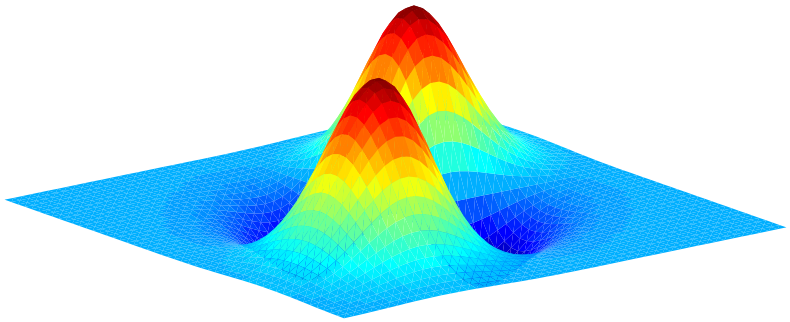
Skript zur Vorlesung

Lineare Algebra II

Private Mitschrift



Algebraische Grundstrukturen
Ausbau der Linearen Algebra

gelesen von

Prof. Dr. Alexander Prestel

Martin Gubisch

Konstanz, Sommersemester 2006

Inhaltsverzeichnis

1. Anordnungen

1.1. Geordnete Mengen

Definition 1.1.

Seien X eine Menge und ρ eine zweistellige Relation auf X, d.h. $\rho \subseteq X \times X = \{(x,y) \;\; x,y \in X\}$.

Wir schreiben oft $x \;\rho\; y$ statt $(x,y) \in \rho$.

ρ heißt eine Ordnung auf X und (X,ρ) heißt eine geordnete Menge, wenn gelten:

1. $x \;\rho\; x$ für alle $x \in X$, (Reflexivität)
2. $(x \;\rho\; y$ und $y \;\rho\; z) \Rightarrow x \;\rho\; z$ für alle $x,y,z \in X$, (Transitivität)
3. $(x \;\rho\; y$ und $y \;\rho\; x) \Rightarrow x = y$ für alle $x,y \in X$. (Antisymmetrie)

Eine Ordnung ρ heißt vollständig oder linear, wenn zusätzlich gilt:

4. $x \;\rho\; y$ oder $y \;\rho\; x$ für alle $x,y \in X$. (Linearität)

Andernfalls heißt ρ partiell.

Beispiel 1.2.

1. Für eine Menge M bezeichne $\mathcal{P}(M) = \{N \mid N \subseteq M\}$ die Potenzmenge. Dann definiert die Inklusionsbeziehung
$$\rho = \{(N_1, N_2) \in \mathcal{P}(M) \times \mathcal{P}(M) \mid N_1 \subseteq N_2\}$$
eine Ordnung auf $\mathcal{P}(M)$. Diese ist im Allgemeinen nicht vollständig: Sei etwa $M = \{m_1, m_2\}$ mit $m_1 \neq m_2$, dann gilt für $N_1 = \{m_1\}$ und $N_2 = \{m_2\}$ weder $N_1 \;\rho\; N_2$ noch $N_2 \;\rho\; N_1$.
2. Entsprechend ist auch $\supseteq$ eine im Allgemeinen unvollständige Ordnung auf $\mathcal{P}(M)$.
3. Die natürlichen Zahlen $\mathbb{N}$ zusammen mit der Teilbarkeitsrelation
$$\rho = \{(n_1, n_2) \in \mathbb{N} \times \mathbb{N} \mid n_1 \text{ teilt } n_2\}$$
eine unvollständige Ordnung auf $\mathbb{N}$, die wir mit $\mid$ bezeichnen.
4. $\leq$ und $\geq$ definieren vollständige Ordnungen auf $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$. ◊

Bemerkung 1.3.

Sei (X,ρ) eine geordnete Menge. Dann ist $(Y, \rho|_Y)$ für jedes $Y \subseteq X$ eine geordnete Menge, wobei
$$\rho|_Y = \{(x,y) \in Y \times Y \mid x \;\rho\; y\} = \rho \cap (Y \times Y).$$
$\rho|_Y$ heißt die Einschränkung von ρ auf Y. ◊

Definition 1.4.

Y heißt eine Kette in (X,ρ), wenn $(Y, \rho|_Y)$ eine vollständig geordnete Menge ist.

Beispiel 1.5.

1. Ist ρ vollständig, dann ist jedes $Y \subseteq X$ eine Kette.
2. $\{2^n \mid n \in \mathbb{N}\}$ ist eine Kette in $(\mathbb{N}, \mid)$.
3. $\{\emptyset, \{1\}, \{1,2\}, \{1,2,3\}, ...\} = \{\{1, ..., n\} \mid n \in \mathbb{N}\}$ bildet eine Kette endlicher Mengen in $(\mathcal{P}(\mathbb{N}), \subseteq)$. ◊

Bemerkung 1.6.

Wir schreiben meist $\preccurlyeq$ statt ρ und setzen $x \prec y$ für $x \preccurlyeq y$ und $x \neq y$. ◊

1.2. Schranken und Extrempunkte

Definition 1.7.

Seien $(X, \preccurlyeq)$ eine geordnete Menge und $A \subseteq X$. Wir nennen $a \in X$ ...

1. maximales Element von A in $(X, \preccurlyeq)$, falls $a \in A$ und es kein $b \in A$ gibt mit $a \prec b$.
2. minimales Element von A in $(X, \preccurlyeq)$, falls $a \in A$ und es kein $b \in A$ gibt mit $a \succ b$.
3. obere Schranke von A, falls $b \preccurlyeq a$ für alle $b \in A$.
4. untere Schranke von A, falls $b \succcurlyeq a$ für alle $b \in A$.
5. größtes Element oder Maximum von A, falls $a \in A$ und $b \preccurlyeq a$ für alle $b \in A$.
6. kleinstes Element oder Minimum von A, falls $a \in A$ und $b \succcurlyeq a$ für alle $b \in A$.
7. Supremum von A, falls a kleinstes Element in der Menge der oberen Schranken von A ist.
8. Infimum von A, falls a größtes Element in der Menge der unteren Schranken von A ist.

Bemerkung 1.8.

1. Kleinste und größte Elemente und damit auch Suprema und Infima sind – sofern wenn vorhanden – eindeutig bestimmt.
2. Wir bezeichnen das Maximum bzw. Minimum einer Menge A mit $\max A$ bzw. $\min A$. Für Supremum bzw. Infimum schreiben wir $\sup A$ bzw. $\inf A$.
3. Es besteht ein Unterschied zwischen maximalem Element und Maximum bzw. minimalem Element und Minimum:

$$\begin{aligned} a \text{ Maximum von } A &\iff a \in A \text{ und } \forall b \in A : a \succcurlyeq b, \text{ dagegen} \\ a \text{ maximales Element von } A &\iff a \in A \text{ und } \forall b \in A : a \not\prec b. \end{aligned}$$

In vollständig geordneten Mengen gibt es diesen Unterschied nicht. ◊

Beispiel 1.9.

1. Sei M eine Menge. In der partiell geordneten Potenzmenge $(\mathcal{P}(M), \subseteq)$ gilt für alle Systeme $A \subseteq \mathcal{P}(M)$: $\sup A = \bigcup A = \bigcup\{M \mid M \in A\}$ und $\inf A = \bigcap A = \bigcap\{M \mid M \in A\}$.
2. Wegen $\bigcup A = \{x \mid \text{für ein } y \in A \text{ gilt } x \in y\}$ und $\bigcap A = \{x \mid \text{für alle } y \in A \text{ gilt } x \in y\}$ ist $\bigcup \emptyset = \emptyset$ und $\bigcap \emptyset = M$, d.h. $\inf \emptyset = M$ und $\sup \emptyset = \emptyset$.
3. Sei $\mathcal{U}$ die Menge aller Untervektorräume eines Vektorraums V. Wir versehen $\mathcal{U}$ mit der partiellen Ordnung $U \preccurlyeq V :\Leftrightarrow U$ ist ein Untervektorraum von V. Dann gelten $\sup\{U, V\} = U + V$ und $\inf\{U, V\} = U \cap V$.
4. $\sup\{6, 10\} = 30$ und $\inf\{6, 10\} = 2$ in $(\mathbb{N}, |\,)$.
5. In $(\mathbb{N}, |\,)$ ist die Menge der Primzahlen zugleich Menge der maximalen und der minimalen Elemente von sich selbst.
6. Für $X = (\mathbb{N}_{\geq 2}, |\,)$ gilt: X besitzt keine maximalen Elemente. Die Menge der minimalen Elemente von X ist die Menge der Primzahlen. ◊

Lemma 1.10.

Seien X nichtleer und endlich und $(X, \preccurlyeq)$ eine geordnete Menge. Dann hat X ein maximales Element.

Beweis. (per Induktion über $n = \sharp X$)

Sei $X = \{x\}$, dann ist x maximales Element von X. Sei $X = \{x_1, ..., x_n\}$, dann besitzt $\{x_1, ..., x_{n-1}\}$ ein maximales Element, Œ x_1. Dann gilt $x_1 \preccurlyeq x_n$, d.h. x_n ist maximal, oder x_1 ist auch maximal in X. □

1.3. Das Lemma von Zorn

Satz 1.11. (Lemma von Zorn)

Sei $(X, \preccurlyeq)$ eine geordnete Menge und jede Kette $Y \subseteq X$ besitze eine obere Schranke in $(X, \preccurlyeq)$.

Dann besitzt X ein maximales Element.

Bemerkung 1.12.

1. Der Beweis kann per transfiniter Induktion geführt werden. Benötigt wird hierfür das *Auswahlaxiom*.
2. Dass die leere Kette $\emptyset$ eine obere Schranke besitzt, heißt nichts anderes, als dass X nichtleer ist. ◊

Korollar 1.13.

Seien $X \subseteq \mathcal{P}(M)$ nichtleer und $\bigcup A \in X$ für alle nichtleeren Ketten A in $(X, \preccurlyeq)$.

Dann besitzt X ein maximales Element in $(X, \preccurlyeq)$.

Beweis.

Folgt unmittelbar aus dem Zornschen Lemma: Die leere Kette hat eine obere Schranke, da $X \neq \emptyset$, und für jede nichtleere Kette A ist $\bigcup A \in X$ eine obere Schranke in $(X, \preccurlyeq)$. Damit sind alle Voraussetzungen erfüllt. □

Bemerkung 1.14.

Zum Zornschen Lemma sind die folgenden Aussagen äquivalent:

1. Zu jeder Menge X gibt es eine Auswahlfunktion $F : \mathcal{P}(X) \to X$, so dass für jedes nichtleere $U \in \mathcal{P}(X)$ gilt $F(U) \in U$. (Auswahlaxiom)
2. Das Produkt einer Familie nichtleerer Mengen ist nicht leer. (Zermelo)
3. In jeder geordneten Menge existiert eine maximale Kette. (Hausdorff)
4. Auf jeder Menge X gibt es eine Ordnung $\preccurlyeq$, so dass jede nichtleere Teilmenge von X bzgl. $\preccurlyeq$ ein kleinstes Element hat. (Wohlordnungssatz)
5. Jede Äquivalenzrelation auf einer Menge besitzt ein Repräsentantensystem. ◊

Definition 1.15.

Seien V ein K-Vektorraum und $S \subseteq V$.

S heißt linear unabhängig, wenn jede endliche Teilmenge von S linear unabhängig ist.

S heißt ein Erzeugendensystem von V (span $E = V$), wenn jedes $v \in V$ als endliche Linearkombination von Elementen aus S geschrieben werden kann.

S heißt eine Basis von V, wenn S ein minimales Erzeugendensystem von V ist, d.h. ein minimales Element von $\{E \subseteq V \mid \text{span } E = V\}$ in der geordneten Menge $(\mathcal{P}(V), \subseteq)$.

Bemerkung 1.16.

1. Für endlichdimensionale Vektorräume gelten: $\mathfrak{V} = (v_1, ..., v_n)$ ist eine Basis von $V \Leftrightarrow \sharp\mathfrak{V} = n$ & $\mathfrak{V}$ ist eine Basis von V, und $\mathfrak{V}$ ist linear unabhängig $\Leftrightarrow \sharp\mathfrak{V} = n$ & $\mathfrak{V}$ ist linear unabhängig.
2. Sei V ein K-Vektorraum. Für $S \subseteq V$ sind äquivalent:

$$\begin{aligned} S \text{ ist eine Basis von } V &\iff S \text{ ist ein linear unabhängiges Erzeugendensystem von } V \\ &\iff S \text{ ist eine maximale linear unabhängige Teilmenge von } V. \end{aligned}$$ ◊

Satz 1.17.

Jeder Vektorraum besitzt eine Basis.

Beweis.

Setze $X = \{S \subseteq V \mid S \text{ ist linear unabhängig }\} \subseteq \mathcal{P}(V)$. Es ist nur zu prüfen, ob $(X, \subseteq)$ die Voraussetzungen des Korollars zu Zorns Lemma erfüllt.

Es gilt: $X \neq \emptyset$, da $\emptyset \in X$. Ist A nun eine nichtleere Kette in $(X, \subseteq)$, so ist $T = \bigcup A \in X$, denn seien $v_1, ..., v_n \in T$ paarweise verschieden. Dann gibt es $S_1, ..., S_n \in A$ mit $v_i \in S_i$. Es ist $\{S_1, ..., S_n\} \subseteq A$ eine Kette in $(X, \subseteq)$, so dass nach allfälliger Umnummerierung etwa $S_1 \subseteq S_2 \subseteq \cdots \subseteq S_n$ gilt. Also sind $v_1, ..., v_n \in S_n \in A \subseteq X$, d.h. $(v_1, ..., v_n)$ ist linear unabhängig. □

2. Gruppentheorie

2.1. Gruppen

Definition 2.1.

Seien M eine nicht-leere Menge und $\circ$ eine Verknüpfung auf M, d.h. eine Abbildung $\circ : M \times M \to M$. Wir schreiben $(x, y) \mapsto x \circ y$.

$(M, \circ)$ heißt ein Monoid, falls für alle $x, y, z \in M$ gilt: $(x \circ y) \circ z = x \circ (y \circ z)$ (Assoziativität).

Bemerkung 2.2.

Es macht also Sinn, $x_1 \circ x_2 \circ x_3 \circ \cdots \circ x_n$ ohne Klammerung zu schreiben. ◊

Beispiel 2.3.

1. $\mathbb{N}_0 = \{0, 1, 2, 3, ...\}$ mit der Addition $+$ oder Multiplikation $\cdot$ als Verknüpfung.
2. Analoges gilt für $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$.
3. $x \circ y := x^y$ ist in $\mathbb{N}$ keine assoziative Verknüpfung: $2^{(3^4)} \neq (2^3)^4$.
4. Seien $A, B, C \in \mathfrak{M}_K(n)$. Es gilt $(AB)C = A(BC)$, d.h. die quadratischen Matrizen über einem Körper K zusammen mit der Matrixmultiplikation bilden einen Monoid.
5. Die Selbstabbildungen auf einer beliebigen Menge X mit der Verkettung

$$f \circ g : x \mapsto f(g(x))$$

als Verknüpfung $(\mathrm{Abb}(X, X), \circ)$ bilden einen Monoid. ◊

Definition 2.4.

$e \in M$ heißt neutrales Element von $(M, \circ)$, falls für alle $x \in M$ gilt: $x \circ e = e \circ x = x$.

Bemerkung 2.5.

Besitzt ein Monoid $(M, \circ)$ ein neutrales Element, so ist dieses stets eindeutig bestimmt:

Seien $e \in M$ neutral und $f \in M$. Gelte $f \circ x = x \circ f = x$ für alle $x \in M$. Dann ist $e = e \circ f = f$. ◊

Definition 2.6.

Seien $x, y \in M$. y heißt invers zu x, falls gilt: $x \circ y = y \circ x = e$.

Wir schreiben x^{-1} für das Inverse zu x.

Bemerkung 2.7.

Jedes Element x eines Monoids $(M, \circ)$ hat höchstens ein Inverses:

Seien $y, z \in M$ invers zu x, dann $y = y \circ e = y \circ (x \circ z) = (y \circ x) \circ z) = e \circ z = z$. ◊

Beispiel 2.8.

1. In $(\mathbb{N}, \cdot)$ hat nur das neutrale Element 1 ein Inverses, nämlich sich selbst.
2. In $(\mathbb{Z}, +)$ hat jedes Element x ein Inverses, nämlich $-x$. ◊

Definition 2.9.

Eine Gruppe ist ein Monoid mit neutralem Element, in dem jedes Element ein Inverses hat.

Beispiel 2.10.

1. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, und $(\mathbb{Q}\backslash\{0\}, \cdot)$ sind Gruppen; $(\mathbb{N}, +)$ und $(\mathbb{N}\backslash\{0\}, \cdot)$ sind keine Gruppen.
2. $(\mathfrak{M}_K(n), +)$ und $(\mathfrak{I}_K(n), \cdot)$, die invertierbaren $n \times n$-Matrizen, sind Gruppen. ◊

Lemma 2.11.

Ein Monoid $(G, \circ)$ ist genau dann eine Gruppe, wenn für beliebige Elemente $a, b \in G$ die Gleichungen $a \circ x = b$ und $x \circ a = b$ lösbar sind.

Beweis.

