\,||w|| > |\langle v,w\rangle|$.

□

Ist $(V, \langle , \rangle)$ ein euklidischer Vektorraum und $B = (b_1, \ldots, b_n)$ eine Orthonormalbasis von V bezüglich $\langle , \rangle$, so ist $\kappa_B : V \to \mathbb{R}^n$ ein Isomorphismus von $\mathbb{R}$-Vektorräumen, der $\langle , \rangle$ in das kanonische Skalarprodukt des $\mathbb{R}^n$ überführt. Es gilt nämlich für $v = \sum_{i=1}^{n} \alpha_i b_i$ und $w = \sum_{i=1}^{n} \beta_i b_i$:

$$\begin{aligned} \langle v, w\rangle &= \left\langle \sum_{i=1}^{n} \alpha_i b_i, \sum_{i=1}^{n} \beta_i b_i \right\rangle \\ &= \sum_{i=1}^{n} \alpha_i\beta_i = (\alpha_i, \ldots, \alpha_n) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \\ &= \kappa_B(v)^t \kappa_B(w). \end{aligned}$$

Auf einem euklidischen Vektorraum $(V, \langle , \rangle)$ kann man nicht-orientierte Winkel definieren.

Definition 6.11 *Für $v, w \in V \setminus \{0\}$ sei $\theta \in [-\pi, \pi[$ eine Zahl mit*

$$\cos\theta = \frac{\langle v, w\rangle}{||v||\ ||w||}.$$

*Dadurch ist θ bis auf das Vorzeichen ± 1 eindeutig festgelegt. Nach der Cauchy-Schwarz'schen Ungleichung ist nämlich $-1 \leq \frac{\langle v,w\rangle}{||v||\ ||w||} \leq 1$. θ heißt **(nicht-orientierter) Winkel** zwischen v und w. Aus dieser Definition folgt sofort $||v+w||^2 = ||v||^2 + ||w||^2 + 2||v||\ ||w|| \cos\theta$.*

In einem euklidischen Vektorraum $(V, \langle , \rangle)$ ist für jeden Unterraum $W \subset V$ das orthogonale Komplement $W^\perp$ ein Komplement im Vektorraumsinne, d.h. es gilt

$$W \oplus W^\perp = V$$

(siehe Übungsaufgaben). Jedes $v \in V$ hat also eine eindeutige Darstellung $v = w + w'$ mit $w \in W, w' \in W^\perp$. Die Abbildung

$$\begin{aligned} p : V &\to W \\ v &\mapsto w \end{aligned}$$

*ist linear. Es ist $p(v) = w$ genau dann, wenn $v - w \in W^\perp$ liegt. p heißt **orthogonale Projektion von V auf W**.*

Lemma 6.12 *Es sei $W \subset V$ ein Unterraum des euklidischen Vektorraums $(V, \langle , \rangle)$ und $(w_1, \ldots, w_r)$ eine Orthonormalbasis von W. Dann gilt für die orthogonale Projektion $p : V \to W$: Für alle $v \in V$ ist $p(v) = \langle v, w_1\rangle w_1 + \cdots + \langle v, w_r\rangle w_r$.*

Beweis : Wir setzen $\tilde{w} = \sum_{i=1}^{r} \langle v, w_i\rangle w_i$. Für alle $j = 1, \ldots, r$ gilt

$$\langle v - \tilde{w}, w_j\rangle = \langle v - \sum_{i=1}^{r} \langle v, w_i\rangle w_i, w_j\rangle = \langle v, w_j\rangle - \langle v, w_j\rangle = 0.$$

Also ist $v - \tilde{w} \in W^\perp$, woraus $p(v) = \tilde{w}$ folgt. □

Es sei $(V, \langle , \rangle)$ ein euklidischer Raum.

Definition 6.13

*i) Eine lineare Abbildung $f : V \to V$ heißt **Isometrie** (bezüglich $\langle , \rangle$), falls für alle $v, w \in V$*

$$\langle f(v), f(w)) \rangle = \langle v, w \rangle$$

gilt.

*ii) Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt **orthogonal**, falls $A^t A = E_n$ gilt.*

Offenbar ist jede Isometrie injektiv, also ein Isomorphismus, falls V endlich-dimensional ist. Eine Matrix A ist definitionsgemäß genau dann orthogonal, wenn $\lambda_A : \mathbb{R}^n \to \mathbb{R}^n$ eine Isometrie bezüglich des kanonischen Skalarproduktes ist. Jede orthogonale Matrix A ist invertierbar mit $A^{-1} = A^t$. Also gilt auch $AA^t = E_n$. Ist V endlich-dimensional, so ist

$$O(V, \langle , \rangle) = \{ f : V \to V : f \text{ Isometrie } \}$$

eine Gruppe bezüglich der Komposition von Abbildungen.

Ebenso bildet die Menge

$$\begin{aligned} O(n, \mathbb{R}) &= \{ A \in \mathbb{R}^{n \times n} : A \text{ orthogonal } \} \\ &= \{ A \in \mathbb{R}^{n \times n} : A^t A = E_n \} \end{aligned}$$

*eine Untergruppe von $GL(n, \mathbb{R})$ bezüglich der Multiplikation von Matrizen, die sogenannte **orthogonale Gruppe**. Jedes $A \in O(n, \mathbb{R})$ erfüllt wegen $A^t A = E_n$ die Gleichung $(\det A)^2 = 1$, also ist $\det A \in \{\pm 1\}$. Die Menge $SO(n, \mathbb{R}) = \{A \in O(n, \mathbb{R}) : \det A = 1\}$ ist eine Untergruppe von $O(n, \mathbb{R})$, sie heißt **spezielle orthogonale Gruppe**.*

7 Hermitesche Formen über $\mathbb{C}$

Ist der Grundkörper $K = \mathbb{C}$, so ist es sinnvoll, statt mit symmetrischen Bilinearformen mit den sogenannnnten hermite'schen Formen zu arbeiten.

Definition 7.1 *Eine Abbildung f, die nur die Eigenschaften i) und ii) besitzt, wird auch **Sesquilinarform** genannt (lat. sesqui bedeutet soviel wie anderthalbfach).*

Beispiel: *$\langle , \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}, \langle x, y \rangle = \overline{x}^t y$ ist eine hermite'sche Form.*

*Sei V ein $\mathbb{C}$-Vektorraum. Eine **hermite'sche Form** auf V ist eine Abbildung*

$$f : V \times V \to \mathbb{C}$$

mit folgenden Eigenschaften:

i) f ist linear in der zweiten Variablen:

$$\begin{aligned} f(v, w_1 + w_2) &= f(v, w_1) + f(v, w_2) \text{ und} \\ f(v, \alpha w_1) &= \alpha f(v, w_1) \end{aligned}$$

für alle $v, w_1, w_2 \in V$ und $\alpha \in \mathbb{C}$.

ii) f ist konjugiert-linear in der ersten Variablen:

$$\begin{aligned} f(v_1 + v_2, w) &= f(v_1, w) + f(v_2, w) \text{ und} \\ f(\alpha v_1, w) &= \overline{\alpha} f(v_1, w) \end{aligned}$$

für alle $v_1, v_2, w \in V$ und $\alpha \in \mathbb{C}$.

iii) f ist konjugiert symmetrisch, d.h.

$$f(v, w) = \overline{f(w, v)} \text{ für alle } v, w \in V.$$

Ist f eine hermite'sche Form auf V, so gilt für alle $v \in V$ nach iii) aus Definition 7.1:

$$f(v, v) = \overline{f(v, v)}, \text{ d.h. } f(v, v) \in \mathbb{R}.$$

Definition 7.2 *Eine hermite'sche Form f auf dem komplexen Vektorraum V heißt **positiv definit**, falls für alle $v \neq 0$ aus V*

$$f(v, v) > 0$$

gilt. Da $f(v, v)$ eine reelle Zahl ist, ist diese Definition sinnvoll.

***Beispiel:** Die Form $\langle x, y \rangle = \overline{x}^t y$ ist positiv definit auf $\mathbb{C}^n$.*

Definition 7.3 *Ein $\mathbb{C}$-Vektorraum V zusammen mit einer positiv definiten hermite'schen Form heißt **unitärer Vektorraum**. In einem unitären Vektorraum $(V, \langle , \rangle)$ ist durch $||v|| = \sqrt{\langle v, v \rangle}$ eine Länge von Vektoren definiert. Diese erfüllt die Eigenschaften i)-iii) aus Lemma 6.10, wobei man in iii) den komplexen Absolutbetrag verwenden muss. (Prüfen Sie das!)*

Nun betrachten wir die Koordinatenmatrizen zu hermite'schen Formen.

Definition 7.4 *Ist f eine hermite'sche Form auf dem $\mathbb{C}$-Vektorraum V und $B = (b_1, \ldots, b_n)$ eine Basis von V, so ist die Koordinatenmatrix $M_{f,B,B}$ von f bezüglich B definiert als die Matrix mit den Einträgen $(f(b_i, b_j))_{i,j}$.*

Lemma 7.5 *Ist f eine hermite'sche Form auf V und $B = (b_1, \ldots, b_n)$ eine Basis von V, so gilt für alle $v, w \in V$ die Gleichung*

$$f(v, w) = \overline{\kappa_b(v)}^t M_{f,B,B} \kappa_B(w).$$

Beweis : analog zu Lemma 5.4 □

Proposition 7.6 *Ist $M = M_{f,B,B}$ die Koordinatenmatrix einer hermite'schen Form f, so gilt*

$$\overline{M}^t = M.$$

Beweis : Der Eintrag von M an der Stelle (i, j) ist definitionsgemäß

$$f(b_i, b_j) = \overline{f(b_j, b_i)},$$

also gleich dem Eintrag von $\overline{M}^t$ an der Stelle (i, j). □

Definition 7.7 *Für $A \in \mathbb{C}^{n \times n}$ schreiben wir $A^* = \overline{A}^t$ und nennen A^* die Adjungierte zu A. Eine Matrix A heißt **hermite'sch**, falls $A = A^*$ gilt.*

Hier muss man mit der Terminologie aufpassen: Die Adjungierte A^* hat nichts mit der adjungierten Matrix aus den Cramer'schen Regeln zu tun.

Ist $A \in \mathbb{C}^{n \times n}$ hermite'sch, so ist A von der Form

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \overline{a_{12}} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \ldots & a_{nn} \end{pmatrix}$$

mit reellen Diagonaleinträgen $a_{11}, \ldots, a_{nn}$.

Beispiel: $A = \begin{pmatrix} 2 & i \\ -i & 1 \end{pmatrix}$ ist hermite'sch. Eine Matrix mit reellen Einträgen ist hermite'sch, wenn sie symmetrisch ist.

Für Koordinatenmatrizen hermite'scher Formen hat man folgendes Basiswechselresultat:

Proposition 7.8 *Sei V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum und $f : V \times V \to \mathbb{C}$ hermite'sch. Sind B und B' Basen von V und $P = P_B^{B'}$ die Übergangsmatrix, so gilt*

$$M_{f,B',B'} = \overline{P}^t M_{f,B,B} P.$$

Beweis : Ist $B = (b_1, \ldots, b_n)$ und $B' = (b'_1, \ldots, b'_n)$, so ist der Koeffizient an der Stelle (i, j) von $M_{f,B',B'}$ gerade $f(b'_i, b'_j) \overset{7.5}{=} \overline{\kappa_B(b'_i)^t} M_{f,B,B} \kappa_B(b'_j)$, also gleich dem Koeffizienten an der Stelle (i, j) von $\overline{P}^t M_{f,B,B} P$. □

Definition 7.9 *Ist $(V, \langle , \rangle)$ ein endlich-dimensionaler unitärer Vektorraum, so heißt eine Basis $B = (b_1, \ldots, b_n)$* ***Orthonormalbasis von V****, falls $\langle b_i, b_j \rangle = \delta_{ij}$ gilt.*

Man kann das Gram-Schmidt-Verfahren auf unitäre Vektorräume übertragen und so die Existenz einer Orthonormalbasis zeigen.

Satz 7.10 *Es sei V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum und $f : V \times V \to \mathbb{C}$ eine positiv definite hermite'sche Form. Dann besitzt V eine Orthonormalbasis bezüglich f.*

Beweis : analog zum Beweis von Satz 6.5 □

Definition 7.11 *Ist $(V, \langle , \rangle)$ ein unitärer Vektorraum, so heißt eine lineare Abbildung $f : V \to V$ eine Isometrie von V (bezüglich $\langle , \rangle$), falls für alle $v, w \in V$*

$$\langle f(v), f(w) \rangle = \langle v, w \rangle$$

gilt.

Ist V endlich-dimensional, so ist jede Isometrie ein Isomorphismus und wie im euklidischen Fall bildet die Menge aller Isometrien $f : V \to V$ eine Gruppe, die auch mit $U(V, \langle \rangle)$ bezeichnet wird.

Definition 7.12 *Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt* ***unitär****, falls $A^* A = E_n$ gilt.*

Jede unitäre Matrix A ist invertierbar mit $A^{-1} = A^*$. Also gilt auch $AA^* = E_n$. Ist A unitär, so ist $|\det A|^2 = 1$.

Definitionsgemäß ist $A \in \mathbb{C}^{n\times n}$ genau dann unitär, wenn $\lambda_A : \mathbb{C}^n \times \mathbb{C}^n$ eine Isometrie bezüglich der kanonischen hermite'schen Form

$$\begin{aligned} \mathbb{C}^n \times \mathbb{C}^n &\to \mathbb{C} \\ (x, y) &\mapsto \overline{x}^t y \end{aligned}$$

ist.

Die reellen unitären Matrizen sind gerade die orthogonalen Matrizen.

Die Menge $U(n, \mathbb{C}) = \{A \in \mathbb{C}^{n\times n} : A \text{ unitär }\}$ ist eine Untergruppe von $\mathrm{GL}(n, \mathbb{C})$, die sogenannte **unitäre Gruppe**.