1. Sei zunächst $(G, \circ)$ eine Gruppe. Dann werden die Gleichungen gelöst durch $x = a^{-1} \circ b$ bzw. $x = b \circ a^{-1}$.
2. Existenz des Neutralen: Sei $a \in G$. Dann gibt es ein $e \in G$ mit $e \circ a = a$. Wir zeigen, dass dann bereits $x \circ e = e \circ x = x$ für alle $x \in G$ gilt. Sei also $x \in G$ beliebig. Dann gibt es ein $y \in G$ mit $a \circ y = x$ und es ist $e \circ x = e \circ (a \circ y) = (e \circ a) \circ y = a \circ y = x$. Analog gibt es ein $f \in G$ mit $x \circ f = x$ für alle x. Wegen $e = e \circ f = f$ folgt schließlich: e ist neutral.
3. Existenz der Inversen: Zu $x \in G$ gibt es $y \in G$ mit $y \circ x = e$ und z mit $x \circ z = e$. Dann ist $y = y \circ e = y \circ (x \circ z) = (y \circ x) \circ z = e \circ z = z$, also $y \circ x = e = x \circ y$. □

Definition 2.12.

Ein Monoid $(M, \circ)$ heißt Kommutativität, falls $x \circ y = y \circ x$ für alle $x, y \in M$.

Eine kommutative Gruppe nennen wir auch eine Abelsche Gruppe.

Bemerkung 2.13.

Für kommutative Verknüpfungen schreiben wir meiste $+$, für nicht kommutative Verknüpfungen $\cdot$. ◊

Lemma 2.14.

Sei $(M, \circ)$ ein Monoid mit neutralem Element e.

Dann bildet die Menge $\mathfrak{I}$ der inversen Elemente von M eine Gruppe.

Beweis.

Seien $x, y \in \mathfrak{I}$. Dann ist auch $xy \in \mathfrak{I}$, denn $(xy)(y^{-1}x^{-1}) = x(yy^{-1})x^{-1} = xx^{-1} = e$; analog ist $(y^{-1}x^{-1})(xy) = e$. Außerdem gilt $e \in \mathfrak{I} : ee = e$, d.h. $\mathfrak{I}$ ist unter $\circ$ abgeschlossen. Wegen $x \in \mathfrak{I} \Rightarrow x^{-1} \in \mathfrak{I}$ und $(x^{-1})^{-1} = x$ folgt schließlich die Behauptung. □

2.2. Untergruppen

Definition 2.15.

Sei $(G, \circ)$ eine Gruppe mit neutralem Element e. $(H, \circ)$ heißt Untergruppe von G, falls $H \subseteq G$ und

1. $\forall x, y \in H : x \circ y \in H$, d.h. $\circ : H \times H \to H$, (Abgeschlossenheit)
2. $e \in H$, (Existenz des Neutralen)
3. $x \in H \Rightarrow x^{-1} \in H$. (Existenz der Inversen)

Beispiel 2.16.

1. $U = (2\mathbb{Z}, +)$ ist Untergruppe von $G = (\mathbb{Z}, +)$, wobei $2\mathbb{Z} = \{2z \mid z \in \mathbb{Z}\} = \{z \in \mathbb{Z} \mid \exists z' \in \mathbb{Z} : z = 2z'\}$.
2. $U = (\mathbb{R}^+, \cdot)$ ist eine Untergruppe von $G = (\mathbb{R}^\times, \cdot)$, wobei $\mathbb{R}^\times = \mathbb{R}\backslash\{0\}$ und $\mathbb{R}^+ = \{x \in \mathbb{R} \mid x > 0\}$. ◊

Bemerkung 2.17.

Für $H \neq \emptyset$ genügt es zu zeigen: $x, y \in H \Rightarrow xy^{-1} \in H$, denn:

1. $x, y \in H \Rightarrow xy^{-1} \in H \Rightarrow xy = x(y^{-1})^{-1} \in H$,
2. $x \in H \Rightarrow xx^{-1} = e \in H$,
3. $e, y \in H \Rightarrow y^{-1} = ey^{-1} \in H$. ◊

Satz 2.18.

Der Durchschnitt von beliebig vielen Untergruppen einer Gruppe G ist selbst eine Untergruppe.

Beweis.

Der Durchschnitt ist nicht leer, da e in allen Untergruppen und damit auch im Durschnitt liegt. Seien nun x, y im Durchschnitt, dann liegen x, y in allen Untergruppen von G, d.h. auch xy^{-1} liegt in allen Untergruppen und damit auch im Durchschnitt. Mit obiger Bemerkung folgt also die Behauptung. □

Satz 2.19.

Sei A eine Teilmenge der Gruppe $(G, \circ)$. Dann gilt:

$$\bigcap_{\substack{A \subseteq H \\ H \text{ Untergruppe}}} H = \{a_1 \circ \cdots \circ a_n \mid n \in \mathbb{N},\ a_i \in A \text{ oder } a_i^{-1} \in A \text{ oder } a_i = e\}.$$

Diese Menge bildet eine Gruppe, die mit $\langle A \rangle$ bezeichnet und die von A erzeugte Untergruppe von G genannt wird.

Beweis.

Es gilt $A \subseteq H \Rightarrow \langle A \rangle \subseteq H \Rightarrow \langle A \rangle \subseteq \bigcap H$. Noch zu zeigen: $\bigcap H \subseteq \langle A \rangle$. Es reicht dazu aus, nachzuweisen, dass $\langle A \rangle$ eine Untergruppe von G ist.

Seien $a = a_1 \circ \cdots \circ a_n \in \langle A \rangle$ und $b = b_1 \circ \cdots \circ b_m \in \langle A \rangle$. Wegen

$$ab^{-1} = (a_1 \circ \cdots \circ a_n)(b_1 \circ \cdots \circ b_m)^{-1} = a_1 \circ \cdots \circ a_n \circ b_m^{-1} \circ \cdots \circ b_1^{-1} \in \langle A \rangle$$

folgt die Behauptung. □

Bemerkung 2.20.

1. Sei $A = \{a_1, ..., a_n\}$. Dann schreiben wir auch $\langle A \rangle = \langle a_1, ..., a_n \rangle$.
2. Ist $(G, +)$ abelsch, dann ist $\langle A \rangle = \mathbb{Z}a_1 + \cdots + \mathbb{Z}a_n$. ◊

2.3. Gruppenordnungen

Definition 2.21.

Seien $(G, \circ)$ eine Gruppe und $a \in G$.

Sei $r \in \mathbb{N} \cup \{\infty\}$ minimal mit $a^r = e$. Dann heißt r die Ordnung von a in G.

Sei $a \in G$ mit $\langle a \rangle = G$. Dann heißt G eine zyklische Gruppe.

Bemerkung 2.22.

Sei $a \in (G, \circ)$. Dann ist $\langle a \rangle = \{..., a^{-2}, a^{-1}, e, a, a^2, ...\} = \{a^n \mid n \in \mathbb{Z}\} = a^{\mathbb{Z}}$. Dabei ist $a^0 = e$ und $a^{-n} = (a^{-1})^n$. Es gilt: $a^n a^m = a^{n+m} = a^{m+n} = a^m a^n$. Es können hier zwei Fälle auftreten:

1. Für alle $n \neq m$ gilt: $a^n \neq a^m$, d.h. die Abbildung $\varphi : \mathbb{Z} \to \langle a \rangle$, $n \mapsto a^n$ ist ein bijektiv und $\varphi(n+m) = \varphi(n) \cdot \varphi(m)$.
2. Es gibt $n \neq m$ mit $a^n = a^m$, Œ $n > m$. Dann ist $a^{n-m} = e$ bzw. es gibt $r \in \mathbb{N}$ mit $a^r = e$. Sei dieses r minimal gewählt, d.h. die (endliche) Ordnung von a. Dann hat $\langle a \rangle = \{a^0, a^1, ..., a^{r-1}\}$ r verschiedene Elemente: Sei $n = rq + s$ mit $0 \leq s \leq r-1$. Dann $a^n = a^{rq+s} = a^s \cdot (a^r)^q = a^s$. Falls es nun k, l gibt mit $0 \leq k < l \leq r-1$ und $a^k = a^l$, d.h. $e = a^{l-k}$. Dies widerspricht aber der Minimalität von r. ◊

Bemerkung 2.23.

Seien M eine Menge und $a, b, c \in M$. $\sim$ ist eine Äquivalenzrelation, wenn gelten:

1. $\sim$ ist reflexiv, d.h. $a \sim a$ für alle $a \in M$;
2. $\sim$ ist symmetrisch, d.h. $a \sim b \Rightarrow b \sim a$ für alle $a, b \in M$;
3. $\sim$ ist transitiv, d.h. $a \sim b$ und $b \sim c \Rightarrow a \sim c$ für alle $a, b, c \in M$.

$[a] = \{b \in M \mid a \sim b\}$ heißt eine Äquivalenzklasse.

Äquivalenzklassen sind disjunkt, d.h. es gilt $[a] \cap [b] \neq \emptyset \Rightarrow [a] = [b]$, denn ist $c \in [a] \cap [b]$, dann $a \sim c$ und $b \sim c$, d.h. $a \sim b$ und damit $[a] = [b]$. ◊

Definition 2.24.

Seien G eine Gruppe, U eine Untergruppe von G und $x \in G$.

$Ux = \{ux \mid u \in U\}$ heißt eine Rechtsnebenklasse in G. Entsprechend heißt $xU = \{xu \mid u \in U\}$ Linksnebenklasse. Ist G Abelsch, dann heißt $xU = Ux$ einfach Nebenklasse.

Beispiel 2.25.

$\mathbb{R}^\times$ lässt sich zerlegen in $(1 \cdot \mathbb{R}^{>0}) \cup ((-1) \cdot \mathbb{R}^{>0})$. $\mathbb{Z}$ lässt sich aufteilen in $\mathbb{Z} = (0 + 2\mathbb{Z}) \cup (1 + 2\mathbb{Z})$. ◊

Satz 2.26.

Sei U eine Untergruppe einer Gruppe G. Dann gilt: $\sharp xU = \sharp U$.

Beweis.

Wir zeigen: $\varphi : U \to xU$, $u \mapsto xu$ ist bijektiv. Wegen $\varphi(u) = ux$ für alle $u \in U$ ist φ surjektiv und wegen $xu_1 = xu_2 \Rightarrow u_1 = u_2$ ist φ injektiv. □

Satz 2.27.

Seien G eine Gruppe, U eine Untergruppe von G und $a, b \in G$. Dann definiert $a \sim b :\Leftrightarrow ab^{-1} \in U$ eine Äquivalenzrelation.

Beweis.

1. Es ist $a \sim a$, denn $aa^{-1} = e \in U$.
2. Gelte $a \sim b$, d.h. $ab^{-1} \in U$, dann $ba^{-1} = (ab^{-1})^{-1} \in U$, also $b \sim a$.
3. Gelten $a \sim b$ und $b \sim c$, d.h. $ab^{-1} \in U$ und $bc^{-1} \in U$. Dann $ac^{-1} = (ab^{-1})(bc^{-1}) \in U$, also $a \sim c$. □

Bemerkung 2.28.

Es gelten $a \sim b \Leftrightarrow ab^{-1} \in U \Leftrightarrow a \in Ub$. Die Äquivalenzklassen bzgl. $\sim$ haben die Gestalt $[b] = Ub$; insbesondere gilt $Ub = Uc$ genau dann, wenn $b \sim c$.

Analog definiert auch $a \sim b :\Leftrightarrow a^{-1}b \in U$ eine Äquivalenzrelation mit $a \sim b \Leftrightarrow b \in aU$. ◊

Satz 2.29. (Kleiner Satz von Fermat)

Seien G eine endliche Gruppe und U eine Untergruppe von G. Dann gilt: $\sharp U$ teilt $\sharp G$.

Beweis.

G lässt sich disjunkt zerlegen in $G = Ub_1 \cup \cdots \cup Ub_m$ für gewisse Elemente $b_1, ..., b_m$ aus G, also gilt $\sharp G = m\sharp U$. □

Definition 2.30.

Sei G eine endliche Gruppe. $\sharp G$ heißt die Ordnung von G.

Bemerkung 2.31.

1. Die Ordnung eines Gruppenelements teilt also stets die Gruppenordnung: $\mathrm{Ord}(a) = \sharp\langle a\rangle$ teilt $\sharp G$.
2. Speziell: Ist G zyklisch von Ordnung r, d.h. $G = \langle a\rangle$ für ein $a \in G$, dann gilt $r = \sharp G$. ◊

Korollar 2.32.

Jede Gruppe von Primzahlordnung ist zyklisch.

Beweis.

Sei $a \in G$, $a \neq e$. Dann gilt: $\mathrm{Ord}(a) = \sharp\langle a\rangle$ teilt $\sharp G = p$. Da p prim ist, folgt $\mathrm{Ord}(a) = p$, also $\langle a\rangle = G$. □

Bemerkung 2.33.

Insbesondere lässt sich jede Gruppe von Primzahlordnung von jedem außer dem neutralem Element erzeugen. ◊

Definition 2.34.

Eine Untergruppe U einer Gruppe G heißt ein Normalteiler von G (in Zeichen: $U \lhd G$), falls für alle $a \in G$ gilt: $aU = Ua$, d.h. $aUa^{-1} = U$.

Bemerkung 2.35.

1. $aUa^{-1} \subseteq U$ ist hinreichend für $U \lhd G$, denn sei $a \in U$, dann folgt aus $aUa^{-1} \subseteq U$, dass auch $U = a^{-1}(aUa^{-1})a \subseteq a^{-1}Ua$, d.h. $U = aUa^{-1}$.
2. $aUa^{-1} = U$ heißt: $\forall u \in U : aua^{-1} \in U$, nicht zwingend $aua^{-1} = u$.
3. In Abelschen Gruppen ist jede Untergruppe ein Normalteiler. ◊

2.4. Gruppenhomomorphismen

Definition 2.36.

Seien $(G, \circ)$ und $(H, +)$ zwei Gruppen. Eine Abbildung $f : G \to H$ heißt Gruppenhomomorphismus, falls gilt:

$$\forall x, y \in G : f(x \circ y) = f(x) + f(y).$$

Einen injektiven Gruppenhomomorphismus nennen wir Gruppenmonomorphismus, einen surjektiven Gruppenhomomorphismus Epimorphismus.

Ist ein Gruppenhomomorphismus sowohl injektiv als auch surjektiv, so nennen wir ihn Isomorphismus und die beiden Gruppen isomorph, in Zeichen: $G \cong H$.

Gilt zusätzlich $G = H$, d.h. ist der Homomorphismus eine Selbstabbildung, so nennen wir einen Automorphismus.

Sei f ein Gruppenhomomorphismus. $f(G) = \{f(a) \mid a \in G\} = \{b \in H \mid b = f(a) \text{ für ein } a \in G\}$ heißt das Bild von f in H. $f^{-1}(\{e\}) = \{a \in G \mid f(a) = e\}$ heißt der Kern von f.

Bemerkung 2.37.

1. Es gelten $f(e_G) = e_H$, d.h. Gruppenhomomorphismen bilden neutrale Elemente stets aufeinander ab, und $f(x^{-1}) = f(x)^{-1}$ für alle x, d.h. das Bild des Inversen ist immer das Inverse des Bildes eines Elements.
2. Ist $f : G \to H$ ein Isomorphismus, dann ist auch $f^{-1} : H \to G$ ein Isomorphismus.
3. Bild(f) ist eine Untergruppe von H und Kern(f) ist eine Untergruppe von G; es gilt sogar $H \lhd \text{Kern}(f)$.
4. Genau dann ist f injektiv, wenn Kern$(f) = \{0\}$ erfüllt ist. ◊

Beispiel 2.38.

1. Die Determinantenabbildung $\det : (\mathfrak{I}_K(n), \cdot) \to (K\backslash\{0\}, \cdot)$ auf der allgemeinen linearen Gruppe $\mathfrak{I}_K(n)$ ist ein Gruppenhomomorphismus. det definiert auch einen Gruppenhomomorphismus zwischen der speziellen linearen Gruppe $\{A \in \mathfrak{I}_K(n) \mid |\det(A)| = 1\}$ und $(\{\pm 1\}, \cdot)$.
2. Seien $G = (\mathbb{Z}, +)$ und $(H, \cdot)$ eine Gruppe. Sei $a \in H$ beliebig. Dann ist $f : \mathbb{Z} \to H$, $f(n) = a^n$ ein Homomorphismus.
3. Jeder Vektorraumhomomorphismus induziert einen Gruppenhomomorphismus zwischen den zugehörigen additiven Gruppen.
4. $\exp : (\mathbb{R}, +) \to (\mathbb{R}^{>0}, \cdot)$ liefert einen Gruppenisomorphismus mit inverser Abbildung ln. ◊

Satz 2.39.

Seien G eine Gruppe und N ein Normalteiler von G. Dann definiert $(aN)(bN) := (ab)N$ eine Gruppenoperation auf $H := \{aN \mid a \in G\}$ und $f : G \to H$ mit $f(a) = aN$ ist ein Gruppenhomomorphismus mit Kern N.

Beweis.

1. Wohldefiniertheit: Seien $a, b \in G$. Wir müssen zeigen, dass aus $aN = a'N$ und $bN = b'N$ schon folgt $(ab)N = (a'b')N$. Wegen $a' = a'e \in a'N = aN$ und analog $b' = b'e \in b'N = bN$ ist $a'b'N \subseteq (aN)(bN)N \subseteq abNNN$ (da N Normalteiler); $abNNN \subseteq abN$. Entsprechend ist $abN \subseteq a'b'N$, insgesamt also $abN = a'b'N$.
2. $H = \{aN \mid a \in G\}$ ist eine Gruppe, denn $aN(bNcN) = a(bc)N = abcN = (ab)cN = (aNbN)cN$, d.h. H ist abgeschlossen; weiter ist eN ist neutrales Element: $aNeN = aeN = aN$ (wobei $eN = N$), und aus $aNa^{-1}N = eN$ folgt $(aN)^{-1} = a^{-1}N$, d.h. jedes Element in H besitzt ein Inverses.
3. f ist ein Gruppenhomomorphismus, denn $f(ab) = abN = aNbN = f(a)f(b)$.
4. Kern$(f) = \{a \mid f(a) = eN\} = \{a \mid aN = eN\} = \{a \mid a \in N\} = N$. □

2.5. Faktorgruppen

Definition 2.40.

Seien G eine Gruppe und N ein Normalteiler von G, dann heißt $G/N := \{aN \mid a \in G\}$ Faktorgruppe von G nach N.

$\phi : G \to G/N$, $a \mapsto aN$ heißt der kanonische Homomorphismus.

Bemerkung 2.41.

Der folgende Satz besagt: Jeder Gruppenhomomorphismus $f : G \to H$ lässt sich "einschränken" auf einen Gruppenisomorphismus $\bar{f} : G/N \to f(G)$ mit $N = \text{Kern}(f)$. Das Diagramm

$$\begin{array}{ccc} G & \xrightarrow{f} & H \\ \downarrow{\scriptstyle \phi} & & \uparrow{\scriptstyle \text{id}} \\ G/N & \xrightarrow{\bar{f}} & f(G) \end{array}$$

kommutiert, d.h. es gilt $f = \bar{f} \circ \phi$. ◊

Satz 2.42. (Homomorphiesatz)

Ist $f : G \to H$ ein Gruppenhomomorphismus, so gilt: $f(G) \cong G/N$ mit $N = \text{Kern}(f)$.

Beweis.

1. $\bar{f} : G/N \to f(G)$, $aN \mapsto f(a)$ ist wohldefiniert, d.h. für $aN = a'N$ gilt $\bar{f}(aN) = \bar{f}(a'N)$ bzw. $f(a) = f(a')$: Sei $u \in \text{Kern}(f)$ mit $a' = au$. Dann gilt $f(a') = f(a)f(u) = f(a)e = f(a)$.
2. $\bar{f}$ ist ein Homomorphismus: $\bar{f}((aN)(bN)) = \bar{f}(abN) = f(ab) = f(a)f(b) = \bar{f}(aN)\bar{f}(bN)$.
3. $\bar{f}$ ist injektiv: Sei $aN \in \text{Kern}(\bar{f})$, d.h. $\bar{f}(aN) = f(a) = e$, dann ist $a \in \text{Kern}(f)$, d.h. $aN = eN$ ist das neutrale Element von G/N. □

Beispiel 2.43.

Seien $a \in G$, $A \neq e$ und $H = \langle a \rangle = a^{\mathbb{Z}} = \{..., a^{-2}, a^{-1}, a^0, a^1, a^2, ...\}$. Dann ist $f : \mathbb{Z} \to H$, $f(n) = a^n$ ein surjektiver Homomorphismus. Also ist nach dem Homomorphiesatz $H \cong \mathbb{Z}/N$ mit $N = \text{Kern}(f)$. Es können zwei Fälle auftreten:

1. $N = \{e\}$, dann ist f injektiv, d.h. für $m \neq n$ gilt $a^n \neq a^m$, und damit $H \cong \mathbb{Z}/\{0\} \cong \mathbb{Z}$.
2. Es gibt $n_1, n_2 \in \mathbb{N}$ mit $a^{n_1} = a^{n_2}$, d.h. a besitzt eine endliche Ordnung $\text{Ord}(a) = n$ (es gilt n minimal mit $a^n = e$. Dann ist $N = \text{Kern}(f) = n\mathbb{Z} = \{nr \mid r \in \mathbb{Z}\}$, denn sei $m = nq + r$ mit $0 \leq r \leq n-1$, dann gilt $m \in \text{Kern}(f) \Leftrightarrow e = f(m) = a^m = (a^n)^q a^r = a^r \Leftrightarrow r = 0$.

   Damit lassen sich die Elemente von $\mathbb{Z}/N$ darstellen als $\{0 + n\mathbb{Z}, ..., (n-1) + n\mathbb{Z}\}$ und $\mathbb{Z}$ lässt sich zerlegen in die Partition $\mathbb{Z} = (0 + n\mathbb{Z}) \cup \cdots \cup ((n-1) + n\mathbb{Z})$. ◊

Bemerkung 2.44.

Sei $n \in \mathbb{N}$ gewählt. Wir schreiben $\overline{m} := m + n\mathbb{Z}$. Es gelten:

1. $\overline{m_1} + \overline{m_2} = \overline{m_1 + m_2}$, denn $(m_1 + n\mathbb{Z}) + (m_2 + n\mathbb{Z}) = (m_1 + m_2)n\mathbb{Z}$ per Definition der Addtition auf $\mathbb{Z}/n\mathbb{Z}$.
2. $\overline{m_1} = \overline{m_2} \Leftrightarrow \overline{m_1 - m_2} = 0 \Leftrightarrow m_1 - m_2 \in n\mathbb{Z} \Leftrightarrow n$ teilt $m_1 - m_2$. Wir schreiben $m_1 \equiv m_2$ oder auch $m_1 = m_2 \mod n$. m_1 und m_2 heißen kongruent.

Wir schreiben $\mathbb{Z}/n\mathbb{Z} = \{\overline{m} \mid m \in \mathbb{Z}\}$. ◊

Beispiel 2.45.

1. Sei $\sharp G = p$ prim. Sei $a \neq e$, dann gilt: $\mathrm{Ord}(a)$ teilt $p \Rightarrow \mathrm{Ord}(a) = p$, d.h. $G = \langle a \rangle \cong \mathbb{Z}/p\mathbb{Z}$. Insbesondere gibt es zu jeder Primzahl bis auf Isomorphie nur eine Gruppe dieser Ordnung.
2. $\sharp G = 4$. Dann können eintreten:
   a) Es gibt $a \neq e$ mit $\mathrm{Ord}(a) = 4$, d.h. $G = \langle a \rangle \cong \mathbb{Z}/4\mathbb{Z}$ ist zyklisch.
   b) Alle $a \neq e$ haben die Ordnung 2. Dann gilt $a^2 = e \Rightarrow a^{-1} = a$ für alle $a \in G$, d.h. jedes Element ist invers zu sich selbst. Sei $b \in G$ verschieden von a, e. Dann ist $ab \neq a, b, e$, d.h. $G = \{e, a, b, ab\}$ und es gilt $ab = (ab)^{-1} = b^{-1}a^{-1} = ba$, d.h. G ist Abelsch. G heißt die Kleinsche Vierergruppe.

   Insbesondere gibt es bis auf Isomorphie nur zwei Gruppen der Ordnung vier. ◊

Bemerkung 2.46. (Dimensionsformel)

Seien V ein K-Vektorraum und U ein Untervektorraum von V. Dann ist $(U, +)$ insbesondere eine Untergruppe von $(V, +)$, d.h. $(v_1 + U) + (v_2 + U)$ definiert eine Gruppenoperation auf $V/U = \{v + U \mid v \in V\}$ und $\phi : V \to V/U$ mit $\phi(v) * v + U$ ist ein surjektiver Gruppenhomomorphismus. Wir zeigen: V/U wird via $\alpha(v + U) := (\alpha v) + U$ zu einem Vektorraum und ϕ definiert einen Vektorraumhomomorphismus zwischen V und V/U.

1. Unabhängigkeit des Vertreters: gelte $v + U = v' + U$, dann ist $v - v' \in U$, also $\alpha(v - v') = \alpha v - \alpha v' \in U$ und damit $\alpha v + U = \alpha v' + U$.
2. ϕ ist K-linear, denn $\phi(\alpha v) = \alpha v + U = \alpha(v + U) = \alpha\phi(v)$.

Außerdem gelten $\mathrm{Kern}(\phi) = U$ und $\mathrm{Bild}(\phi) = V/U$. Damit erhalten wir eine Dimensionsformel für Quotientenvektorräume:

$$\dim V = \dim \mathrm{Kern}\phi + \dim \mathrm{Bild}\phi = \dim U + \dim V/U \qquad \Longrightarrow \qquad \dim V/U = \dim V - \dim U. \quad ◊$$

Bemerkung 2.47.

Eine Gruppe G nennen wir einfach, falls G nur die trivialen Normalteiler $\{e\}$ und G selbst hat.

Beispielsweise ist $\mathbb{Z}/n\mathbb{Z}$ genau dann einfach, wenn n eine Primzahl ist: $\mathbb{Z}/n\mathbb{Z}$ ist stets Abelsch, d.h. jede Untergruppe von $\mathbb{Z}/n\mathbb{Z}$ ist ein Normalteiler von $\mathbb{Z}/n\mathbb{Z}$. Nun hat $\mathbb{Z}/n\mathbb{Z}$ genau dann nichttriviale Untergruppen, wenn n nicht prim ist. ◊

2.6. Symmetrische Gruppen

Definition 2.48.

Sei X eine Menge. Dann heißt $S_X = \mathrm{Aut}(X) = \{f : X \to X \mid f \text{ bijektiv}\}$ die symmetrische Gruppe bzw. Permutationsgruppe der Menge X.

Beispiel 2.49.

Es bezeichne S_n die symmetrische Gruppe der Menge $X = \{1, ..., n\}$. Wir schreiben

$$f = \begin{pmatrix} a_1 & \cdots & a_n \\ b_1 & \cdots & b_n \end{pmatrix} \qquad \Longleftrightarrow \qquad f(a_1) = b_1,\ ...,\ f(a_n) = b_n.$$

1. Wegen $\sharp S_2 = \sharp\{\mathrm{id}, \left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)\} = 2$ prim ist $S_2 \cong \mathbb{Z}/2\mathbb{Z}$, insbesondere kommutativ.
2. $S_3 = \{\mathrm{id},\ \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{smallmatrix}\right),\ \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix}\right),\ ...\}$ ist nicht kommutativ:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

   Insbesondere ist zwar $\sharp S_3 = 3! = 6$, aber $S_3 \neq \mathbb{Z}/6\mathbb{Z}$. Damit kann es auch kein $a \in S_3$ geben mit $\mathrm{Ord}(a) = 6$, sonst $G = \langle a \rangle \cong \mathbb{Z}/6\mathbb{Z}$. ◊

Definition 2.50.

In S_n heißt $\sigma = (b_1, ..., b_m)$ ein Zyklus der Länge m, falls $\sigma(b_1) = b_2,\ \sigma(b_2) = b_3,\ ...,\ \sigma(b_m) = b_1$.

Ein Zyklus der Länge 2 heißt eine Transposition.

Beispiel 2.51.

Die Zyklen $(1,2,3), (2,3,1), (3,1,2)$ in S_3 sind per Definition identisch; ebenso $(1,3,2), (2,1,3), (3,2,1)$. Untergruppen von S_3 sind damit:

$$\{e\};\ \{e,(1,2)\},\ \{e,(1,3)\},\ \{e,(2,3)\};\ \{e,(1,2,3)\},\ \{e,(1,3,2)\}.$$

$\{e,(1,2)\}$ ist kein Normalteiler von S_3, denn

$$(1,3)\ \{e,(1,2)\}\ (1,3)^{-1} = (1,3)\ \{e,(1,2)\}\ (1,3) = \{e,(2,3)\}.$$

Analog sind auch $\{e,(1,3)\}$ und $\{e,(2,3)\}$ keine Normalteiler von S_3. Aber $U := \{e,(1,2,3),(1,3,2)\}$ ist Normalteiler von S_3: aUa^{-1} ist Untergruppe der Ordnung 3, da $x \mapsto axa^{-1}$ ein Automorphismus von S_3 ist. Also $aUa^{-1} = U$. ◊

Lemma 2.52.

Jedes $\sigma \in S_n$ ist ein Produkt von Transpositionen, d.h. $\sigma = \tau_1 \circ \cdots \circ \tau_k$ mit Transpositionen $\tau_i \in S_n$.

Beweis.

Für $\sigma = \mathrm{id}$ gilt: $\mathrm{id} = \tau \circ \tau$ für ein beliebiges $\tau \in S_n$. Sei nun $\sigma \neq \mathrm{id}$ und sei i_1 minimal mit $\sigma(i_1) \neq i_1$ ($\Rightarrow \sigma(i_1) > i_1$). Setze $\tau_1 = (i_1, \sigma(i_1))$ und $\sigma_1 = \tau_1 \circ \sigma$. Sei i_2 minimal mit $\sigma_1(i_1) \neq i_2$ ($\Rightarrow i_1 < i_2$). Setze $\tau_2 = (i_2, \sigma_1(i_2))$ und $\sigma_2 = \tau_2 \circ \sigma_1 = \tau_2 \circ \tau_1 \circ \sigma$ u.s.w.. Schließlich ist $\sigma_k = \tau_k \circ \cdots \circ \tau_1 \circ \sigma = \mathrm{id}$, d.h. $\sigma = \tau_1^{-1} \circ \cdots \circ \tau_k^{-1} = \tau_1 \circ \cdots \circ \tau_k$. ◊

Satz 2.53.

Sei $\tau_0 = (1,2)$. Zu jeder Transposition $\tau \in S_n$ gibt es dann ein $\sigma \in S_n$ mit $\tau = \sigma \circ \tau_0 \circ \sigma^{-1}$.

Wir sagen dann, τ ist konjugiert zu τ_0.

Beweis.

Sei $\tau = (l,k)$ mit $l \neq k$. Setze σ so, dass $\sigma(1) = k$ und $\sigma(2) = l$. Dann gelten für k, l und alle $j \neq k, l$:

$$(\sigma \circ \tau \circ \sigma^{-1})(k) = l,\ \text{denn } k \mapsto 1 \mapsto 2 \mapsto l,$$
$$(\sigma \circ \tau \circ \sigma^{-1})(l) = k,\ \text{denn } l \mapsto 2 \mapsto 1 \mapsto k,$$
$$(\sigma \circ \tau \circ \sigma^{-1})(j) = j,\ \text{denn } j \mapsto \sigma^{-1}(j) \mapsto \sigma^{-1}(j) \mapsto j.$$ □

2.7. Fehlstände und Signaturen

Definition 2.54.

Ein Paar (i,j) heißt Fehlstand von $\sigma \in S_n$, falls $i < j$ und $\sigma(i) > \sigma(j)$ gelten.

Die Signatur eines Zyklus in S_n ist definiert als

$$\mathrm{sgn}(\sigma) = \begin{cases} +1 & \text{falls } \sigma \text{ eine gerade Anzahl von Fehlständen hat} \\ -1 & \text{falls } \sigma \text{ eine ungerade Anzahl von Fehlständen hat} \end{cases}.$$

Bemerkung 2.55.

Es gilt $\mathrm{sgn}(\tau_0) = -1$. Wir werden zeigen, dass sogar für jede Transposition τ in S_n gilt $\mathrm{sgn}(\tau) = -1$. ◊

Lemma 2.56.

Die Signaturabbildung sgn besitzt die Repräsentation

$$\operatorname{sgn}(\sigma) = \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Beweis.

Bezeichne m die Anzahl der Fehlstände von σ. Dann gilt:

$$\begin{aligned}\prod_{i<j}(\sigma(j) - \sigma(i)) &= \prod_{\substack{i<j \\ \sigma(i)<\sigma(j)}} (\sigma(j) - \sigma(i)) \prod_{\substack{i<j \\ \sigma(i)>\sigma(j)}} |\sigma(j) - \sigma(i)| \cdot (-1)^m \\ &= (-1)^m \cdot \prod_{i<j} |\sigma(j) - \sigma(i)| \quad = \quad (-1)^m \cdot \prod_{i<j}(j - i),\end{aligned}$$

also ist $\operatorname{sgn}(\sigma) = (-1)^m = \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i}$. □

Satz 2.57.

Es gilt $\operatorname{sgn}(\tau \circ \sigma) = \operatorname{sgn}(\tau)\operatorname{sgn}(\sigma)$, d.h. $\operatorname{sgn} : S_n \to \{\pm 1\}$ ist ein Gruppenhomomorphismus.

Beweis.

Nach der Darstellungsformel gilt

$$\operatorname{sgn}(\tau \circ \sigma) = \prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{j - i} = \prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \frac{\sigma(j) - \sigma(i)}{j - i} = \prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \operatorname{sgn}(\sigma).$$

Weiter ist

$$\begin{aligned}\prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{j - i} &= \prod_{\substack{i<j \\ \sigma(i)<\sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \prod_{\substack{i<j \\ \sigma(i)>\sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \\ &= \prod_{\substack{i<j \\ \sigma(i)<\sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \prod_{\substack{i \leftrightarrow j \\ i<j \\ \sigma(i)<\sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \\ &= \prod_{\sigma(i)<\sigma(j)} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \quad = \quad \prod_{i<j} \frac{\tau(j) - \tau(i)}{j - i} \quad = \quad \operatorname{sgn}(\tau).\end{aligned}$$

□

Korollar 2.58.

Jede Transposition $\tau \in S_n$ besitzt eine negative Signatur: $\operatorname{sgn}(\tau) = -1$.

Beweis.

Es gilt $\tau = \sigma \circ \tau_0 \circ \sigma^{-1}$ für einen geeigneten Zyklus $\sigma \in S_n$, also

$$\operatorname{sgn}(\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau_0)\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau_0)\operatorname{sgn}(\sigma)^{-1} = -1.$$

□

Definition 2.59.

$A_n = \text{Kern}(\text{sgn}) = \{\sigma \in S_n \mid \text{sgn}(\sigma) = +1\}$ heißt die alternierende Gruppe auf $\{1, ..., n\}$.

Bemerkung 2.60.

1. Es gilt $\{\pm 1\} \cong \frac{S_n}{A_n}$, d.h. $\sharp S_n = \sharp A_n \cdot \sharp \frac{S_n}{A_n} = 2\sharp A_n$.
2. A_n ist für $n \geq 5$ einfach; Beweis im nächsten Semester in der Vorlesung Algebra. ◊

Satz 2.61.