Die Menge $\mathrm{SU}(n, \mathbb{C}) = \{A \in U(n, \mathbb{C}) : \det A = 1\}$ der unitären Matrizen der Determinante 1 bildet eine Untergruppe von $U(n, \mathbb{C})$, sie heißt **spezielle unitäre Gruppe**.

Ab jetzt sei $(V, \langle , \rangle)$ ein n-dimensionaler unitärer Raum, d.h. $\langle , \rangle$ ist eine positiv definite hermite'sche Form auf V. Wir wollen nun für geeignete Endomorphismen $f : V \to V$ eine Orthonormalbasis von $(V, \langle , \rangle)$ finden, so dass die Koordinatenmatrix von f Diagonalgestalt hat.

Satz 7.13 *Es sei $f : V \to V$ linear und es sei B eine Orthonormalbasis von $(V, \langle , \rangle)$. Wir betrachten die Koordinatenmatrix $A = A(f)_B^B$ der linearen Abbildung f.*

*i) A ist genau dann hermite'sch, wenn $\langle v, f(w)\rangle = \langle f(v), w\rangle$ für alle $v, w \in V$ gilt. Die lineare Abbildung f heißt dann **selbstadjungiert** bezüglich $\langle , \rangle$.*

ii) A ist genau dann unitär, wenn $\langle v, w\rangle = \langle f(v), f(w))\rangle$ für alle $v, w \in V$ gilt, also genau dann, wenn f eine Isometrie von $(V, \langle , \rangle)$ ist.

Beweis : Da B eine Orthonormalbasis ist, ist die Koordinatenmatrix von $\langle , \rangle$ bezüglich B gerade die Einheitsmatrix. Nach Lemma 7.5 gilt also für alle $x, y \in V$, dass $\langle x, y\rangle = \overline{\kappa_B(x)^t}\kappa_B(y)$ ist. Ferner gilt nach [LA 1], Proposition 6.11 für alle $x \in V$, dass $\kappa_B(f(x)) = A\kappa_B(x)$ ist.

i) Ist A hermite'sch, d.h. gilt $\overline{A}^t = A$, so folgt für alle $v, w \in V$:

$$\begin{aligned}
\langle v, f(w)\rangle &= \overline{\kappa_B(v)}^t \kappa_B(f(w)) \\
&= \overline{\kappa_B(v)}^t A\kappa_B(w) \\
&= \overline{\kappa_B(v)}^t \overline{A}^t \kappa_B(w) \\
&= \overline{(A\kappa_B(v))}^t \kappa_B(w) \\
&= \langle f(v), w\rangle.
\end{aligned}$$

Ist umgekehrt f selbstadjungiert bezüglich $\langle , \rangle$, so folgt für Elemente der Orthonormalbasis $B = (b_1, \ldots, b_n)$:

$$\langle b_i, f(b_j)\rangle = \langle f(b_i), b_j\rangle,$$

also gilt

$$\begin{aligned}
e_i A e_j &= \kappa_B(b_i)^t A\kappa_B(b_j) \\
&= \kappa_B(b_i)^t \kappa_B(f(b_j)) \\
&= \langle b_i, f(b_j)\rangle = \langle f(b_i), b_j\rangle \\
&= \overline{\kappa_B(f(b_i))}^t \kappa_B(b_j) \\
&= \overline{A\kappa_B(b_i))}^t \kappa_B(b_j) \\
&= \overline{\kappa_B(b_i)}^t \overline{A}^t \kappa_B(b_j) \\
&= e_i^t \overline{A}^t e_j,
\end{aligned}$$

woraus $A = \overline{A}^t$, d.h. A hermite'sch, folgt.

ii) Ist A unitär, d.h. gilt $A^*A = E_n$, so folgt für alle $v, w \in V$:

$$\begin{aligned}
\langle f(v), f(w)\rangle &= \overline{\kappa_B(f(v))}^t \kappa_B(f(w)) \\
&= \overline{A\kappa_B(v))}^t (A\kappa_B(w)) \\
&= \overline{\kappa_B(v)}^t \overline{A}^t A\kappa_B(w) \\
&= \overline{\kappa_B(v)}^t A^* A\kappa_B(w) \\
&= \overline{\kappa_B(v)}^t \kappa_B(w) = \langle v, w\rangle.
\end{aligned}$$

Ist umgekehrt f eine Isometrie bezüglich $\langle , \rangle$, so gilt für alle $i, j \in \{1, \ldots, n\}$:

$$\begin{aligned} e_i^t A^* A e_j &= \overline{\kappa_B(b_i)}^t \overline{A}^t A \kappa_B(b_j) \\ &= \overline{A\kappa_B(b_i)})^t (A\kappa_B(b_j)) \\ &= \overline{\kappa_B(f(b_i))}^t \kappa_B(f(b_j)) \\ &= \langle f(b_i), f(b_j) \rangle \\ &= \langle b_i, b_j \rangle \\ &= \delta_{ij}, \end{aligned}$$

d.h. $A^* A = E_n$.

□

Jetzt können wir folgenden wichtigen Satz zeigen:

Satz 7.14 (Spektralsatz, auch Satz von der Hauptachsentransformation genannt)

i) Sei $(V, \langle\rangle)$ ein endlich-dimensionaler unitärer Vektorraum und $f\colon V \to V$ eine selbstadjungierte lineare Abbildung, d.h. es gilt $\langle f(v), w\rangle = \langle v, f(w)\rangle$ für alle $v, w \in V$. Dann gibt es eine Orthonormalbasis von $(V\langle , \rangle)$, die nur aus Eigenvektoren von f besteht. Insbesondere ist f diagonalisierbar.

ii) Sei A eine hermite'sche Matrix in $\mathbb{C}^{n\times n}$. Dann gibt es eine unitäre Matrix $P \in U(n, \mathbb{C})$, so dass

$$PAP^* \text{ eine Diagonalmatrix}$$

mit reellen Einträgen ist.

Beweis : (mit Induktion nach $n = \dim V$)

i) Ist $\dim V = 0$, so ist die Behauptung klar. Wir nehmen also an, $n = \dim V \geq 1$ und die Behauptung gilt für alle Vektorräume der Dimension $< n$. Das charakteristische Polynom $\chi_f(X) \in \mathbb{C}[X]$ hat nach dem Fundamentalsatz der Algebra eine Nullstelle $a \in \mathbb{C}$. Diese ist ein Eigenwert von f, also besitzt f einen Eigenvektor b_1. Wir normieren b_1 so, dass $\langle b_1, b_1\rangle = 1$ ist (indem wir gegebenenfalls b_1 durch $\frac{b_1}{||b_1||}$ ersetzen). Indem wir b_1 zu einer beliebigen Basis von V ergänzen und dann das Gram-Schmidt-Verfahren anwenden, können wir b_1

zu einer Orthonormalbasis $B = (b_1, \ldots, b_n)$ ergänzen. Die Koordinatenmatrix $A(f)_B^B$ hat dann folgende Gestalt:

$$A(f)_B^B = \begin{pmatrix} a & * & \ldots & * \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$$

mit einer Matrix $A' \in \mathbb{C}^{(n-1)\times(n-1)}$.

Da f selbstadjungiert ist, ist nach Satz 7.13 die Matrix $A(f)_B^B$ hermite'sch, d.h. es gilt $\overline{A(f)_B^B}^t = A(f)_B^B$.

Also folgt $A(f)_B^B = \begin{pmatrix} a & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$ mit einer hermite'schen Matrix A'.

Somit bildet f den von $B' = (b'_2, \ldots, b'_n)$ erzeugten Unterraum W von V in sich ab, und es gilt $A(f|_W)_{B'}^{B'} = A'$. Da A' hermite'sch ist, ist $f_{|W}$ nach Satz 7.13 selbst-adjungiert. Also gibt es nach Induktionsvoraussetzung eine Orthonormalbasis $C = (c_2, \ldots, c_n)$ von W, so dass $A(f|_W)_C^C$ eine Diagonalmatrix ist. Aber dann ist $(b_1, c_2, \ldots, c_n)$ eine Orthonormalbasis von V, so dass die Koordinatenmatrix von f bezüglich dieser Basis eine Diagonalmatrix ist.

ii) Wir betrachten die lineare Abbildung $\lambda_A : \mathbb{C}^n \to \mathbb{C}^n$. Da A hermite'sch ist, ist λ_A nach Satz 7.13 selbstadjungiert bezüglich des kanonischen Skalarprodukts $\langle , \rangle$ auf $\mathbb{C}^n$.

Nach i) existiert also eine Orthonormalbasis $B = (b_1, \ldots, b_n)$ von $\mathbb{C}^n$ bezüglich $\langle , \rangle$, so dass die Koordinatenmatrix von λ_A bezüglich B Diagonalgestalt hat.

Ist P die Übergangsmatrix von der kanonischen Basis zu B, so ist also PAP^{-1} eine Diagonalmatrix. Nun ist P^{-1} die Übergangsmatrix von B zur kanonischen Basis, also die Matrix mit den Spalten $b_1, \ldots, b_n$. Da B eine Orthonormalbasis bezüglich $\langle , \rangle$ ist, folgt also $(P^{-1})^* P^{-1} = E_n$, d.h. P^{-1} ist unitär. Somit ist auch P unitär mit $P^{-1} = P^*$, woraus die Behauptung folgt.

□

Korollar 7.15 *Es sei $(V, \langle , \rangle)$ ein n-dimensionaler unitärer Vektorraum und $q : V \times V \to \mathbb{C}$ eine weitere hermite'sche Form auf V. Dann gibt es eine Orthonormalbasis B von $(V, \langle , \rangle)$, die gleichzeitig Orthogonalbasis von q ist. Die Formen $\langle , \rangle$ und q lassen sich also simultan diagonalisieren.*

Beweis : Es sei C eine beliebige Orthonormalbasis von $(V, \langle\rangle)$. Nach Proposition 7.6 ist die Koordinatenmatrix $M = M_{q,C,C}$ hermite'sch. Also existiert nach Satz 7.14 ii) ein $P \in U(n, \mathbb{C})$, so dass PMP^* eine Diagonalmatrix ist. Die Matrix P ist nach [LA 1], Lemma 6.13 die Übergangsmatrix von C zu einer Basis $B = (b_1, \ldots, b_n)$ von V. Also ist $P^* = P^{-1} = P_C^B$, so dass aus Proposition 7.8 $PMP^* = M_{f,B,B}$ folgt. Da diese Matrix eine Diagonalmatrix ist, ist B eine Orthogonalbasis bezüglich q. Ferner ist $M_{\langle,\rangle,C,C} = E_n$, also folgt wieder mit Proposition 7.8: $M_{\langle,\rangle,B,B} = PE_nP^* = E_n$. Somit ist B eine Orthonormalbasis bezüglich $\langle , \rangle$. □

Korollar 7.16 *Ist $f : V \to V$ eine selbstadjungierte Abbildung eines endlichdimensionalen unitären $\mathbb{C}$-Vektorraums, so sind Eigenvektoren zu verschiedenen Eigenwerten orthogonal zueinander. Insbesondere sind sie linear unabhängig.*

Beweis : Nach dem Spektralsatz gibt es eine Orthonormalbasis $B = (b_1, \ldots, b_n)$ von V, die aus Eigenvektoren von f besteht. Wir nummerieren die b_i so, dass für die paarweise verschiedenen Eigenwerte $\alpha_1, \ldots, \alpha_r$ von f gilt:

$$\langle b_1, \ldots, b_{m_1}\rangle = V_{\alpha_1}, \langle b_{m_1+1}, \ldots, b_{m_1+m_2}\rangle = V_{\alpha_2}, \ldots, \langle b_{m_1+\cdots+m_{r-1}+1}, \ldots, b_n\rangle = V_{\alpha_r}.$$

Sind v und w Eigenvektoren zu verschiedenen Eigenwerten, liegt also $v \in V_{\alpha_i}$ und $w \in V_{\alpha_j}$ für $\alpha_i \neq \alpha_j$, so folgt $v \perp w$. □

Ist f also eine selbstadjungierte Abbildung mit paarweise verschiedenen Eigenwerten $\alpha_1, \ldots, \alpha_n$, so ist jede Basis B aus Eigenvektoren von f eine Orthogonalbasis, die man durch Normieren der Basisvektoren in eine Orthonormalbasis überführen kann

Beispiel: Es sei $A = \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}$. Dann ist

$$\chi_A(X) = (X-2)^2 - 1 = X^2 - 4X + 3 = (X-1)(X-3),$$

d.h. A hat die Eigenwerte 1 und 3. Lösen der Gleichungssysteme $(A - E_2)x = 0$ und $(A - 3E_2)x = 0$ liefert Eigenvektoren

$$v_1' = \begin{pmatrix} 1 \\ -i \end{pmatrix} \text{ zu } 3 \text{ und } v_2' = \begin{pmatrix} 1 \\ i \end{pmatrix} \text{ zu } 1.$$

Sei $v_1 = \frac{1}{||v_1'||} v_1' = \begin{pmatrix} 1/\sqrt{2} \\ -i/\sqrt{2} \end{pmatrix}$ und $v_2 = \frac{1}{||v_2'||} v_2' = \begin{pmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{pmatrix}$. Dann ist (v_1, v_2) eine Orthonormalbasis für $(\mathbb{C}^2, \langle , \rangle)$ aus Eigenvektoren von A. Für

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & i/\sqrt{2} \end{pmatrix}^{-1} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & i/\sqrt{2} \end{pmatrix}^{*} = \begin{pmatrix} 1/\sqrt{2} & i/\sqrt{2} \\ 1/\sqrt{2} & -i/\sqrt{2} \end{pmatrix},$$

gilt also

$$PAP^* = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}.$$

Satz 7.17 *Es sei $(V, \langle , \rangle)$ ein unitärer Vektorraum. Dann sind alle Eigenwerte eines selbstadjungierten Endomorphismus f von V reell. Insbesondere sind also auch alle Eigenwerte einer hermite'schen Matrix reell.*

Beweis : Die Aussage folgt aus dem Spektralsatz, wir können sie aber auch direkt zeigen. Ist v ein Eigenvektor von f zum Eigenwert α, so gilt nämlich $\langle f(v), v\rangle = \langle v, f(v)\rangle$, da f selbstadjungiert ist. Also folgt

$$\overline{\alpha}\langle v, v\rangle = \langle \alpha v, v\rangle = \langle f(v), v\rangle = \langle v, f(v)\rangle = \langle v, \alpha v\rangle = \alpha\langle v, v\rangle.$$

Da $v \neq 0$ ist, ist $\langle v, v\rangle \neq 0$, also ist $\alpha \in \mathbb{R}$. □

Korollar 7.18 *Jede symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ hat nur reelle Eigenwerte, d.h. das charakteristische Polynom $\chi_A(X)$ zerfällt über $\mathbb{R}$ vollständig in Linearfaktoren.*

Beweis : Das folgt aus dem vorhergehenden Satz, wenn wir A als Matrix mit komplexen Einträgen betrachten. Dann ist sie hermite'sch. □

Mit genau demselben Argument wie über den komplexen Zahlen kann man einen analogen Spektralsatz für euklidische Vektorräume beweisen.