Seien K Körper und $e^{(i)} \in K^{1\times n}$ der i-te kanonische Basisvektor des $K^{1\times n}$. Dann gilt für alle $\sigma \in S_n$:

$$\det \begin{pmatrix} e^{\sigma(1)} \\ \cdots \\ e^{\sigma(n)} \end{pmatrix} = \text{sgn}(\sigma)$$

Beweis. (formal per Induktion)

Sei $\sigma = \tau_1 \circ \cdots \circ \tau_k$ eine Zerlegung von σ in Transpositionen und sei $\sigma_1 = \tau_2 \circ \cdots \circ \sigma_k$. Dann gilt:

$$\det \begin{pmatrix} e^{\sigma(1)} \\ \cdots \\ e^{\sigma(n)} \end{pmatrix} = \det \begin{pmatrix} e^{\tau_1\circ\sigma_1(1)} \\ \cdots \\ e^{\tau_1\circ\sigma_1(n)} \end{pmatrix} = (-1) \det \begin{pmatrix} e^{\sigma_1(1)} \\ \cdots \\ e^{\sigma_1(n)} \end{pmatrix} = \cdots$$
$$= (-1)^k \det \begin{pmatrix} e^{(1)} \\ \cdots \\ e^{(n)} \end{pmatrix} = (-1)^k = \text{sgn}(\sigma). \qquad \square$$

Satz 2.62. (Regel von Leibnitz)

Für eine Matrix $A \in \mathfrak{M}_K(n, n)$ mit den Zeilen $A_i = (a_{i1}, \ ..., \ a_{in})$ gilt:

$$\det A = \sum_{\sigma\in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \cdot \text{sgn}(\sigma).$$

Beweis. (formal per Induktion)

Nach der Multilinearität der Determinantenabbildung gilt

$$\det A = \det \begin{pmatrix} A_1 \\ A_2 \\ \cdots \\ A_n \end{pmatrix} = \sum_{j_1=1}^{n} a_{1,j_1} \det \begin{pmatrix} e^{(j_1)} \\ A_2 \\ \cdots \\ A^n \end{pmatrix} = \sum_{j_1,\ldots,j_n=1}^{n} a_{1,j_1} \cdots a_{n,j_n} \det \begin{pmatrix} e^{(j_1)} \\ e^{(j_2)} \\ \cdots \\ e^{(j_n)} \end{pmatrix}.$$

Wir definieren $j : \{1, , ..., n\} \to \{1, ..., n\}$ mit $j(i) = j_i$. Falls nun j nicht injektiv ist, dann gibt es $i_1, i_2 \in \{1, ..., n\}$ mit $j_{i1} = j_{i2}$ und damit gilt $\det(e^{(j_1)}; \ ...; \ e^{(j_n)}) = 0$, also auch $\det A = 0$. Andernfalls ist j ein Zyklus und die verbleibenden Terme summieren sich auf zu

$$\det A = \sum_{\sigma\in S_n} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \det \begin{pmatrix} e^{\sigma(1)} \\ \cdots \\ e^{\sigma(n)} \end{pmatrix} = \sum_{\sigma\in S_n} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \cdot \text{sgn}(\sigma). \qquad \square$$

3. Ringtheorie

3.1. Ringe und Ideale

Definition 3.1.

Das Tripel $(A, +, \cdot)$, bestehend aus einer Menge A und zwei Operationen $+, \cdot$, ist ein Ring, falls gelten:

1. $(A, +)$ ist eine Abelsche Gruppe. $+$ heißt die Addition auf A.
2. $(A, \cdot)$ ist ein Monoid. $\cdot$ heißt die Multiplikation auf A.
3. Für alle $a, b, c \in A$ gelten: $a(b + c) = ab + ac$ und $(a + b)c = ac + bc$ (Distributivgesetze).

$(A, +, \cdot)$ heißt kommutativ, falls $\cdot$ kommutativ ist, d.h. falls für alle $a, b \in A$ gilt $ab = ba$.

$(A, +, \cdot)$ heißt ein Schiefkörper, falls $(A\backslash\{0\}, \cdot)$ eine Gruppe ist. Ein neutrales Element 1 von $\cdot$ heißt Einselement. Ein kommutativer Schiefkörper ist ein Körper.

Bemerkung 3.2.

1. Es gilt: $(a_1 + \cdots + a_r)(b_1 + \cdots + b_s) = a_1b_1 + a_1b_2 + \cdots + a_rb_s = \sum_{i=1}^{r} \sum_{j=1}^{s} a_ib_j$.
2. Für alle $a \in A$ gilt $a \cdot 0 = 0$, denn $0 + 0 = 0$, also $a \cdot (0 + 0) = a \cdot 0$, d.h. $a \cdot 0 + a \cdot 0 = a \cdot 0$ und damit $0 = a \cdot 0$. Analog ist $0 = 0 \cdot a$.
3. $A = \{0\}$ ist nach dieser Definition ein Ring mit Einselement 0. Ab jetzt fordern wir allerdings stets $1 \neq 0$ in A, falls A ein Einselement besitzt. ◊

Definition 3.3.

Seien $A = (A, +, \cdot)$ ein Ring und $B \subseteq A$. Dann heißt B ein Unterring von A, falls gelten:

1. B ist abgeschlossen unter $+$ und $\cdot$, d.h. aus $x, y \in B$ folgen $x + y \in B$ und $x \cdot y \in B$.
2. B enthält das neutrale Element der Addition: $0 \in B$.
3. B ist abgeschlossen unter additiv inversen Elementen, d.h. mit $x \in B$ ist stets auch $-x \in B$.

Beispiel 3.4.

1. Sei A ein Ring. Dann ist $\mathfrak{M}_A(n, n)$, die Menge der $(n \times n)$-Matrizen mit Einträgen aus A, ein Ring. $\mathfrak{M}_A(n, n)$ ist kein kommutativer Ring für $n \geq 2$. Hat A ein Einselement, so ist die Einheitsmatrix Id_n das Einselement von $\mathfrak{M}_A(n, n)$.
2. Die Menge $\mathrm{End}_K(V)$ der Endomorphismen auf einem Vektorraum V mit Addition $+$ und Multiplikation $\circ$ ist ein nicht kommutativer Ring mit Einselement id für $\dim_K(V) \geq 2$.
3. Seien M eine Menge und $\mathcal{P}(M) = \{B \mid B \subseteq M\}$ die Potenzmenge von M. Dann ist $(\mathcal{P}(M), \triangle, \cap\}$ ein kommutativer Ring mit Einselement M, wobei die symmetrische Differenz $\triangle$ definiert ist als $A \triangle B = (A \cup B)\backslash(A \cap B) = (A\backslash B) \cup (B\backslash A)$. Dabei ist $\emptyset$ additiv neutral: $A \triangle \emptyset = A$; außerdem ist jedes Element zu sich selbst additiv invers: $A \triangle A = \emptyset$. ◊

Definition 3.5.

Sei A ein Ring mit Eins. Die invertierbaren Elemente des Monoids $(A, \cdot)$ heißen Einheiten von A.

Bemerkung 3.6.

1. Die Menge $A^\times$ der Einheiten von A bildet eine multiplikative Gruppe, die Einheitengruppe von A.
2. Es gelten beispielsweise $\mathbb{Z}^\times = \{-1, 1\}$ und $\mathbb{Q}^\times = \mathbb{Q}\backslash\{0\}$.
3. $A^\times = A\backslash\{0\}$ gilt genau dann, wenn A ein Schiefkörper ist. ◊

Definition 3.7.

Sei I eine Untergruppe von $(A,+)$. Gilt $AI \subseteq I$, dann heißt I ein Linksideal in A.

Gilt $IA \subseteq I$, dann heißt I ein Rechtsideal in A.

Ist I sowohl Links- als auch Rechtsideal in A, dann heißt I zweiseitiges Ideal oder auch einfach nur Ideal in A.

Bemerkung 3.8.

1. Ist I ein zweiseitiges Ideal von A, so definiert $(a+I)(b+I) := ab+I$ eine Multiplikation auf $A+I$. Wieder ist die Unabhängigkeit vom Vertreter nachzuweisen. Seien hierfür $a, a', b, b' \in A$ mit $a + I = a' + I$ und $b + I = b' + I$. Dann liegen $a - a'$ und $b - b'$ in I, d.h. $(ab - a'b') = (a - a')b + a'(b - b') \in I$, da I zweiseitiges Ideal ist. Also ist $ab + I = a'b' + I$.
2. $A + I$ mit dieser Multiplikation und der üblichen Addition $(a + I) + (b + I) := (a + b) + I$ definiert einen Ring. Eigenschaften wie Existenz von Eins und Einheiten oder Kommutativität vererben sich von A auf $A + I$.
3. Ideale in $\mathbb{Z}$ sind von der Gestalt $n\mathbb{Z} = \{nm \mid m \in \mathbb{Z}\}$, da Ideale auch Untergruppen sind. ◊

3.2. Ringhomomorphismen und Quotientenringe

Definition 3.9.

Seien A, B Ringe. $f : A \to B$ heißt ein Ringhomomorphismus, falls für alle $a, b \in A$ gilt:

$$f(a + b) = f(a) + f(b) \qquad \text{und} \qquad f(ab) = f(a)f(b).$$

Sei f ein Ringhomomorphismus. Dann nennen wir f einen Ringmonomorphismus, falls f injektiv ist, und einen Ringepimorphismus, falls f surjektiv ist. Ist f beides, dann heißt f Ringisomorphismus.

Ist f eine Selbstabbildung, d.h. gilt $A = B$, dann heißt f ein Ringendomorphismus; ist f zusätzlich bijektiv, dann heißt f ein Ringautomorphismus.

$f^{-1}(\{0\}) \subseteq A$ heißt der Kern von f und $f(A) \subseteq B$ das Bild von A.

Bemerkung 3.10.

1. Wieder gelten: $f(0) = 0$ und $f(-x) = -f(x)$. Ist x eine Einheit, dann ist $f(x^{-1}) = f(x)^{-1}$.
2. Genau dann ist f injektiv, wenn $\mathrm{Kern}(f) = \{0\}$ ist. $\mathrm{Kern}(f)$ ist stets ein beidseitiges Ideal. ◊

Definition 3.11.

Sei I ein Ideal von A. $A/I := A + I$ heißt der Restklassenring oder Quotientenring von A nach I.

$\phi : I \mapsto A/I$ mit $\phi(a) = a + I$ heißt der kanonische Homomorphismus oder die Restklassenabbildung.

Bemerkung 3.12.

1. ϕ definiert einen Ringhomomorphismus: $\phi(ab) = ab + I = (a + I)(b + I) = \phi(a)\phi(b)$.
2. Wegen $\mathrm{Kern}(\phi) = I$. Sind die zweiseitigen Ideale von A genau die Kerne von Homomorphismen von A irgendwohin.
3. Sei $f : A \to B$ ein Ringepimorphismus, dann gilt: Ist A kommutativ, so auch B: Seien $b_1, b_2 \in B$ dann gibt es $a_1, a_2 \in A$ mit $f(a_1) = b_1$, $f(a_2) = b_2$ und es folgt

$$b_1b_2 = f(a_1)f(a_2) = f(a_1a_2) = f(a_2a_1) = f(a_2)f(a_1) = b_2b_1.$$

4. Jeder Ringhomomorphismus $f : A \to B$ lässt sich auf einen Ringisomorphismus $\bar{f} : A/I \to f(A)$ mit $I = \text{Kern}(f)$ unf $\bar{f}(a+I) = f(a)$ einschränken. Das Diagramm

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \downarrow{\scriptstyle \phi} & & \uparrow{\scriptstyle \text{id}} \\ A/I & \xrightarrow{\bar{f}} & f(A) \end{array}$$

kommutiert, d.h. es gilt $f = \bar{f} \circ \phi$: ◊

Satz 3.13. (Homomorphiesatz)

Ist $f : A \to B$ ein Ringhomomorphismus, dann gilt $f(A) \cong A/\text{Kern}(f)$.

Beweis.

Setze $I = \text{Kern}(f)$ und $\bar{f} : A/I \to f(A)$, $\bar{f}(a+I) = f(a)$. Es ist nur noch zu zeigen, dass für alle $a, b \in I$ gilt $\bar{f}((a+I)(b+I)) = \bar{f}(a+I)\bar{f}(b+I)$:

$$\bar{f}((a+I)(b+I)) = \bar{f}((ab)+I) = f(ab) = f(a)f(b) = \bar{f}(a+I)\bar{f}(b+I). \qquad \square$$

Bemerkung 3.14.

1. Sei I ein Ideal in A. Dann gilt $a + I = b + I \Leftrightarrow a - b \in I \Leftrightarrow a \equiv b \bmod I \Leftrightarrow \overline{a} = \overline{b}$.
2. Gelte $a \equiv b \bmod I$ und $c \equiv d \bmod I$, dann ist $ac \equiv bd \bmod I$ und $a + c \equiv b + d \bmod I$, denn nach den Rechengesetzen im Ring A/I folgen aus $\overline{a} = \overline{b}$ und $\overline{c} = \overline{d}$, dass $\overline{a} + \overline{c} = \overline{b} + \overline{d}$ und $\overline{ac} = \overline{bd}$.
3. Rechenbeispiel: Wir zeigen, dass die Zahl $2^{32}+1$ keine Primzahl ist; genauer, dass 641 diese Zahl teilt. Es ist $641 = 2^7 \cdot 5 + 1 = 5^4 + 2^4$. Also gilt $2^7 \cdot 5 \equiv -1 \bmod 641$. Setze $I = 641\mathbb{Z}$. Dann gelten $\overline{2^7} \cdot \overline{5} = \overline{-1}$ und $\overline{5^4} = \overline{-2^4}$. Damit ist

$$\overline{(2^7 \cdot 5)^4} = \overline{2^{28}} \cdot \overline{5^4} = \overline{(-1)^4} = \overline{1} = \overline{-2^{32}} \quad \Longrightarrow \quad \overline{2^{32}} = \overline{-1} \quad \Longrightarrow \quad \overline{2^{32}+1} = \overline{0}.$$

Also gilt: 641 teilt $2^{32} + 1$. ◊

3.3. Teilbarkeit in Integritätsbereichen

Bemerkung 3.15.

1. In Schiefkörpern gilt: $ab = 0 \Rightarrow a = 0$ oder $b = 0$, d.h. Schiefkörper und insbesondere Körper sind nullteilerfrei, denn angenommen, es ist $a \neq 0$, dann ist $a^{-1}(ab) = a^{-1}0 = 0$, d.h. $b = 0$. Ist dagegen $b \neq 0$, dann $(ab)b^{-1} = 0b^{-1} = 0$, d.h. $a = 0$.
2. Sei $A = \mathbb{Z}/nm\mathbb{Z}$ mit $n, m > 1$. Dann gilt zwar $nm \equiv 0 \bmod nm\mathbb{Z}$, aber es ist weder $n \equiv 0$ noch $m \equiv 0 \bmod nm\mathbb{Z}$, d.h. $\mathbb{Z}/nm\mathbb{Z}$ ist nicht nullteilerfrei. Andererseits ist $\mathbb{Z}/p\mathbb{Z}$ für jede Primzahl p nullteilerfrei. ◊

Definition 3.16.

Ein $a \in A$ heißt ein Linksnullteiler, falls $a \neq 0$ und $ba = 0$ für ein $b \neq 0$. b ist dann ein Rechtsnullteiler.

Ein Integritätsbereich ist ein nullteilerfreier, kommutativer Ring mit Eins.

Sei R Integritätsbereich mit $a, b \in R$. a teilt b in R (in Zeichen: $a \mid b$), falls es ein $c \in R$ gibt mit $ac = b$. a heißt assoziiert zu b (in Zeichen: $a \sim b$), falls es eine Einheit $c \in R^\times$ gibt mit $ac = b$.

a heißt ein echter Teiler von b, falls $a \mid b$ und $b \nmid a$. $a \neq 0$ heißt irreduzibel oder unzerlegbar, falls $a \notin \mathbb{R}^\times$ und für $a = bc$ immer $b \in R^\times$ oder $c \in R^\times$ erfüllt ist. Andernfalls heißt a reduzibel oder zerlegbar.

$a_0, a_1, ..., a_n, ...$, heißt eine absteigende Teilerkette, falls für alle i gilt: a_{i+1} ist ein echter Teiler von a_i.

Bemerkung 3.17.

1. $\mathbb{Z}$ ist ein Integritätsbereich, ebenso alle Körper K und deren Polynomringe $K[X]$.
2. In einem Integritätsbereich R gelten $a \mid a$; $(a \mid b \;\&\; b \mid c) \Rightarrow a \mid c$ und $(a \mid b \;\&\; b \mid a) \Leftrightarrow a \sim b$: Gelten zunächst $a \mid b$ und $b \mid a$, dann sind $ac = b$ und $bd = a$ für gewisse $c, d \in R$, also $bdc = b$ bzw. $b(dc-1) = 0$. Wegen der Nullteilerfreiheit ist dann $b = 0$, d.h. auch $a = 0$ und damit $a \sim b$, oder aber $dc - 1 = 0$, d.h. $d, c \in R^\times$ sind Einheiten und wiederum $a \sim b$. Gilt umgekehrt $a \sim b$, d.h. $ac = b$ für ein $c \in R^\times$, dann $bc^{-1} = a$.
3. Weiter gelten: $a \sim a$, $a \sim b \Rightarrow b \sim a$ und $(a \sim b \;\&\; b \sim c) \Rightarrow a \sim c$, d.h. $\sim$ ist eine Äquivalenzrelation.
4. Es sind stets $a \mid 0$, $a \mid 1$ genau dann, wenn $a \in R^\times$ und $0 \mid a$ genau dann, wenn $a = 0$. ◊

Satz 3.18.

Hat ein Integritätsbereich keine unendliche Teilerkette, so ist jedes von Null verschiedene Element eine Einheit oder ein Produkt irreduzibler Elemente.