Satz 7.19 *(**Spektralsatz für euklidische Vektorräume**, auch Hauptachsentransformation genannt)*

i) Sei $(V, \langle , \rangle)$ ein endlich-dimensionaler euklidischer Vektorraum und $f : V \to V$ eine selbstadjungierte Abbildung, d.h. es gilt $\langle f(v), w\rangle = \langle v, f(w)\rangle$ für alle $v, w \in V$. Dann gibt es eine Orthonormalbasis von $(V, \langle , \rangle)$, die aus Eigenvektoren von f besteht. Insbesondere ist f diagonalisierbar.

ii) Sei A eine symmetrische Matrix in $\mathbb{R}^{n\times n}$. Dann gibt es eine orthogonale Matrix $P \in O(n,\mathbb{R})$, so dass

$$PAP^t$$

eine Diagonalmatrix ist.

Beweis : Ein analoges Induktionsargument wie im Beweis von Satz 7.14 lässt sich hier durchführen, wenn wir Korollar 7.18 verwenden, um einen Eigenvektor zu finden. □

8 Euklidische und faktorielle Ringe

Definition 8.1 *Ein Integritätsring R mit einer Abbildung $\delta : R \setminus \{0\} \to \mathbb{N}_0$ heißt **euklidischer Ring**, wenn gilt:*

Zu allen $f, g \in R$ mit $g \neq 0$ gibt es Elemente $q, r \in R$ mit

$$f = q \cdot g + r \text{ und } (\delta(r) < \delta(g) \text{ oder } r = 0).$$

Die Abbildung δ wird als Grad- oder Normabbildung des euklidischen Rings R bezeichnet.

Beispiel:

i) Jeder Körper ist zusammen mit der Abbildung $\delta = 0$ ein euklidischer Ring.

ii) Ist K ein Körper, so ist $K[X]$ mit der Abbildung $\delta(g) = \text{grad}(g)$ ein euklidischer Ring nach Satz 4.8.

iii) $\mathbb{Z}$ ist ein euklidischer Ring mit $\delta(a) = |a|$ und der gewöhnlichen Division mit Rest.

iv) Wir betrachten $\mathbb{Z}[i] = \{a + ib : a, b \in \mathbb{Z}\} \subset \mathbb{C}$. $\mathbb{Z}[i]$ ist ein Unterring von $\mathbb{C}$ (ÜA). Wir definieren eine Gradabbildung

$$\begin{aligned} \delta : \mathbb{Z}[i] \setminus \{0\} &\to \mathbb{N} \\ (a + ib) &\mapsto |a + ib|^2 = a^2 + b^2. \end{aligned}$$

Für alle $x, y \in \mathbb{R}$ finden wir offenbar $a, b \in \mathbb{Z}$ mit $|x - a| \leq \frac{1}{2}, |y - b| \leq \frac{1}{2}$. Es sei $z = x + iy \in \mathbb{C}$.

Dann gilt für $a + ib \in \mathbb{Z}[i]$:

$$|z - (a + ib)|^2 = |(x - a) + i(y - b)|^2 \leq 2 \cdot \frac{1}{4} < 1.$$

Insbesondere gibt es für beliebiges $f, g \in \mathbb{Z}[i]$ mit $g \neq 0$ ein $q = a + ib \in \mathbb{Z}[i]$ mit

$$\left| \frac{f}{g} - q \right|^2 < 1.$$

Setzen wir $r = f - qg$, so gilt, falls $r \neq 0$ ist, $\delta(r) = |f - qg|^2 < |g|^2 = \delta(g)$.

$\mathbb{Z}[i]$ heißt der **Ring der ganzen Gauß'schen Zahlen.**

v) Allgemeiner sei $d \neq 0, 1$ eine quadratfreie ganze Zahl, d.h. kein Quadrat eines $n > 1, n \in \mathbb{N}$, ist ein Teiler von d. Dann ist

$$R_d = \begin{cases} \mathbb{Z} + \sqrt{d}\mathbb{Z}, & \text{falls } d \equiv 2, 3 \mod 4 \\ \mathbb{Z} + \frac{1}{2}(1 + \sqrt{d})\mathbb{Z}, & \text{falls } d \equiv 1 \mod 4. \end{cases}$$

ein Unterring von $\mathbb{C}$. Für $d = -1$ gilt $R_{-1} = \mathbb{Z}[i]$. Die Ringe R_d sind in der Zahlentheorie von besonderem Interesse. Insbesondere möchte man wissen, ob R_d euklidisch oder faktoriell (s.u.) ist.

Satz 8.2 *Jeder euklidische Ring ist ein Hauptidealring.*

Beweis : Es sei $\mathfrak{a} \subset R$ ein Ideal. Wir können annehmen, dass $\mathfrak{a} \neq 0$ ist. Wir wählen unter den Elementen $a \neq 0$ aus $\mathfrak{a}$ eines mit minimalem Grad $\delta(a)$. Jetzt zeigen wir $\mathfrak{a} = (a)$. Die Inklusion „$\supset$ “ ist trivial. Sei also $b \in \mathfrak{a}$. Dann können wir b schreiben als $b = q \cdot a + r$ mit $q, r \in R$, so dass $r = 0$ oder $\delta(r) < \delta(a)$ ist. Es ist

$$r = b - q \cdot a \in \mathfrak{a},$$

also gilt aufgrund der Definition von a schon $\delta(r) \geq \delta(a)$, woraus $r = 0$ folgt. Daher ist $b = q \cdot a \in (a)$. □

Korollar 8.3 *Die Ringe $\mathbb{Z}, \mathbb{Z}[i]$ und der Polynomring $K[X]$ über einem Körper K sind als euklidische Ringe auch Hauptidealringe.*

Als nächstes wollen wir Primfaktorzerlegungen in Hauptidealringen studieren. Wir sagen, dass in einem Integritätsring R ein Element $x \in R$ das Element $y \in R$ teilt und schreiben $x \mid y$, falls $y = x \cdot a$ für ein $a \in R$ gilt. Also ist $x \mid y$ äquivalent zu $y \in (x)$. Ist diese Bedingung nicht erfüllt, so schreiben wir $x \nmid y$.