Beweis.

Sei $a \notin R^\times$, $a \neq 0$. Dann ist a irreduzibel oder besitzt eine Darstellung $a = bc$. b, c selbst sind irreduzibel besitzen Zerlegungen u.s.w.. Dies muss abbrechen, sonst hätte R eine unendliche Teilerkette. □

Definition 3.19.

Sei R Integritätsbereich. Ein $p \in R \backslash R^\times$, $p \neq 0$, heißt Primelement, falls $p \mid ab \Rightarrow p \mid a$ oder $p \mid b$.

Bemerkung 3.20.

1. Die Primelemente von $\mathbb{Z}$ werden auch als Primzahlen bezeichnet.
2. Jedes Primelement p ist irreduzibel: Gelte $p = ab$. Zu zeigen: $a \in R^\times$ oder $b \in R^\times$. Aus $p \mid ab$ folgt Œ $p \mid a$, d.h. $pc = a$, also $p = pcb$ und damit $1 = cb$, d.h. $b \in R^\times$.

Definition 3.21.

Sei R ein Integritätsbereich, in dem alle irreduziblen Elemente prim sind und in dem jede von Null verschiedene Nichteinheit bis auf Reihenfolge und Einheiten eindeutiges Produkt irreduzibler Elemente ist. Dann heißt R ein faktorieller Ring oder ZPE-Ring („Zerlegung in Primelemente eindeutig“).

Seien R faktoriell und $a, b \in R \backslash R^\times$, $a, b \neq 0$, $a = e_1 \prod_{i=1}^{n} p_i^{\nu_i}$ und $b = e_2 \prod_{i=1}^{n} p_i^{\mu_i}$. Wir setzen

$$\mathrm{ggT}(a,b) = \prod_{i=1}^{n} p_i^{\min\{\nu_i,\mu_i\}}, \qquad \mathrm{kgV}(a,b) = \prod_{i=1}^{n} p_i^{\max\{\nu_i,\mu_i\}}.$$

ggT heißt der größte gemeinsame Teiler und kgV das kleinste gemeinsame Vielfaches von a und b.

a und b heißen teilerfremd, falls $\mathrm{ggT}(a,b) = 1$.

3.4. Teilbarkeit in Hauptidealringen

Definition 3.22.

Seien A ein kommutativer Ring und $a \in A$. Für $a \in R$ heißt $Ra = \{ba \mid b \in R\}$ ein Hauptideal in R.

A heißt ein Hauptidealring, falls zu jedem Ideal I in A ein $a \in A$ existiert mit $I = Ra$.

Sind $a_1, ..., a_n \in A$, dann heißt $Aa_1 + \cdots + Aa_n$ das von $a_1, ..., a_n$ erzeugte Ideal.

Bemerkung 3.23.

1. Insbesondere sind Hauptideale die von einem Element erzeugten Ideale.
2. Besitzt A ein Einselement 1, dann ist A stets ein Hauptideal, da A von 1 erzeugt wird.
3. Sei R ein Integritätsbereich. Dann gelten: $a \mid b$ gdw. $b \in Ra$ gdw. $Rb \subseteq Ra$ und $a \sim b$ gdw. $Rb \subseteq Ra$ & $Ra \subseteq Rb$ gdw. $Ra = Rb$. ◊

Definition 3.24.

Seien R ein Integritätsbereich und $w : R\backslash\{0\} \to \mathbb{N}$ mit

1. Aus $a \mid b$, $b \neq 0$ folgt $w(a) \leq w(b)$,
2. Zu jedem $b \neq 0$ gibt es $q, r \in R$ mit $a = bq + r$ und $r = 0$ oder $w(r) < w(b)$.

Dann heißt w eine euklidische Wertefunktion.

Bemerkung 3.25.

1. $\mathbb{Z}$ und $K[X]$ für einen beliebigen Körper K besitzen eine euklidische Wertefunktion: In $\mathbb{Z}$ definiere $w(m) = |m|$, in $K[X]$ setze $w(p(x)) = \deg p$.
2. $\mathbb{Z}$ und $K[X]$ sind Hauptidealringe und besitzen keine unendlichen Teilerketten: ◊

Satz 3.26.

1. Jeder Integritätsbereich mit euklidischer Wertefunktion ist ein Hauptidealring.
2. Kein Integritätsbereich, der ein Hauptidealring ist, hat eine unendliche Teilerkette.

Beweis.

1. Sei $I \subseteq R$ ein Ideal. Wir suchen ein $a \in R$ mit $I = Ra$. Wähle $a \in I$ mit $w(a)$ minimal (also $a \neq 0$). Dann ist zunächst $Ra \subseteq I$. Sei $b \in I$, $b \neq 0$, $b = ca + r$ mit $w(r) < w(a)$ oder $r = 0$. Im Fall $r = 0$ folgt $b = ca$, also $b \in Ra$; sonst $r = b - ca \in I$. Dann gilt $w(r) < w(a)$, was im Widerspruch zur Minimalität von $a \in I$ bzgl. w steht. Damit gilt auch $I \subseteq Ra$.
2. Sei $a_0, a_1, ...,$ eine unendliche Teilerkette, d.h. a_{i+1} ist ein echter Teiler von a_i. Damit sind $Ra_i \subseteq Ra_{i+1}$ und $Ra_{i+1} \nsubseteq Ra_i$, d.h. $Ra_0 \subsetneq Ra_1 \subsetneq \cdots \subsetneq Ra_i \subsetneq Ra_{i+1} \subsetneq \cdots$ ... ist eine echt aufsteigende Kette von Idealen. Wir definieren $I = \bigcup\{Ra_i \mid i \in \mathbb{N}_0\}$. I ist ein Ideal in R, also abgeschlossen bzgl. $+$ und $\cdot$ mit Elementen aus R). Nach Voraussetzung ist $I = Ra$ ein Hauptideal. Wegen $a \in I$ ist $a \in Ra_n$ für ein gewisses $n \in \mathbb{N}_0$. Dann $a_{n+1} \in I = Ra \subseteq Ra_n$, d.h. $a_n \mid a_{n+1}$, ein Widerspruch. □

Lemma 3.27.

Sei I ein Ideal des kommutativen Ringes R mit Eins. Dann gilt:

$$I \subsetneq R \text{ ist maximal} \quad \Longleftrightarrow \quad R/I \text{ ist ein Körper.}$$

Beweis.

1. Sei $I \subseteq R$ ein maximales Ideal, d.h. R ist das einzige Ideal J mit $I \subseteq J$ in R. Sei $\overline{a} \in R/I$ mit $\overline{a} \neq \overline{0}$, d.h. $a \notin I$. Betrachte das Ideal $I + Ra$.Aus $a \in (I + Ra)\backslash I$ folgt $I \subsetneq I + Ra$, also $I + Ra = R$, d.h. $1 = b + ca$ für gewisse $b \in I$, $c \in R$. Somit ist $\overline{1} = \overline{0} + \overline{ca}$, d.h. $\overline{c}$ ist invers zu $\overline{a}$. Also ist jedes von Null verschiedene Element in R/I eine Einheit und R/I somit ein Körper.
2. Umgekehrt sei J ein Ideal in R mit $I \subsetneq J$. Wir zeigen, dass $J = R$. Wähle hierfür $a \in J\backslash I$. Dann gibt es ein $b \in R$ mit $\overline{1} = \overline{a}\overline{b} = \overline{ab}$, d.h. $1 - ab \in I$. Da $I \subsetneq J$, liegen sowohl ab als auch $1 - ab$ in J, also auch $1 \in J$ und damit $J = R$, d.h. I ist maximal. □

Lemma 3.28.

Sind $a, b \in \mathbb{Z}$ teilerfremd, so gibt es $x, y \in \mathbb{Z}$ mit $ax + by = 1$.

Beweis.

Da $\mathbb{Z}$ Hauptidealring ist, gilt $I = \mathbb{Z}a + \mathbb{Z}b = \mathbb{Z}c$ für ein $c \in \mathbb{Z}$, also $x'a + y'b = c$. $\mathbb{Z}a \subseteq \mathbb{Z}c$ impliziert $c \mid a$ und aus $\mathbb{Z}b \subseteq \mathbb{Z}c$ folgt $c \mid b$. Da a, b teilerfremd sind, ist damit $c \in \mathbb{Z}^\times = \{\pm 1\}$. Also $x'a + y'b = \pm 1$. □

Satz 3.29.

Für $n > 1$ gilt: $\mathbb{Z}/n\mathbb{Z}$ ist nullteilerfrei $\Leftrightarrow$ $\mathbb{Z}/n\mathbb{Z}$ ist ein Körper $\Leftrightarrow$ n ist eine Primzahl.

Beweis.

Es ist nur noch zu zeigen, dass $\mathbb{Z}/n\mathbb{Z}$ für n prim stets ein Körper ist. Sei $a \in \mathbb{Z}/n\mathbb{Z}$ nicht das Nullelement, d.h. $a \notin n\mathbb{Z}$. Dann gibt es nach dem letzten Lemma $x, y \in \mathbb{Z}$ mit $ax - pn = 1$, d.h. $ax - 1 \in n\mathbb{Z}$ bzw. $ax \equiv 1 \bmod n$ und somit ist a eine Einheit in $\mathbb{Z}/n\mathbb{Z}$. □

Satz 3.30.

Sei der Integritätsbereich R ein Hauptidealring. Dann gelten:

1. Jedes irreduzible Element in R ist prim.
2. Jede von Null verschiedene Nichteinheit in R ist (bis auf Reihenfolge und Einheiten) eindeutiges Produkt von irreduziblen Elementen.

Insbesondere ist jeder Hauptidealring ein faktorieller Ring.

Beweis.

1. Sei $a \in R$ irreduzibel, dann ist $Ra \subsetneq Rb \subsetneq R$ nicht möglich, sonst wäre b ein echter Teiler von a. Also ist $Ra \subsetneq I \subsetneq R$ für kein Ideal I möglich, da I stets ein Hauptideal ist. Damit ist Ra ein maximales Ideal. Somit ist R/Ra ein Körper. Schreibe $\overline{x}$ für $x + Ra \in R/Ra$. Beachte: $\overline{x} = \overline{0}$ gilt genau dann, wenn $x \in Ra$ bzw. $a \mid x$. Gelte nun $a \mid xy$, dann $\overline{xy} = \overline{0}$, d.h. $\overline{x} = \overline{0}$ oder $\overline{y} = \overline{0}$ bzw. $a \mid x$ oder $a \mid y$. Damit ist a ein Primelement.
2. Sei $a \in R$ nicht die Null und keine Einheit. Besitze a die Zerlegungen $a = p_1 \cdots p_n = q_1 \cdots q_m$ mit irreduziblen Elementen p_i, q_i, Œ $n \leq m$. Aus $p_1 \mid a$ folgt $p_1 \mid q_1 \cdots q_m$, da die q_i prim sind, also etwa $p_1 \mid q_1$. Da q_1 irreduzibel, folgt $p_1 \sim q_1$, d.h. $q_1 = e_1 \cdot p_1$ für eine Einheit e_1, also $p_1 \cdots p_n = e_1 \cdot p_1 \cdot q_2 \cdots q_m$ und ergo $p_2 \cdots p_n = e_1 \cdot q_2 \cdots q_m$ u.s.w. bis $1 = e_1 \cdot e_2 \cdots e_n \cdot q_n + 1 \cdots q_m$, d.h. $n = m$ und $q_i = e_i \cdot p_i$. □

Bemerkung 3.31.

Nicht jeder faktorielle Ring ist ein Hauptidealring: Die Ringe $\mathbb{R}[X, Y] = \mathbb{R}[X][Y]$ und $\mathbb{Z}[X]$ besitzen eine euklidische Wertefunktion, aber nicht jedes ihrer Ideale lässt sich von einem einzigen Element erzeugen. ◊

4. Modultheorie

4.1. Moduln

Definition 4.1.

Sei $(A, +, \cdot)$ ein kommutativer Ring mit Eins. Ein Modul M über A ist eine Abelsche Gruppe $(M, \oplus)$ mit einer Operation $\otimes : A \times M \to M$, so dass für alle $a, b \in A$ und alle $x, y \in M$ gelten:

$$a \otimes (x \oplus y) = (a \otimes x) \oplus (a \otimes y); \qquad (a \cdot b) \otimes x = a \otimes (b \otimes x);$$
$$(a + b) \otimes x = (a \otimes x) \oplus (b \otimes x); \qquad 1 \otimes x = x.$$

Beispiel 4.2.

1. Jeder kommutative Ring $(A, +, \cdot)$ mit Eins wird via $x \oplus y := x \times y$ und $a \otimes x := a \cdot x$ zu einem A-Modul.
2. Jeder Vektorraum $(V, \oplus, \otimes)$ über einem Körper $(K, +, \cdot)$ ist ein K-Modul.
3. Sei $(G, \oplus)$ eine Abelsche Gruppe. Dann wird G via $a \otimes y := n \cdot x = x + \cdots + x$ zu einem $\mathbb{Z}$-Modul.
4. Seien $(V, +, \cdot)$ ein K-Vektorraum und $f \in \mathrm{End}_K(V)$. Sei $K[f] = \{p(f) \mid p(X) \in K[X]\}$. Dann wird V via $x \oplus y := x + y$ und $a \otimes x := p(f)(x)$ zu einem $K[f]$-Modul. Dabei ist $p(f) = a_0\mathrm{id} + a_1 f + a_1 f \circ f + \cdots$.

   V kann auch als $K[X]$-Modul aufgefasst werden via $p \otimes x := p(f)(x)$. Formal wird hierfür der Einsetzungshomomorphismus $\mu : K[X] \to K[f]$ mit $p \mapsto p(f)$ benötigt. In dem Fall ist V also ein Modul über einem Hauptidealring. Solche Moduln sind für uns später von besonderem Interesse. ◊

Bemerkung 4.3.

$\varphi_f : K[X] \to \mathrm{End}_K(V)$ mit $\varphi_f(p) = p(f)$ ist ein Ringhomomorphismus:

$$\begin{aligned} \varphi_f(p+q)(x) &= (p+q)(f)(x) = p(f)(x) + q(f)(x) = \varphi_f(p)(x) + \varphi_f(q)(x), \\ \varphi_f(p \cdot q)(x) &= (p \cdot q)(f)(x) = p(f)(x) \cdot q(f)(x) = \varphi_f(p)(x) \cdot \varphi_f(q)(x), \\ \varphi_f(0)(x) &= 0(x). \end{aligned}$$

Zusätzlich gilt hier $\varphi_f(1) = \mathrm{id}$, d.h. auch die Einselemente der Ringe werden aufeinander abgebildet. ◊

Satz 4.4.

Seien A, B Ringe und $\varphi : A \to B$ ein Ringhomomorphismus mit $\varphi(1_A) = 1_B$.

Ist dann M ein B-Modul, so kann M auch als A-Modul aufgefasst werden, d.h. durch $(\otimes_B, \varphi)$ wird in kanonischer Weise eine Modulmultiplikation $\otimes_A$ auf M induziert.

Beweis.

Definiere $\otimes_A : A \times M \to M$ durch $(a, x) \mapsto a\varphi(a) \otimes_B x$. Dann gelten:

1. $a \otimes_A (x \oplus y) = \varphi(a) \otimes_B (x \oplus y) = (\varphi(a) \otimes_B x) \oplus (\varphi(a) \otimes_B y) = (a \otimes_A x) \oplus (a \otimes_A y)$.
2. $(a+b) \otimes_B x = \varphi(a+b) \otimes_B x = (\varphi(a) + \varphi(b)) \otimes_B x = (\varphi(a) \otimes_B x) \oplus (\varphi(b) \otimes_B x) = (a \otimes_A x) \oplus (b \otimes_A x)$.
3. $(ab) \otimes_A x = \varphi(ab) \otimes_B x = (\varphi(a)\varphi(b)) \otimes_B x = \varphi(a) \otimes_B (\varphi(b) \otimes_B x) = \varphi(a) \otimes_B (b \otimes_A x) = a \otimes_A (b \otimes_A x)$.
4. $1 \otimes_A x = \varphi(1) \otimes_B x = 1 \otimes_A x = x$. □

Definition 4.5.

Seien A ein Ring und M ein A-Modul. $L \subseteq M$ heißt ein Untermodul, falls gelten:

1. $(L, \oplus)$ ist eine Untergruppe von $(M, \oplus)$,
2. $(L, \otimes)$ ist abgeschlossen, d.h. für alle $a \in A$ und alle $x \in L$ gilt $a \otimes x \in L$.

Bemerkung 4.6.

1. Es genügt dabei zu zeigen, dass mit $a, b \in A$ und $x, y \in L$ gilt $(a \otimes x) \oplus (b \otimes y) \in L$.
2. Ab jetzt unterscheiden wir bei der Notation nicht mehr zwischen $+, \oplus$ bzw. zwischen $\cdot, \otimes$. ◊

Satz 4.7.

Sei V ein K-Vektorraum, aufgefasst als $K[f]$-Modul für ein $f \in \mathrm{End}_K(V)$, und sei U ein Untervektorraum von V.

U ist genau dann ein $K[f]$-Untermodul von V, wenn U invariant unter f ist.

Beweis.

$$\begin{aligned} U \text{ ist ein } K[f]\text{-Modul} &\iff U \text{ ist abgeschlossen unter } A \times U \to U,\ (a,u) \mapsto au \\ &\iff U \text{ ist abgeschlossen unter } K[X] \times U \to U,\ (f,u) \mapsto f(u) \\ &\iff f(U) \subseteq U. \end{aligned}$$

□

Bemerkung 4.8.