Definition 8.4 *Es sei R ein Integritätsring und $p \in R \setminus \{0\}$ keine Einheit*

*i) p heißt **irreduzibel**, falls für jede Zerlegung $p = xy$ mit $x, y \in R$ gilt: $x \in R^*$ oder $y \in R^*$. Ein nicht irreduzibles Element nennen wir auch **reduzibel**.*

*ii) p heißt **prim** oder **Primelement**, falls für alle $x, y \in R$ aus $p \mid xy$ schon $p \mid x$ oder $p \mid y$ folgt, mit anderen Worten, falls das Hauptideal (p) ein Primideal ist (ÜA).*

***Beispiel**: In $\mathbb{Z}$ entsprechen die irreduziblen Elemente genau den Primelementen. Dies sind die Zahlen $\pm p, p$ Primzahl.*

Bemerkung 8.5 *Es sei R ein Integritätsring und $p \in R \setminus \{0\}$ keine Einheit.*

i) Ist $(p) \subset R$ ein maximales Ideal, so ist p ein Primelement.

ii) Primelemente sind irreduzibel.

Beweis :

i) Maximale Ideale sind prim nach Satz 4.21.

ii) Sei p ein Primelement und gelte $p = xy$ in R. Dann teilt p das Produkt xy, woraus $p \mid x$ oder $p \mid y$ folgt. Ist $x = pa$, so folgt $p = xy = pay$, also, da R ein Integritätsring ist, $ay = 1$. Somit ist $y \in R^*$. Falls $p \mid y$, so folgt mit demselben Argument $x \in R^*$.

□

In Hauptidealringen gilt sogar

Satz 8.6 *Sei R ein Hauptidealring und $p \in R \setminus \{0\}$ keine Einheit. Dann sind äquivalent:*

i) p *ist irreduzibel.*

ii) p *ist Primelement.*

iii) (p) *ist ein maximales Ideal.*

Beweis : iii) $\Rightarrow$ ii) $\Rightarrow$ i) folgt allgemein mit Bemerkung 8.5.

i) $\Rightarrow$ iii): Angenommen, p ist irreduzibel und $\mathfrak{a} \subset R$ ist ein Ideal mit $(p) \subset \mathfrak{a}$. Da R ein Hauptidealring ist, ist $\mathfrak{a} = (a)$ für ein $a \in R$. Also existiert ein $c \in R$ mit $p = ac$, woraus $a \in R^*$ oder $c \in R^*$ folgt. Im ersten Fall ist $\mathfrak{a} = (a) = R$, im zweiten Fall ist $\mathfrak{a} = (a) = (p)$. Daher ist (p) maximal $\square$

Als nächstes wollen wir zeigen, dass in Hauptidealringen stets ein Analogon der Primfaktorzerlegung in $\mathbb{Z}$ existiert.

Definition 8.7 *i) Ein Element a eines Ringes R besitzt eine **Zerlegung in irreduzible Faktoren**, wenn sich a schreiben lässt als*

$$a = \varepsilon p_1 \cdots p_r$$

mit $\varepsilon \in R^*$ und irreduziblen $p_1, \ldots p_r$ in R, $r \in \mathbb{N}_0$.

*ii) Wir sagen, $a \in R$ hat eine **eindeutige Zerlegung in irreduzible Faktoren**, falls a eine Zerlegung in irreduzible Faktoren hat, die in folgender Weise eindeutig ist: Sind $a = \varepsilon p_1 \cdots p_r$ und $a = \varepsilon' q_1 \cdots q_s$ zwei Zerlegungen in irreduzible Faktoren, so gilt $r = s$ und nach geeigneter Umnummerierung unterscheidet sich für alle $i = 1, \ldots, r$ das Element p_i von q_i um eine Einheit.*

*iii) Ein Integritätsring R heißt **faktoriell**, falls jedes $a \neq 0$ aus R eine eindeutige Zerlegung in irreduzible Faktoren besitzt.*

Satz 8.8 *Es sei R ein Integritätsring.*

i) R ist faktoriell genau dann, wenn sich jedes Element $a \neq 0$, das keine Einheit ist, als Produkt von Primelementen schreiben kann.

ii) Ist R faktoriell, so ist $a \in R$ genau dann irreduzibel, wenn es prim ist.

Beweis : Wir zeigen zuerst

ii) Alle Primelemente sind irreduzibel nach Bemerkung 8.5. Sei umgekehrt $a \in R$ irreduzibel. Falls $a \mid xy$ gilt, so müssen wir $a \mid x$ oder $a \mid y$ zeigen. Wir können annehmen, dass weder x noch y Einheiten sind. Es seien $x = \varepsilon x_1 \cdots x_r$ und $y = \varepsilon' y_1 \cdots y_s$ Zerlegungen in irreduzible Elemente. Es gibt ein $z \in R$ mit $xy = az$, das wir ebenfalls in irreduzible Elemente zerlegen können. Aus der Eindeutigkeit der Zerlegung folgt dann, dass das irreduzible Element a bis auf eine Einheit gleich einem der x_i oder einem der y_i ist. Also folgt $a \mid x$ oder $a \mid y$.

i) Ist R faktoriell, so ist nach ii) jedes irreduzible Element prim. Daher ist jedes $a \neq 0$, das keine Einheit ist, ein Produkt von Primelementen.

   Umgekehrt nehmen wir an, dass sich jede Nichteinheit $a \neq 0$ in R schreiben lässt als

$$a = \pi_1 \cdots \pi_r.$$

   mit Primelementen π_i. Da die π_i nach Bemerkung 8.5 irreduzibel sind, ist das insbesondere eine Zerlegung von a in irreduzible Elemente. Gilt $a = \varepsilon p_1 \cdots p_s$ mit $\varepsilon \in R^*$ und $p_1, \ldots, p_s$ irreduzibel, so folgt aus $\pi_1 \mid \varepsilon p_1 \cdots p_s$, dass $\pi_1 \mid \varepsilon$ oder $\pi_1 \mid p_j$ für ein $j = 1, \ldots, s$ gilt. Da π_1 keine Einheit ist, gibt es ein $j = 1, \ldots, s$ und ein $c \in R$ mit $p_j = \pi_1 \cdot c$. Da p_j irreduzibel und $\pi_1 \notin R^*$ ist, folgt hieraus $c \in R^*$. Da R ein Integritätsring ist, folgt $\pi_2 \cdots \pi_r = \varepsilon' \prod_{i \neq j} p_i$ für ein $\varepsilon' \in R^*$. Setzt man das Verfahren induktiv fort, so ergibt sich die Eindeutigkeit der Zerlegung von a.

□

Satz 8.9 *Jeder Hauptidealring ist faktoriell.*

Insbesondere ist also jeder euklidische Ring faktoriell, für jeden Körper K also etwa der Ring $K[X]$.

Für den Beweis dieser Aussage benötigen wir das folgende Lemma.