Sei der $K[f]$-Modul V zerlegt in die direkte Summe $V_1 \oplus \cdots \oplus V_r$, wobei die V_i invariante $K[f]$-Untermoduln seien. Wählt man eine K-Vektorraumbasis $\mathfrak{W}_i$ für V_i und setzt $A_i = \mathrm{Mat}_{\mathfrak{W}_i}^{\mathfrak{W}_i}(f_{|V_i})$, $\mathfrak{W} = (\mathfrak{W}_1, ..., \mathfrak{W}_r)$, so gilt:

$$\mathrm{Mat}_{\mathfrak{W}}^{\mathfrak{W}}(f) = \mathrm{diag}(A_1, ..., A_r)$$

Unser Ziel ist, V in möglichst einfache $K[f]$-Untermoduln zu zerlegen und so eine allgemeine Normalform für beliebige Matrizen über einem Körper K zu erhalten. Speziell für orthogonale und für symmetrische Matrizen haben wir bereits solche Normalformen kennen gelernt: die Diagonalmatrizen bzw. die Drehmatrizen um Hauptachsen. ◊

4.2. Modulhomomorphismen

Definition 4.9.

Seien M und L zwei A-Moduln. $f : L \to M$ heißt ein A-Modul-Homomorphismus, falls für alle $a \in A$ und alle $x, y \in M$ gelten:

$$f(x+y) = f(x) + f(y) \qquad \text{und} \qquad f(ax) = af(x).$$

$\mathrm{Bild}(f) = f(L) \subseteq M$ heißt das Bild und $\mathrm{Kern}(f) = f^{-1}(\{0\}) \subseteq L$ der Kern von f.

Bemerkung 4.10.

1. $f(L)$ ist ein Untermodul von M und $\mathrm{Kern}(f)$ ist ein Untermodul von L.
2. Es gilt: f ist injektiv gdw. $\mathrm{Kern}(f) = \{0\}$.
3. Ist L ein Untermodul von M, so wird die Faktorgruppe $M/L = \{x + L \mid x \in M\}$ (der additiven Gruppen) zu einem A-Modul durch: $a(x+L) = (ax+L)$.

   Unabhängigkeit des Vertreters: Sei $x + L = x' + L$, dann ist $x - x' \in L$, also auch $a(x-x') \in L$, d.h. $ax - ax' \in L$ und damit $ax + L = ax' + L$.
4. Jeder Modulhomomorphismus $f : L \to M$ induziert einen Modulisomorphismus $\bar{f} : L/N \to f(M)$ mit $N = \mathrm{Kern}(f)$. Das Diagramm

$$\begin{array}{ccc} L & \xrightarrow{\ f\ } & M \\ \downarrow{\scriptstyle \phi} & & \uparrow{\scriptstyle \mathrm{id}} \\ L/N & \xrightarrow{\ \bar{f}\ } & f(L) \end{array}$$

   kommutiert, d.h. es gilt $f = \bar{f} \circ \phi$. ◊

Satz 4.11. (Homomorphiesatz)

Ist $f : L \to M$ ein A-Modulhomomorphismus, so ist $L/\mathrm{Kern}(f) \cong f(L)$ als A-Moduln.

Beweis.

Es bleibt nur zu zeigen, dass $\bar{f}$ verträglich mit der skalaren Multiplikation ist:

$$\bar{f}(a(x+N)) = \bar{f}(ax+N) = f(ax) = af(x) = a\bar{f}(x+N).$$

□

4.3. Quotientenkörper

Bemerkung 4.12.

Wir werden in diesem Abschnitt zeigen, dass sich jeder Integritätsbereich durch Hinzunahme der fehlenden Einheiten in kanonischer Weise zu einem Körper vervollständigen lässt. ◊

Lemma 4.13.

Seien R ein Integritätsbereich und $a, b, c, d \in R$, $b, d \neq 0$. Wir definieren

$$(a, b) \sim (c, d) \qquad \Longleftrightarrow \qquad ad = bc.$$

Dann ist $\sim$ eine Äquivalenzrelation auf $R \times R$.

Beweis.

1. Für alle $a, b \in R$ gilt $ab = ba$, d.h. $(a, b) \sim (a, b)$.
2. Für alle a, b, c, d gilt $ad = bc \Rightarrow cb = da$, also auch $(a, b) \sim (c, d) \Rightarrow (c, d) \sim (a, b)$.
3. Für alle $a, b, c, d, e, f \in R$ gilt: $ad = bc$ & $cf = de \Rightarrow adf = bcf$ & $bcf = bde \Rightarrow adf = bde \Rightarrow af = be$, also auch $(a, b) \sim (e, f)$. □

Definition 4.14.

Wir definieren

$$\frac{a}{b} := [(a, b)]_\sim = \{(c, d) \mid ad = bc\}, \qquad \mathrm{Quot}(R) = \left\{ \frac{a}{b} \mid a, b \in R \right\} \subseteq R \times R.$$

$\mathrm{Quot}(R)$, versehen mit den Operationen

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

wird zu einem Körper, dem Quotientenkörper von R.

Bemerkung 4.15.

1. Wie üblich ist die Wohldefiniertheit dieser Operationen nachzurechnen. Gelten hierfür $(a, b) \sim (c, d)$ und $(a', b') \sim (c', d')$, d.h. $ad = bc$ und $a'd' = b'c'$, dann
$$\frac{a}{b} + \frac{a'}{b'} = \frac{ab' + ba'}{bb'} = \frac{dab'd' + dba'd'}{dbb'd'} = \frac{bcb'd' + dbb'c'}{dbb'd'} = \frac{cd' + dc'}{dd'} = \frac{c}{d} + \frac{c'}{d'}$$
und
$$\frac{a}{b}\frac{a'}{b'} = \frac{aa'}{bb'} = \frac{daa'd'}{dbb'd'} = \frac{bcb'c'}{dbb'd'} = \frac{c}{d}\frac{c'}{d'}.$$
2. Wir fassen R auf als Teilmenge von $\mathrm{Quot}(R)$ via der Abbildung $\iota : r \mapsto \frac{r}{1}$. ι ist eine Einbettung, d.h. ein Ringmonomorphismus. Wir schreiben $R \hookrightarrow \mathrm{Quot}(R)$ via ι und identifizieren R mit $\iota(R)$. ◊

4.4. Erzeugendensysteme und Basen von Moduln

Definition 4.16.

Seien A ein Ring, M ein A-Modul und $W \subseteq M$. Dann gilt

$$\langle W \rangle_A = \mathrm{span}_A(W) := \{a_1 x_1 + \cdots + a_n x_n \mid n \in \mathbb{N}_0,\ a_i \in A,\ x_i \in W\} = \bigcap_{\substack{W \subseteq L \\ L \text{ Untermodul}}} L$$

und $\langle W \rangle_A$ ist selbst ein Untermodul von M, der von W erzeugte Untermodul.

M heißt ein endlich erzeugter Modul, falls es $x_1, ..., x_r$ gibt mit $\langle \{x_1, .., x_r\} \rangle_A = M$.

Bemerkung 4.17.

Ist $M = \langle\{x_1, ..., x_r\}\rangle_A = M$, dann lässt sich M darstellen als $M = Ax_1 + \cdots + Ax_r$. ◊

Definition 4.18.

Seien $M_1, ..., M_r$ Untermoduln des A-Moduls M. M ist die direkte Summe von $M_1, ..., M_4$, in Zeichen: $M = M_1 \oplus \cdots \oplus M_r$, falls jedes $x \in M$ genau eine Darstellung $x = x_1 + \cdots + x_r$ mit $x_i \in M_i$ besitzt.

Bemerkung 4.19.

1. Ist M_1 ein Untermodul von M, so muss es keinen Untermodul M_2 mit $M = M_1 \oplus M_2$ geben. Im Gegensatz dazu existiert zu jedem Untervektorraum U eines K-Vektorraumes V ein Unterraum $U^{\oplus}$ mit $V = U \oplus U^{\oplus}$.
2. Genau dann ist $M = M_1 \oplus M_2$, wenn $M = M_1 + M_2$ und $M_1 \cap M_2 = \{0\}$, Beweis wie bei Vektorräumen.
3. Bei K-Vektorräumen gilt: $\alpha x = 0 \Leftrightarrow \alpha = 0$ oder $x = 0$. Dies gilt nicht mehr für Moduln:

   Fasse die Abelsche Gruppe $(\mathbb{Z}/m\mathbb{Z}, +)$ auf als $\mathbb{Z}$-Modul mit $m\overline{x} = \overline{mx}$, dann gilt $m\overline{x} = 0$ für alle $\overline{x} \in \mathbb{Z}/m\mathbb{Z}$. ◊

Definition 4.20.

$(x_1, ..., x_n)$ heißt eine Modulbasis von M, falls $M = Ax_1 + \cdots + Ax_n$ und für alle $a_1, ..., a_n \in A$ gilt: Ist $a_1x_1 \oplus \cdots \oplus a_nx_n = 0$, dann sind bereits alle $a_i = 0$.

M heißt in diesem Fall frei und $\{x_1, ..., x_n\}$ heißt A-linear unabhängig.

Bemerkung 4.21.

1. Ist $\{x_1, ..., x_n\}$ eine Basis von M, dann gilt $M = Ax_1 \oplus \cdots \oplus Ax_n$. Die Umkehrung ist im Allgemeinen falsch, denn ist $M = Ax_1 + \cdots + Ax_n$, dann folgt aus $a_1x_1 + \cdots + a_nx_n = 0$ nur, dass alle $a_ix_i = 0$, nicht jedoch notwendigerweise $a_i = 0$.
2. Jeder endlich-dimensionale K-Vektorraum ist ein endlich erzeugter, freier K-Modul, aber nicht jeder endlich erzeugte Modul ist frei. ◊

Satz 4.22.

Seien R ein Integritätsbereich und M ein R-Modul.

Seien $(x_1, ..., x_n)$ und $(y_1, ..., y_m)$ zwei R-Basen von M. Dann gilt $n = m$.

Beweis.

Wir betten M in einen Vektorraum über $K = \mathrm{Quot}(R)$ ein via des Modulmonomorphismus

$$\sigma : M \to K^n, \qquad u = a_1x_1 + \cdots + a_nx_n \mapsto (a_1, ..., a_n) :$$

1. σ ist wohldefiniert, da die Darstellung jedes Elements aus M als Linearkombination der x_i eindeutig.
2. Seien $u = a_1x_1 + \cdots + a_nx_n$, $v = b_1x_1 + \cdots + b_nx_n$ und $a \in R$, dann gelten
$$\sigma(u + v) = (a_1 + b_1, ..., a_n + b_n) = \sigma(u) + \sigma(v), \qquad \sigma(au) = (au_1, ..., au_n) = a\sigma(u).$$
3. σ ist injektiv: Gelte $\sigma(a_1x_1 + \cdots + a_nx_n) = 0$, dann ist $(a_1, ..., a_n) = 0$, d.h. alle $a_i = 0$. Damit ist $\mathrm{Kern}(\sigma) = \{0\}$.

Anngenommen, $m > n$. Dann sind $\sigma(y_1), ..., \sigma(y_m) \in K^n$ linear unabhängig über K, es gibt also $a_1, ..., a_n \in R$ und $b \in R\backslash\{0\}$ mit $\frac{a_1}{b}\sigma(y_1) + \cdots + \frac{a_m}{b}\sigma(y_m) = 0$, und nicht alle $a_i = 0$. Dann ist $a_1\sigma(y_1) + \cdots + a_m\sigma(y_m) = 0$, d.h. $\sigma(a_1y_1 + \cdots + a_my_m) = 0$, also $a_1y_1 + \cdots + a_my_m = 0$, da σ injektiv. Da aber nicht alle $a_i = 0$, kann $(y_1, ..., y_m)$ keine R-Basis sein, ein Widerspruch. □

4.5. Hauptsatz für endlich erzeugte Moduln über Hauptidealringen

Bemerkung 4.23.

Seien R ein Hauptidealring und Integritätsbereich und M ein freier R-Modul mit Basis $v_1, ..., v_n \in M$. Weiter sei $K = \text{Quot}(R)$ der Quotientenkörper von R. Dann definieren die Projektionen $\pi_i : M \to R$, $\pi_i(a_1v_1 + \cdots + a_nv_n) = a_i$ Modulepimorphismen und die Koordinatenabbildung $\Psi : M \to R^n$ mit $\Psi(a_1v_1 \cdots + a_nv_n) = (a_1, ..., a_n)$ einen Modulisomorphismus. Wir identifizieren M mit $R^n \subseteq K^n$. ◊

Definition 4.24.

Seien R ein Integritätsbereich, M ein freier R-Modul und $U \subseteq M$ ein R-Untermodul von M. Dann bezeichnet $\text{rang}(U) = \dim_K \langle U \rangle_R$ den Rang von U im K^n. Elemente $u_1, ..., u_m \in M$ heißen R-linear abhängig, falls $u_1, ..., u_m$ K-linear abhängig sind.

Bemerkung 4.25.

1. Sind $u_1, ..., u_m$ R-linear unabhängig, dann auch K-linear unabhängig: Gelte $0 = \frac{a_1}{b}u_1 + \cdots + \frac{a_m}{b}u_m$, dann auch $0 = a_1u_1 + \cdots + a_mu_m$, d.h. alle $a_i = 0$ und damit auch alle $\frac{a_i}{b} = 0$.
2. Insbesondere: Ist $(u_1, ..., u_m)$ eine R-Basis von U, dann auch eine K-Basis von $\langle U \rangle_R$.
3. Die Umkehrung ist falsch: Nicht jede K-Basis von $\langle U \rangle_R$ ist auch eine R-Basis von U. ◊

Satz 4.26. (Hauptsatz für endlich erzeugte Moduln über Hauptidealringen)

Seien R ein Hauptidealring und Integritätsbereich, M ein freier R-Modul mit Basis $(v_1, ..., v_n)$ und M' ein R-Untermodul von M. Dann gelten:

1. M' ist frei vom Rang $m \leq n$.
2. Ist $M' \neq \{0\}$, so gibt es eine Basis $e_1, ..., e_n$ von M und $\alpha_1, ..., \alpha_m \in R\backslash\{0\}$, so dass $(\alpha_1e_1, ..., \alpha_me_m)$ eine Basis von M' ist und $\alpha_i \mid \alpha_{i+1}$ für alle $1 \leq i \leq m-1$.

Beweis.

Definiere $\mathcal{M} = \{f(M') \mid f \in \text{Hom}_R(M, R)\}$. Dann sind alle Elemente von $\mathcal{M}$ Ideale von R, also Hauptideale, und wegen $\pi_i \in \mathcal{M}$ ist $\mathcal{M} \neq \{0\}$. Die Relation $R\beta \subseteq R\gamma :\Leftrightarrow \gamma \mid \beta$ definiert eine partielle Ordnung auf $\mathcal{M}$. $\mathcal{M}$ besitzt nach Zorns Lemma ein maximales Element $h(M') = R\alpha$ und wegen $\mathcal{M} \neq 0$ ist $\alpha \neq 0$. Sei $e' \in M'$ mit $h(e') = \alpha$.

1. Für alle $f \in \text{Hom}_R(M, R)$ gilt $\alpha \mid f(e')$:

   Sei δ ein Erzeuger des Hauptideals $R\delta = R\alpha + Rf(e')$. Dann gibt es $\beta, \gamma \in R$ mit $\beta\alpha + \gamma f(e') = \delta$. Setze $g = \beta h + \gamma f$, dann ist $g \in \text{Hom}_R(M, R)$ mit $g(e') = \delta$, d.h. $R\delta \subseteq g(M')$. Da $R\alpha$ maximal in $\mathcal{M}$, folgt aus $R\alpha \subseteq R\delta \subseteq g(M') \subseteq R\alpha$, dass $R\alpha = R\delta$, d.h. $f(e') \in R\alpha$ und damit $\alpha \mid f(e')$. ◊

Speziell gilt $\alpha \mid \pi_i(e')$ für alle Projektionen π_i, d.h. $e = \frac{1}{\alpha}e' = \sum \frac{1}{\alpha}\pi_i(e')v_i$ liegt in M. Es gilt $h(e) = 1$, denn $\alpha h(e) = h(\alpha e) = h(e') = \alpha$.

2. $M = \text{Kern}(h) \oplus Re$:

   Sei $v \in M$ beliebig. Dann ist $v = h(v)e + (v - h(v)e)$ und wegen $h(v - h(v)e) = h(v) - h(v)h(e) = 0$ ist $v - h(v)e \in \text{Kern}(h)$, d.h. $v \in \text{Kern}(h) + Re$. Sei nun $v \in \text{Kern}(h) \cap Re$, dann ist $v = \gamma e$ für ein $\gamma \in R$ und $0 = h(v) = \gamma h(e) = \gamma$, d.h. $v = 0$. Also ist $\text{Kern}(h) \cap Re = \{0\}$ und damit $M = \text{Kern}(h) \oplus Re$. ◊

3. $M' = (M' \cap \text{Kern}(h)) \oplus Re'$

   Sei $v \in M'$, dann ist $h(v) = \beta\alpha$ für ein $\beta \in R$. Zerlege $v = \beta\alpha e + (v - h(v)e)$, dann ist $\beta\alpha e = \beta e' \in Re'$, es gilt $v - h(v)e \in \text{Kern}(h)$ und $h(v)e = \beta e' \in M'$. Sei nun $v \in (M' \cap \text{Kern}(h)) \cap Re'$, d.h. $v = \beta e'$ und $h(v) = 0$, dann ist $h(\beta e') = \beta h(e') = \beta\alpha = 0$ und wegen $\alpha \neq 0$ somit $\beta = 0$, d.h. $v = 0$. ◊

Per Induktion über $m = \text{rang}(M')$ folgt: M' ist frei. Sei nämlich jeder Untermodul von kleinerem Rang als m frei. Dann ist insbesondere $M' \cap \text{Kern}(h)$ frei nach (3). Sei $(v_1, ..., v_{m-1})$ eine Basis von $M' \cap \text{Kern}(h)$, dann ist $(v_1, ..., v_{m-1}, e')$ eine Basis von M'.

4. Es gibt $\alpha_1, ..., \alpha_m \in R$, so dass $(\alpha_1 e_1, ..., \alpha_m e_m)$ eine Basis von M' ist:

   Wiederum per Induktion, diesmal über $n = \text{rang}(M)$. Sei die Behauptung für jeden Modul mit kleinerem Rang als n erfüllt, dann besitzt nach (2) insbesondere $\text{Kern}(h)$ eine Basis $(e_2, ..., e_n)$ und es gibt $\alpha_2, ..., \alpha_m \in R\backslash\{0\}$, so dass $(\alpha_2 e_2, ..., \alpha_m e_m)$ eine Basis von $M' \cap \text{Kern}(h)$ isr mit $\alpha_i \mid \alpha_{i+1}$. Wähle nun $e_1 = e$ und $\alpha_1 = \alpha$, dann ist $(e_1, ..., e_n)$ nach (2) eine Basis von M und $(\alpha_1 e_1, ..., \alpha_m e_m)$ ist nach (3) eine Basis von M', denn $\alpha_1 e_1 = \alpha e = e'$. ◊

5. $\alpha_1 \mid \alpha_2$:

   Definiere $f \in \text{Hom}_R(M, R)$ durch $f(e_1) = 1$, $f(e_2) = 1$ und $f(e_i) = 0$ für $i \geq 3$. Dann folgt aber $\alpha_1 = f(\alpha_1 e_1) = f(\alpha e) = f(e')$, d.h. $R\alpha \subseteq f(M')$. Wegen der Maximalität von $R\alpha$ erhalten wir $R\alpha = f(M')$. Nun ist $\alpha_2 = \alpha_2 f(e_2) = f(\alpha_2 e_2) \in f(M') = R\alpha_1$, d.h. $\alpha_1 \mid \alpha_2$. □

4.6. Struktursätze für endlich erzeugte Moduln über Hauptidealringen

Definition 4.27.

Seien M ein Modul über dem Hauptidealring und Integritätsbereich R, $v \in M$ und $f : R \to Rv$, $f(\alpha) = \alpha v$. Dann heißt

$$\text{Ann}(v) = \text{Kern}(f) = \{\alpha \in R \mid \alpha v = 0\}$$

der Annihilator oder auch der Annulator von v in M.

Beispiel 4.28.

In $\mathbb{Z}$-Moduln gilt $\text{Ann}(v) = \text{Ord}(v)\mathbb{Z}$. Beispielsweise ist für $M = \mathbb{Z}/p\mathbb{Z}$ mit p prim stets $\text{Ann}(v) = p\mathbb{Z}$ für $v \neq \overline{0}$ und $\text{Ann}(\overline{0}) = \mathbb{Z}$ (klar). In $M = \mathbb{Z}/pq\mathbb{Z}$ mit p, q prim dagegen sind beispielsweise $\text{Ann}(\overline{p}) = q\mathbb{Z}$ und $\text{Ann}(\overline{q}) = p\mathbb{Z}$. ◊

Satz 4.29. (Erster Struktursätze für endlich erzeugte Moduln über Hauptidealringen)

Sei M ein endlich erzeugter Modul über dem Hauptidealring und Integritätsbereich R.

Dann gibt es $v_1, ..., v_n \in M$ und $\alpha_1, ..., \alpha_n \in R$ mit $\alpha_1 \mid \cdots \mid \alpha_n$, so dass gelten $M = Rv_1 \oplus \cdots \oplus Rv_n$ und $\text{Ann}(v_i) = R\alpha_i$.

Es gilt also speziell $\sum \gamma_i v_i = \sum \delta_i v_i$ genau dann, wenn $\gamma_i \equiv \delta_i \mod R\alpha_i$ für alle i.

Beweis.

Sei $M = Rw_1 + \cdots + Rw_n$. Dann definiert $\Phi : R^n \to M$, $\Phi(e^{(i)}) = w_i$, d.h. $((a_1, ..., a_n) \mapsto a_1 w_1 + \cdots + a_n w_n$, einen R-Modulhomomorphismus. Da R^n frei ist mit kanonischer Basis $(e^{(1)}, ..., e^{(n)})$, existieren nach dem Hauptsatz auch eine Basis $(e_1, ..., e_n)$ von R^n und $\alpha_1, ..., \alpha_m \in R\backslash\{0\}$, so dass $(\alpha_1 e_1, ..., \alpha_m e_m)$ eine Basis von $\text{Kern}(\Phi)$ ist (mit $m \leq n$) und $\alpha_1 \mid \cdots \mid \alpha_n$. Setze $v_1 = \Phi(e_1)$, ..., $v_n = \Phi(e_n)$. Sei $v \in M$ beliebig. Da Φ surjektiv ist, gibt es $\gamma_1, ..., \gamma_n \in R$ mit $f = \Phi(\gamma_1 e_1 + \cdots + \gamma_n e_n) = \gamma_1 v_1 + \cdots + \gamma_n v_n$, d.h. $M = Rv_1 + \cdots + Rv_n$. Setze $\alpha_{m+1} = 0$, ..., $\alpha_n = 0$, dann ist

$$\begin{aligned}
\sum_{i=1}^n \gamma_i v_i = \sum_{i=1}^n \delta_i v_i \quad &\iff \quad 0 = \sum_{i=1}^n (\gamma_i - \delta_i) v_i = f\left(\sum_{i=1}^n (\gamma_i - \delta_i) e_i\right) \\
&\iff \quad \sum_{i=1}^n (\gamma_i - \delta_i) e_i \in \text{Kern}(\Phi) \\
&\iff \quad \sum_{i=1}^n (\gamma_i - \delta_i) e_i = \sum_{i=1}^n \beta_i (\alpha_i e_i) \text{ für gewisse } \beta_1, ..., \beta_n \in R \\
&\iff \quad \gamma_i - \delta_i \in R\alpha_i, \text{ da } R\alpha_1 \supseteq \cdots \supseteq R\alpha_n.
\end{aligned}$$

Schließlich gilt

$$\begin{aligned} \alpha \in \mathrm{Ann}(v_i) &\iff 0 = \alpha v_i = \Phi(\alpha e_i) \\ &\iff \alpha e_i \in \mathrm{Kern}(\Phi) \\ &\iff \alpha_i \mid \alpha \\ &\iff \alpha \in R\alpha_i, \end{aligned}$$

d.h. $\mathrm{Ann}(v_i) = R\alpha_i$ für alle $i = 1, ..., n$. □

Satz 4.30. (Zweiter Struktursatz für endlich erzeugte Moduln über Hauptidealringen)

Sei R Hauptidealring und Integritätsbereich und sei M ein endlich erzeugter R-Modul.

Dann existieren $v_1, ..., v_m \in M$ und $p_1, ..., p_m \in R$ prim und $\nu_1, ..., \nu_m \in \mathbb{N}_0$ mit $M = Rv_1 \oplus \cdots \oplus Rv_m$ und $\mathrm{Ann}(v_i) = Rp_i^{\nu_i}$ oder $\{0\}$.

Beweis.

Nach dem letzten Satz ist $M = Rw_1 \oplus \cdots \oplus Rw_n$ mit $\mathrm{Ann}(w_1) \supseteq \cdots \supseteq \mathrm{Ann}(w_n)$. Als Hauptideal besitzt jedes $\mathrm{Ann}(w_k)$ eine Darstellung $\mathrm{Ann}(w_k) = Rp_1^{\nu_1} \cdots p_s^{\nu_s}$ mit $\nu_1, ..., \nu_s \neq 0$. Nach dem Chinesischen Restsatz existieren dazu $x_1, ..., x_s \in R$ mit $x_j \equiv 1 \bmod Rp_j^{\nu_j}$, $1 \neq j \neq s$, und $x_j \equiv 0 \bmod Rp_i^{\nu_i}$ für $i \neq j$.

Wir setzen $v_j = x_j w_k$, dann gilt $\alpha v_j = 0 \Leftrightarrow \alpha x_j \in \mathrm{Ann}(w_k) \Leftrightarrow p_j^{\nu_j} \mid \alpha$, also ist $\mathrm{Ann}(v_j) = Rp_j^{\nu_j}$. Es gilt $\sum x_j \equiv 1 \bmod Rp_i^{\nu_i} \equiv 1 \bmod Rp_1^{\nu_1} \cdots p_s^{\nu_s}$ für alle i. Da sich $w_k = (\sum x_j w_k) = \sum v_j$ als Linearkombination der v_j darstellen lässt, ist $Rw = Rv_1 + \cdots + Rv_s$. Falls nun $\sum \beta_j v_j = 0$, so ist $\sum(\beta_j x_j)w_k = 0$, also $\sum \beta_j x_j \in \mathrm{Ann}(w_k)$ und damit $p_j^{\nu_j} \mid \beta_j$, d.h. $\beta_j \in \mathrm{Ann}(v_j)$. Insbesondere ist $Rw_k = Rv_1 \oplus \cdots \oplus Rv_s$ und aus $Rw_1 \supseteq \cdots \supseteq Rw_m$ folgt $M = Rv_1 \oplus \cdots \oplus Rv_m$. □

Satz 4.31. (Struktursatz für endlich erzeugte, Abelsche Gruppen)

Sei G eine endlich erzeugte, Abelsche Gruppe. Dann ist G direkte Summe von endlichen zyklischen Gruppen von Primzahlpotenzordnung oder $\mathbb{Z}$.

Beweis.

G kann aufgefasst werden als endlich erzeugter $\mathbb{Z}$-Modul $G = \mathbb{Z}v_1 \oplus \cdots \oplus \mathbb{Z}v_m$. Nach dem Zweiten Struktursatz ist dann $\mathbb{Z}v_i \cong \mathbb{Z}/p_i^{\nu_i}\mathbb{Z}$, falls $\mathrm{Ann}(v_i) = \mathbb{Z}p_i^{\nu_i}$ endlich ist, oder $\mathbb{Z}v_i \cong \mathbb{Z}/\{0\} \cong \mathbb{Z}$ im Falle $\mathrm{Ann}(v_i) = \{0\}$. □

5. Normalformen von Matrizen

5.1. Allgemeine Normalform über K

Bemerkung 5.1.

Seien K ein Körper, V ein endlichdimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$. Wir suchen eine K-Basis $\mathfrak{V}$ von V, so dass die Darstellungsmatrix $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f)$ von möglichst einfacher Form ist.

1. Ist $V = Kv_1 \oplus ... \oplus Kv_n$ mit $\mathfrak{V} = (v_1, ..., v_n)$ Basis aus Eigenvektoren von f zu Eigenwerten $\lambda_1, ..., \lambda_n$, dann hat $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f)$ Diagonalgestalt:

   $$\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \mathrm{diag}(\lambda_1, ..., \lambda_n).$$

   Die von den Eigenvektoren aufgespannten Eigenräume $\langle v_i \rangle$ sind stets f-invariant, d.h. $f(\langle v_i \rangle) \subseteq \langle v_i \rangle$.

2. Ist $V = \mathbb{R}^n$ oder $V = \mathbb{C}^n$ und ist f selbstadjungiert, d.h. gilt $\langle f(v), w \rangle = \langle v, f(w) \rangle$ für alle $v, w \in V$, dann besitzt f immer einen Basis aus Eigenvektoren zu f.

3. Ist $V = \mathbb{C}^n$ und ist f unitär, d.h. $\langle f(v), f(w) \rangle = \langle v, w \rangle$ für alle $v, w \in V$, dann sind sogar alle Eigenwerte ± 1, d.h. bzgl. einer geeigneten Basis $\mathfrak{V}$ ist $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \mathrm{diag}(\pm 1, ..., \pm 1)$.

4. Ist $V = \mathbb{R}^n$ und ist f orthogonal, d.h. $\langle f(v), f(w)\rangle = \langle v, w\rangle$ für alle $v, w \in V$, dann existieren Winkel $\theta_1, ..., \theta_k$ und eine Basis $\mathfrak{V}$ von V, sodass $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \mathrm{diag}(\pm 1, ..., \pm 1, \mathrm{Rot}(\theta_1), ..., \mathrm{Rot}(\theta_k))$, wobei die Drehmatrix $\mathrm{Rot}(\theta)$ um den Winkel θ definiert ist als $\mathrm{Rot}(\theta) = \left(\begin{smallmatrix} +\cos\theta & -\sin\theta \\ +\sin\theta & +\cos\theta \end{smallmatrix}\right)$. Dabei sind auch die von den zu den Rotationskästchen gehörigen Basisvektoren aufgespannten, zweidimensionalen Unterräume $\langle v_1(\theta_i), v_2(\theta_i)\rangle$ f-invariant.

5. Sei allgemeiner $V = V_1 \oplus \cdots \oplus V_m$ mit unter f invarianten Unterräumen $V_1, ..., V_m$ von V. Sei $\mathfrak{V}_i$ eine beliebige Basis von V_i und sei $M_i = \mathrm{Mat}_{\mathfrak{V}_i}^{\mathfrak{V}_i}(f|_{V_i})$. Dann gilt bzgl. der Basis $\mathfrak{V} = (\mathfrak{V}_1, ..., \mathfrak{V}_m)$ von V: $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \mathrm{diag}(M_1, ..., M_m)$. ◊

Bemerkung 5.2.

Seien ab jetzt K ein beliebiger Körper, V ein endlichdimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$.

1. Via des Einsetzungshomomorphismus $\mu : K[X] \to K[f] \subseteq \mathrm{End}(V)$, $p(X) \mapsto p(f)$ kann V als $K[X]$-Modul aufgefasst werden mit Multiplikation $K[X] \times V \to V$, $p(X)v := \mu(p)v = p(f)(v)$.

2. Genau dann ist ein Untervektorraum U von V invariant unter f, wenn U ein $K[X]$-Modul ist: Ist $f(U) \subseteq U$, dann auch $p(f)(U) \subseteq U$, d.h. U ist abgeschlossen unter der Modulmultiplikation und damit ein Untermodul des $K[X]$-Moduls V. Ist dagegen U ein $K[X]$-Modul, dann gilt insbesondere für das Polynom $p(X) = X$, dass $p(X)U = f(U) \subseteq U$, d.h. U ist f-invariant.

3. Sei $(w_1, ..., w_n)$ eine Vektorraumbasis von V, d.h. $V = Kw_1 \oplus \cdots \oplus Kw_n$. Dann besitzt V die Darstellung $V = K[X]v_1 + \cdots + K[X]v_n$ (nicht unbedingt $V = K[X]v_1 \oplus \cdots \oplus K[X]v_n$), d.h. V ist als $K[X]$-Modul eindlich erzeugt. Außerdem ist $K[X]$ ein Integritätsbereich und Hauptidealring. Nach dem Ersten Struktursatz für endlich erzeugte Moduln über Hauptidealringen gibt es dann $v_1, ..., v_r \in V$, so dass $V = V_1 \oplus \cdots \oplus V_r$ mit $V_i = K[X]v_i$ und $\mathrm{Ann}(v_i) = K[X]q_i$, $q_i \in K[X]$, $q_1 \mid \cdots \mid q_r$. ◊

Satz 5.3. (Struktursatz für endlichdimensionale K-Vektorräume)

Sei $w \in V$ beliebig, $w \neq 0$, mit $\mathrm{Ann}(w) = K[X]q$, $q(X) = \alpha_0 + \alpha_1 X + \cdots + \alpha_{d-1}X^{d-1} + X^d$.

Dann ist $\mathfrak{W} = (w, f(w), ..., f^{d-1}(w))$ eine K-Basis von $K[X]w$ und $\mathrm{Mat}_{\mathfrak{W}}^{\mathfrak{W}}(f|_{K[X]w})$ hat die Form

$$\mathrm{Mat}_{\mathfrak{W}}^{\mathfrak{W}}(f|_{K[X]w}) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & \ddots & \vdots & -\alpha_2 \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -\alpha_{d-1} \end{pmatrix}.$$

$\mathrm{Mat}_{\mathfrak{W}}^{\mathfrak{W}}(f|_{K[X]w})$ heißt die Allgemeine Normalform von $f|_{\mathrm{Ann}(w)}$.

Beweis.

1. Der Fall $q = 0$ kann nicht auftreten, denn andernfalls wäre $\mathrm{Ann}(w) = \{0\}$, d.h. aus $p(X)(w) = 0$ folgt stets $p = 0$. Seien dann $\beta_0, ..., \beta_n$ beliebig, so dass $\beta_0 w + \cdots + \beta_n f^n(w) = 0$. Wir definieren $p(X) = \beta_0 + \cdots + \beta_n X^n$, dann ist $p(X)w = 0$, d.h. $p = 0$ und somit $\beta_0 = \cdots = \beta_n = 0$, d.h. die $(n+1)$-elementige Menge $\{w, ..., f^n(w)\}$ ist linear unabhängig. Dies ist unmöglich, da $\dim V = n$.

2. Ebenso kann der Fall $q = \alpha_0$ ausgeschlossen werden, denn dann wäre $q \in K[X]^\times$ eine Einheit und somit $\mathrm{Ann}(w) = K[X]q = K[X]$, d.h. $p(X)w = 0$ für alle $p \in K[X]$. Insbesondere wäre $1w = 0$, ein Widerspruch.

3. Also ist $d = \deg q > 1$. Wir zeigen zunächst: $\{w, f(w), ..., f^{d-1}(w)\}$ ist linear unabhängig. Angenommen, es gäbe $\beta_0, \beta_1 ..., \beta_{d-1} \in K$ mit $\beta_0 w + \beta_1 f(w) + \cdots + \beta_{d-1} f^{d-1}(w) = 0$ und nicht alle $\beta_i = 0$. Dann ist $\deg p \geq 1$ für $p(X) = \beta_0 + \beta_1 X + \cdots + \beta_{d-1}X^{d-1}$ und $p(X)w = 0$, d.h. $p \in \mathrm{Ann}(w) = K[X]q$. Also $q \mid p$ und damit $d = \deg q \leq \deg p = d - 1$, ein Widerspruch.