Lemma 8.10 *Jeder Hauptidealring R ist noethersch, d.h. jede aufsteigende Kette von Idealen $\mathfrak{a}_1 \subset \mathfrak{a}_2 \subset \cdots \subset R$ wird stationär in dem Sinne, dass es ein $n \in \mathbb{N}$ mit $\mathfrak{a}_i = \mathfrak{a}_n$ für alle $i \geq n$ gibt.*

Beweis : Die Vereinigung $\bigcup_{i\geq 1} \mathfrak{a}_i = \mathfrak{a}$ einer aufsteigenden Kette von Idealen ist ein Ideal in R (ÜA). Nach Voraussetzung existiert ein $a \in R$ mit $\mathfrak{a} = (a)$. Es gibt ein $n \geq 1$ mit $a \in \mathfrak{a}_n$. Also gilt

$$(a) \subset \mathfrak{a}_n \subset \mathfrak{a} = (a),$$

woraus $\mathfrak{a} = \mathfrak{a}_n$ folgt. Somit ist $\mathfrak{a}_i = \mathfrak{a}_n$ für $i \geq n$, die Idealkette wird also stationär. □

Beweis von 8.9 : Es sei S die Menge der Hauptideale in R, die von all denjenigen $a \neq 0$ erzeugt werden, die keine Zerlegung in irreduzible Faktoren besitzen. Ist $S = \emptyset$, so hat jedes $a \neq 0$ in R eine Zerlegung in irreduzible Faktoren. Nach Satz 8.6 lässt sich dann jede Nichteinheit $a \neq 0$ als Produkt von Primelementen schreiben, und nach Satz 8.8 ist R daher faktoriell. Also müssen wir nur $S = \emptyset$ zeigen.

Wir nehmen an, $S \neq \emptyset$. Dann hat S ein maximales Element $\mathfrak{a}$, d.h. für alle Ideale $\mathfrak{b}$ in R mit $\mathfrak{a} \subsetneq \mathfrak{b}$ ist $\mathfrak{b} \notin S$. Gäbe es nämlich kein solches maximales Element, so könnte man eine Kette von Idealen

$$\mathfrak{a}_1 \subsetneq \mathfrak{a}_2 \subsetneq \cdots \subsetneq \mathfrak{a}_i \subsetneq \ldots$$

für $i \in \mathbb{N}$ konstruieren, so dass alle $\mathfrak{a}_i \in S$ sind. Da R nach Lemma 8.10 noethersch ist, wird diese Kette stationär, d.h. für ein $n \in \mathbb{N}$ ist $\mathfrak{a}_n = \mathfrak{a}_{n+1} = \cdots = \mathfrak{a}_i = \ldots$ Das steht im Widerspruch zur Konstruktion der $\mathfrak{a}_i$.

Das maximale Element $\mathfrak{a}$ von S ist ein Hauptideal, also $\mathfrak{a} = (a)$. Definitionsgemäß kann a keine Einheit und nicht irreduzibel sein. Also können wir a als Produkt $a = bc$ mit $b, c \in R \setminus R^*$ schreiben. Wegen $(a) \subsetneq (b)$ und $(a) \subsetneq (c)$ ist. $(b) \notin S$ und $(c) \notin S$. Somit haben b und c eine Zerlegung in irreduzible Faktoren. Dann hat aber auch das Produkt $a = bc$ eine Zerlegung zu irreduziblen Faktoren, was zu einem Widerspruch führt. □

Man kann die Primfaktorzerlegung in einem faktoriellen Ring R weiter standardisieren, indem man ein Vertretersystem P der Primelemente modulo R^* auswählt. Man wählt also aus jeder Äquivalenzklasse bezüglich der Äquivalenzrelation

$$p \sim q \text{ genau dann, wenn } p = \varepsilon q \text{ für ein } \varepsilon \in R^*$$

genau ein Element aus.

Dann hat jedes $a \in R \setminus \{0\}$ eine Zerlegung

$$a = \varepsilon \prod_{p \in P} p^{v_p(a)}$$

mit eindeutig bestimmten $\varepsilon \in R^*$ und $v_p(a) \in \mathbb{N}_0$, so dass fast alle $v_p(a) = 0$ sind.

In $\mathbb{Z}$ etwa nimmt man üblicherweise $P =$ Menge der Primzahlen. In $K[X]$ kann man $P =$ Menge der irreduziblen normierten Polynome nehmen, da $K[X]^* = K^*$ nach Lemma 4.7 gilt.

Beispiel: Wir haben gesehen, dass euklidische Ringe Hauptidealringe und dass Hauptidealringe faktoriell sind. Insbesondere ist für jeden Körper K der Polynomring $K[X]$ ein Hauptidealring. Jedes irreduzible Polynom $f \in K[X]$ erzeugt also nach Satz 8.6 ein maximales Ideal, d.h.

$$L = K[X]/(f)$$

ist ein Körper. Der kanonische Ringhomomorphismus

$$K \to K[X] \to K[X]/f$$

ist nach Bemerkung 4.14 injektiv. Man kann K also als Teilkörper von L auffassen. Ist $a = X + (f)$ die Restklasse von X in L, so gilt: $f(a)$ (ausgerechnet in L) stimmt mit der Restklasse von $f(X)$ überein, d.h. $f(a) = 0$. In L existiert also eine Nullstelle des irreduziblen Polynoms f.

Diese Konstruktion wollen wir nun benutzen, um endliche Körper zu studieren.

Lemma 8.11 *Es sei $\mathbb{F}$ ein endlicher Körper. Dann ist char $(\mathbb{F}) = p$ eine Primzahl, und es gibt ein $n > 0$, so dass*

$$|\mathbb{F}| = p^n.$$

Beweis : Da $\mathbb{F}$ endlich ist, gibt es ein $n \in \mathbb{N}$ mit $n \cdot 1 = 0$ in $\mathbb{F}$. Also ist char $(\mathbb{F}) \neq 0$ und damit eine Primzahl. $\mathbb{F}$ ist als Körpererweiterung von $\mathbb{F}_p$ ein $\mathbb{F}_p-$ Vektorraum.

Wir setzen $n = \dim_{\mathbb{F}_p} \mathbb{F}$. Dann folgt $|\mathbb{F}| = p^n$ durch Abzählen aller Linearkombinationen einer Basis. □

Für jedes irreduzible Polynom $f \in \mathbb{F}_p[X]$ vom Grad n ist

$$\mathbb{F}_p[X]/(f)$$

nach dem obigen Beispiel ein Erweiterungskörper von $\mathbb{F}_p$, der p^n Elemente enthält.

Beispiel: Das Polynom $X^2 + X + 1 \in \mathbb{F}_2[X]$ ist irreduzibel, denn es hat keine Nullstelle in $\mathbb{F}_2$.

$$\mathbb{F}_2[X]/(X^2 + X + 1)$$

ist also ein Körper mit 4 Elementen. Bezeichnen wir mit x die Restklasse von X modulo $(X^2 + X + 1)$, so enthält $\mathbb{F}_2[X]$ die vier Elemente

$$0, 1, x, 1 + x.$$

(Übungsaufgabe: Stellen Sie eine Additions- und eine Multiplikationstabelle für diesen Körper auf.)

Proposition 8.12 *Ist $\mathbb{F}_q$ ein endlicher Körper mit $q = p^n$ Elementen, so gilt für jedes $x \in \mathbb{F}_q$ die Gleichung $x^q = x$. Ferner ist die multiplikative Gruppe $\mathbb{F}_q^*$ zyklisch der Ordnung $q - 1$.*

Beweis : Für $x = 0$ gilt offenbar $x^q = x$. Wir betrachten die multiplikative Gruppe $\mathbb{F}_q^* = \mathbb{F}_q \backslash \{0\}$. Sie hat $q - 1$ Elemente, also erfüllt nach dem Satz 2.20 jedes $x \in \mathbb{F}_q^*$ die Gleichung $x^{q-1} = 1$, woraus $x^q = x$ folgt.

Für jedes $x \in \mathbb{F}_q^*$ ist die Ordnung $\text{ord}(x)$ ein Teiler von $q - 1$. Wir setzen $m = \max\{\text{ord}(x) : x \in \mathbb{F}_q^*\}$ und wählen ein $\alpha \in \mathbb{F}_q^*$ mit $\text{ord}(\alpha) = m$. Die von α erzeugte zyklische Gruppe $\langle\alpha\rangle$ ist eine Untergruppe der Ordnung m von $\mathbb{F}_q^*$. Ist $x \in \mathbb{F}_q^*$ ein Element mit $\text{ord}(x)|m$, so gilt $x^m = 1$. Das Polynom $x^m - 1$ hat nach Satz 4.9 höchstens m Nullstellen in $\mathbb{F}_q$.

Die Menge $H_m = \{x \in \mathbb{F}_q^* : \text{ord}(x)|m\}$ ist also eine Untergruppe von $\mathbb{F}_q^*$ mit höchstens m Elementen, die $\langle\alpha\rangle$ enthält. Da $\langle\alpha\rangle$ ebenfalls m Elemente enthält, folgt $\langle\alpha\rangle = H_m$.

Angenommen, $z \in \mathbb{F}_q^*$ läge nicht in $\langle\alpha\rangle = H_m$. Dann ist $\text{ord}(z)$ kein Teiler von m, daher ist $\text{kgV}\big(\text{ord}(z), m\big) > m$.. Nach Lemma 2.22 existiert dann ein $y \in \mathbb{F}_q^*$ mit

$$\text{ord}(y) = \text{kgV}\big(\text{ord}(z), m\big) > m$$

in Widerspruch zur Definition von m. Also folgt $\langle\alpha\rangle = \mathbb{F}_q^*$ und daher ist $\mathbb{F}_q^*$ in der Tat zyklisch. □

In der Algebra studiert man Zerfällungskörper von Polynomen über Körpern. Wir wollen dies hier am Beispiel

$$X^q - X \in \mathbb{F}_p[X]$$

für eine Primzahl p und eine $p-$Potenz $q = p^n$ kennenlernen.

Nach Korollar 8.9 lässt sich $X^q - X$ in ein Produkt irreduzibler Faktoren zerlegen. Insbesondere gibt es ein irreduzibles normiertes Polynom $f_1 \in \mathbb{F}_p[X]$ mit

$$(X^q - X) = f_1 \cdot g$$

für ein $g \in \mathbb{F}_p[X]$.

Damit ist $F_1 := \mathbb{F}_p[X]/(f_1)$ ein Erweiterungskörper von $\mathbb{F}_p$ mit $p^{\text{grad}(f_1)}$ vielen Elementen. Mit demselben Argument finden wie einen irreduziblen normierten Faktor f_2 von g in $F_1[X]$ und einen Erweiterungskörper $F_2 = F_1[X]/(f_2)$ von F_1.

Nach endlich vielen Schritten haben wir einen Körper F gefunden, der $\mathbb{F}_p$ enthält, in dem das Polynom $X^q - X$ vollständig in Linearfaktoren zerfällt und in dem alle Elemente die Gleichung $x^q = x$ erfüllen.

Daher hat F genau q Elemente. Für jede Primzahlpotenz $q = p^n$ gibt es also einen Körper mit q Elementen. Man kann zeigen, dass je zwei Körper mit q Elementen isomorph sind (ÜA).

Definition 8.13 *Sei R ein Integritätsring und $x_1, \ldots, x_n \in R$.*

*i) $d \in R$ heißt **größter gemeinsamer Teiler** von $x_1, \ldots, x_n$, d.h. $d = ggT(x_1, \ldots x_n)$, falls $d \mid x_i$ für alle $i = 1, \ldots, n$ und falls aus $a \mid x_i$ für alle $i = 1, \ldots, n$ schon $a \mid d$ folgt.*

*ii) $v \in R$ heißt **kleinstes gemeinsames Vielfaches** von $x_1, \ldots, x_n$, d.h. $v = kgV(x_1, \ldots, x_n)$, falls $x_i \mid v$ für alle $i = 1, \ldots, n$ und falls aus $x_i \mid a$ für alle $i = 1, \ldots, n$ schon $v \mid a$ folgt.*

Falls der ggT oder das kgV existieren, so sind sie eindeutig bestimmt bis auf Multiplikation mit einer Einheit in R (ÜA).

Satz 8.14 *Es sei R faktoriell und P ein Vertretersystem der Primelemente von R. Zu beliebigen $x_1, \ldots, x_n \in R$ existieren $ggT(x_1, \ldots, x_n)$ und $kgV(x_1, \ldots, x_n)$. Ist $x_i = \varepsilon_i \prod_{p \in P} p^{v_p(x_i)}$ für $i = 1, \ldots, n$, so ist (bis auf Multiplikation mit einer Einheit)*

$$\begin{aligned} ggT(x_1, \ldots, x_n) &= \prod_{p \in P} p^{\min(v_p(x_1), \ldots, v_p(x_n))} \\ kgV(x_1, \ldots, x_n) &= \prod_{p \in P} p^{\max(v_p(x_1), \ldots, v_p(x_n))} \end{aligned}$$

Beweis : Geht genau wie für $R = \mathbb{Z}$ (ÜA). □

Satz 8.15 *Seien $x_1, \ldots, x_n$ Elemente eines Integritätsringes R.*

i) Falls $(x_1, \ldots, x_n) = (d)$ ein Hauptideal ist, so ist $d = ggT(x_1, \ldots, x_n)$.

ii) Falls $(x_1) \cap \cdots \cap (x_n) = (v)$ *ein Hauptideal ist, so ist* $v = kgV(x_1, \ldots, x_n)$.

Beweis :

i) Ist $(x_1, \ldots, x_n) = (d)$, so gilt $d \mid x_i$ für alle i. Da $d \in (x_1, \ldots, x_n)$ ist, gilt $d = a_1x_1 + \cdots + a_nx_n$ für $a_1, \ldots, a_n \in R$. Jeder gemeinsame Teiler der x_i teilt also d.

ii) Ist $(x_1) \cap \cdots \cap (x_n) = (v)$, so ist $v \in (x_i)$ für alle i, d.h. $x_i \mid v$. Falls a ein gemeinsames Vielfaches aller x_i ist, so ist $a \in (x_1) \cap \cdots \cap (x_n) = (v)$, d.h. $v \mid a$.

□

Literatur

[LA 1] A. Werner. *Lineare Algebra 1.* Vorlesungsskript.