4. Es gilt $Kw + Kf(w) + \cdots + Kf^{d-1}(w) = K[X]w$: Offenbar ist $Kw + Kf(w) + \cdots + Kf^{d-1}(w) \subseteq K[X]w$. Umgekehrt hat $f^d(w)$ wegen $q(f)(w) = 0$ die Darstellung $f^d(w) = (-\alpha_0 w - \alpha_1 f(w) - \cdots - \alpha_{d-1} f^{d-1})$,

d.h. $f^d(w) \in Kw + Kf(w) + \cdots + Kf^{d-1}(w)$. Iterativ erhalten wir, dass auch $f^{d+1}(w)$, $f^{d+2}(w)$, ... in $Kw + Kf(w) + \cdots + Kf^{d-1}(w)$ liegen, d.h. $K[X]w \subseteq Kw + Kf(w) + \cdots + Kf^{d-1}(w)$.

Also ist $\mathfrak{W} = (w, f(w), ..., f^{d-1}(w))$ eine K-Basis von $K[X]w$ und $\mathrm{Mat}_{\mathfrak{W}}^{\mathfrak{W}}(f|_{K[X]w})$ hat die Darstellung

$$\mathrm{Mat}_{\mathfrak{W}}^{\mathfrak{W}}(f|_{K[X]w}) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & \ddots & \vdots & -\alpha_2 \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -\alpha_{d-1} \end{pmatrix}.$$

Damit ist alles gezeigt. □

Lemma 5.4.

Es existiert genau ein normiertes Polynom $p_f \in K[X]$ mit

$$\mathrm{Ann}(V) = \{p \in K[X] \mid p(f)(v) = 0 \text{ für alle } v \in V\} = K[X]p_f.$$

p_f heißt das Minimalpolynom von f und teilt das charakteristischen Polynoms $P_f = \det(f - X\mathrm{Id})$.

Beweis.

1. Da $\mathrm{Ann}(V)$ ein $K[X]$-Ideal ist, besitzt es einen (nornmierten) Erzeuger $p_f \in K[X]$. Dieser ist eindeutig: Haben p, q die Eigenschaft eines Minimalpolynoms, dann $K[X]p = K[X]q$, d.h. $p \sim q$ und damit unterscheiden sich p, q nur um eine Einheit. Aus der Normiertheit folgt dann sofort $p = q$.
2. Nach dem Satz von Hamilton-Cayley ist $P_f(f) = 0$, d.h. für alle $v \in V$ gilt $P_f(f)(v) = 0$ und damit $P_f \in \mathrm{Ann}(V)$, d.h. $p_f \mid P_f$. □

Satz 5.5. (Normalform für Darstellungsmatrizen über K)

Es existieren Vektoren $v_1, ..., v_r$ und Grade $d_1, ..., d_r$, so dass sich V zerlegen lässt in $V = V_1 \oplus \cdots \oplus V_r$ mit zugehörigen Basen $\mathfrak{V}_1, ..., \mathfrak{V}_r$, $\mathfrak{V}_i = (v_i, f(v_i), ..., f^{d_i}(v_i))$ und f bzgl. der Basis $\mathfrak{V} = (\mathfrak{V}_1, ..., \mathfrak{V}_r)$ die Darstellung

$$\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \mathrm{diag}(\mathrm{Mat}_{\mathfrak{V}_1}^{\mathfrak{V}_1}(f|_{V_1}), ..., \mathrm{Mat}_{\mathfrak{V}_r}^{\mathfrak{V}_r}(f|_{V_r}))$$

besitzt, wobei alle $\mathrm{Mat}_{\mathfrak{V}_i}^{\mathfrak{V}_i}(f|_{V_i})$ Allgemeine Normalform haben.

Bemerkung 5.6.

Sei $V = K[X]v_1 \oplus \cdots \oplus K[X]v_r$ mit $\mathrm{Ann}(v_i) = K[X]q_i$, dann gilt $q_i \mid p_f$ und damit insbesondere $q_i \mid P_f$. Wegen $p_f(f)v = 0$ für alle $v \in V$ ist nämlich speziell $p_f(f)v = 0$ für alle $v \in K[X]v_i$, d.h. $p_f \in \mathrm{Ann}(v_i) = K[X]q_i$ und damit $q_i \mid p_f$. ◊

5.2. Normalform über $\mathbb{C}$

Bemerkung 5.7.

Sei ab jetzt $K = \mathbb{C}$ oder allgemeiner K algebraisch abgeschlossen, d.h. das charakteristische Polynom P_f eines Homomorphismus $f \in \mathrm{End}_{\mathbb{C}}(V)$ zerfalle in Linearfaktoren:

$$P_f := (-1)^n \prod_{j=1}^{s} (X - \lambda_j')^{\mu_j},$$

wobei $\lambda_1', ..., \lambda_s'$ die verschiedenen Eigenwerte von f sind und sich die algebraischen Vielfachheiten $\mu_1, ..., \mu_s$ zu n aufsummieren.

1. Stimmen die algebraischen Vielfachheiten mit den geometrischen $\nu_i = \dim \mathrm{Eig}(f, \lambda_i')$ überein, so ist besitzt V eine Basis aus Eigenvektoren zu f und f hat somit eine Darstellungsmatrix in Diagonalform. Der Raum V lässt sich dann in die Eigenräume zerlegen: $V = Kv_1 \oplus \cdots \oplus Kv_n$.
2. Andernfalls existieren $r < n$, $v_1, ..., v_r \in V$ und $q_1, ..., q_r \in K[X]$ mit $V = K[X]v_1 \oplus \cdots \oplus K[X]v_r$ und $\mathrm{Ann}(v_i) = q_i$. Wegen $q_i \mid P_f$ gibt es ein $d_i \leq \mu_i$ und ein $\lambda_i \in \{\lambda_1', ..., \lambda_s'\}$ mit $q_i(X) = (X - \lambda_i)^{d_i}$. ◊

Satz 5.8. (Strukturatz für endlichdimensionale $\mathbb{C}$-Vektorräume)

Seien $v \in \{v_1, ..., v_r\}$ und $\lambda \in \{\lambda_1', ..., \lambda_s'\}$ mit $\mathrm{Ann}(v) = K[X]q$, $q = (X - \lambda)^d$.

Dann ist $\mathfrak{V} = (v, (f - \lambda)(v), ..., (f - \lambda)^{d-1}(v))$ eine Basis von $K[X]v$ und $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f|_{K[X]v})$ hat die Form

$$\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f|_{K[X]v}) = \begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 1 & \lambda & 0 & \cdots & 0 \\ 0 & 1 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & \lambda \end{pmatrix}.$$

$\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f|_{K[X]v})$ heißt der Jordanblock zu $f|_{K[X]v}$ und $K[X]v$ der Hauptraum von v.

Beweis.

1. Wir wissen bereits, dass $\mathfrak{V}' = (v, f(v), ..., f^{d-1}(v))$ eine $\mathbb{C}$-Basis von $K[X]v$ bildet. Es genügt daher zu zeigen, dass die Darstellungsmatrix der Abbildung $\Phi : K[X]v \to K[X]v$, $\Phi(f^i(v)) = (f - \lambda)^i(v)$, invertierbar ist, d.h. einen Basiswechsel vermittelt:

$$\begin{cases} \mathfrak{v}_1 = (f-\lambda)^0(v) = v = \mathfrak{v}_1' \\ \mathfrak{v}_2 = (f-\lambda)^1(v) = f(v) - \lambda v = -\lambda\mathfrak{v}_1' + \mathfrak{v}_2' \\ \mathfrak{v}_3 = (f-\lambda)^2(v) = f^2(v) - 2\lambda f(v) + \lambda^2 v = \lambda^2\mathfrak{v}_1' - 2\lambda\mathfrak{v}_2' + \mathfrak{v}_3' \\ \quad\vdots \qquad\qquad\quad \vdots \qquad\qquad\qquad\qquad\quad \vdots \\ \mathfrak{v}_d = (f-\lambda)^{d-1}(v) = \cdots + f^d(v) = \cdots + \mathfrak{v}_d' \end{cases},$$

also ist

$$\mathrm{Mat}_{\mathfrak{V}'}^{\mathfrak{V}'}(\Phi) = \begin{pmatrix} 1 & -\lambda & \lambda^2 & \cdots & * \\ 0 & 1 & -2\lambda & \cdots & * \\ 0 & 0 & 1 & \cdots & * \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} = \mathrm{Mat}_{\mathfrak{V}'}^{\mathfrak{V}}(\mathrm{Id})$$

und insbesondere $\det(\mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{E}}(\Phi)) = 1$.

Die Matrizen $A = \mathrm{Mat}_{\mathfrak{V}'}^{\mathfrak{V}'}(\Phi)$ und $B = \mathrm{Mat}_{\mathfrak{V}'}^{\mathfrak{V}}(\mathrm{Id})$ erhält man so: Sei $A^{(i)}$ die i-te Spalte von A und sei $B^{(i)}$ die i-te Spalte von B, dann gelten

$$A^{(i)} = (\Psi_{\mathfrak{V}'} \circ \Phi \circ \Psi_{\mathfrak{V}'}^{-1})(e^{(i)}) = (\Psi_{\mathfrak{V}'} \circ \Phi)(\mathfrak{v}_i') = \Psi_{\mathfrak{V}'}(\mathfrak{v}_i) = \Psi_{\mathfrak{V}'}(a_{1i}\mathfrak{v}_1' + \cdots + a_{di}\mathfrak{v}_d') = (a_{1i}, ..., a_{di}),$$
$$B^{(i)} = (\Psi_{\mathfrak{V}'} \circ \mathrm{Id} \circ \Psi_{\mathfrak{V}}^{-1})(e^{(i)}) = (\Psi_{\mathfrak{V}'} \circ \mathrm{Id})(\mathfrak{v}_i) = \Psi_{\mathfrak{V}'}(\mathfrak{v}_i) = \Psi_{\mathfrak{V}'}(a_{1i}\mathfrak{v}_1' + \cdots + a_{di}\mathfrak{v}_d') = (a_{1i}, ..., a_{di}),$$

wobei $\Psi_{\mathfrak{W}} : W \to K^n$, $\mathfrak{w}_i \mapsto e^{(i)}$ die Koordinatenabbildung zu einer K-Basis $\mathfrak{W} = (\mathfrak{w}_1, ..., \mathfrak{w}_n)$ eines n-dimensionalen K-Vektorraumes W bezeichnet.

Um nun die Darstellungsmatrix von $f|_{K[X]v}$ bzgl. $\mathfrak{V}$ zu erhalten, betrachte

$$\begin{aligned} f(\mathfrak{v}_i) &= f(f-\lambda)^{i-1}(v) = (f-\lambda)(f-\lambda)^{i-1}(v) + \lambda(f-\lambda)^{i-1}(v) \\ &= (f-\lambda)^i(v) + \lambda(f-\lambda)^{i-1}(v) = \mathfrak{v}_{i+1} + \lambda\mathfrak{v}_i \end{aligned}$$

für $i \leq d-1$ und $f(\mathfrak{v}_d) = \lambda\mathfrak{v}_d$, da $(f-\lambda)^d(v) = 0$. Somit ist die i-te Spalte von $f|_{K[X]v}$ gegeben durch

$$\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f|_{K[X]v})^{(i)} = \Psi_{\mathfrak{V}}(f(\Psi_{\mathfrak{V}}^{-1}(e^{(i)}))) = \Psi_{\mathfrak{V}}(f(\mathfrak{v}_i)) = \Psi_{\mathfrak{V}}(\mathfrak{v}_{i+1} + \lambda\mathfrak{v}_i) = e^{(i+1)} + \lambda e^{(i)},$$

falls $i \leq d-1$, und $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f|_{K[X]v})^{(d)} = \lambda e^{(d)}$. □

Satz 5.9. (Normalform für Darstellungsmatrizen über $\mathbb{C}$)

Es existieren Vektoren $v_1, ..., v_r$, Eigenwerte $\lambda_1, ..., \lambda_r$ zu f und Grade $d_1, ..., d_r$, so dass sich V zerlegen lässt in $V = V_1 \oplus \cdots \oplus V_r$ mit zugehörigen Basen $\mathfrak{V}_1, ..., \mathfrak{V}_r$, $\mathfrak{V}_i = (v_i, (f-\lambda_i)(v_i), ..., (f-\lambda_i)^{d_i}(v_i))$ und f bzgl. der Basis $\mathfrak{V} = (\mathfrak{V}_1, ..., \mathfrak{V}_r)$ die Darstellung

$$\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) = \mathrm{diag}(\mathrm{Mat}_{\mathfrak{V}_1}^{\mathfrak{V}_1}(f|_{V_1}), ..., \mathrm{Mat}_{\mathfrak{V}_r}^{\mathfrak{V}_r}(f|_{V_r}))$$

besitzt, wobei alle $\mathrm{Mat}_{\mathfrak{V}_i}^{\mathfrak{V}_i}(f|_{V_i})$ Jordanblöcke sind. $\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f)$ heißt die Jordansche Normalform von f.

Bemerkung 5.10. (Berechnung einer Jordanbasis)

Wir fassen diejenigen Komponenten von $V = V_1 \oplus \cdots \oplus V_r$ zusammen, die zum gleichen Eigenwert gehören, und erhalten so $V = V(\lambda_1') \oplus \cdots \oplus V(\lambda_s')$ mit Basis $\mathfrak{V} = (\mathfrak{V}(\lambda_1'), ..., \mathfrak{V}(\lambda_s'))$. Seien $\lambda \in \{\lambda_1', ..., \lambda_s'\}$ und $W = V(\lambda)$, $\mathfrak{W} = \mathfrak{V}(\lambda)$. Dann setzt sich $\mathrm{Mat}_{\mathfrak{W}}^{\mathfrak{W}}(f|_W)$ aus mehreren Jordanblöcken zum Eigenwert λ zusammen. Bisher wissen wir allerdings noch nicht, wie die $v_1, ..., v_r$ gewählt werden müssen.

1. Sei $U_i = \mathrm{Kern}(f-\lambda)^i$, dann gibt es ein $d \in \mathbb{N}$, $1 \le d \le \mu(\lambda)$, mit $\{0\} = U_0 \subsetneq U_1 \subsetneq \cdots \subsetneq U_d = W$. Wir zerlegen $U_d = W$ zunächst in $U_d = U_{d-1} \oplus W_d$ mit $W_d = U_{d-1}^{\oplus}$ ein komplementärer Unterraum zu U_{d-1} in $U_d \subseteq V$.

   Es gelten $(f-\lambda)W_d \subseteq U_{d-1}$ und $(f-\lambda)W_d \cap U_{d-2} = \{0\}$, d.h. U_{d-1} ist zerlegbar in $U_{d-1} = U_{d-2} \oplus W_{d-1}$ mit $W_{d-1} = U_{d-2}^{\oplus}$ in U_{d-1}.

   Iterativ erhalten wir die Zerlegung

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

   Speziell ist $W_1 = U_1$ wegen $U_0 = \{0\}$.

2. Wir müssen noch geeignete Basen $\mathfrak{W}_1, ..., \mathfrak{W}_d$ von $W_1, ..., W_d$ finden.

   a) Wir wählen zunächst eine beliebige Basis von U_{d-1} und ergänzen diese durch $\mathfrak{W}_d = (w_1^d, ..., w_{s_d}^d)$ zu einer Basis von W. Dann ist $\mathfrak{W}_d$ eine Basis von W_d.

   b) Eine geeignete Basis von W_{d-1} ist dann durch $\mathfrak{W}_{d-1} = ((f-\lambda)w_1^d, ..., (f-\lambda)w_{s_d}^d, w_1^{d-1}, ..., w_{s_d-1}^{d-1})$ gegeben, wobei $w_1^{d-1}, ..., w_{s_d-1}^{d-1}$ so gewählt sind, dass sie $(f-\lambda)w_1^d, ..., (f-\lambda)w_{s_d}^d$ zu einer Basis von W_{d-1} ergänzen u.s.w.

   c) bis $\mathfrak{W}_1 = ((f-\lambda)^{d-1}w_1^d, ..., (f-\lambda)^{d-1}w_{s_d}^d, (f-\lambda)^{d-2}w_1^{d-1}, ..., (f-\lambda)^{d-2}w_{s_d-1}^{d-1}, ..., ..., w_1^1, ..., w_{s_1}^1)$ eine geeignete Basis von $W_1 = \mathrm{Kern}(f-\lambda)$ aus Eigenvektoren von f zu λ bildet.

3. Bezeichne nun $\mathrm{Jor}(\lambda, i, j) \in \mathbb{C}^{j \times j}$ den i-ten Jordanblock zu λ, wobei $j = 1, ..., d$ und $i = 1, ..., s_j$. Ordnen wir die Blöcke in $\mathrm{Mat}_{\mathfrak{W}}^{\mathfrak{W}}(f|_W)$ der Größe nach absteigend an, d.h.

$$\mathrm{Mat}_{\mathfrak{W}}^{\mathfrak{W}}(f|_W) = \mathrm{diag}(\mathrm{Jor}(\lambda, 1, d), ..., \mathrm{Jor}(\lambda, s_d, d), ..., ..., \mathrm{Jor}(\lambda, 1, 1), ..., \mathrm{Jor}(\lambda, s_1, 1)),$$

   dann setzt sich eine zugehörige Basis zusammen aus $(w_1^d, (f-\lambda)w_1^d, ..., (f-\lambda)^{d-1}w_1^d)$ (für den ersten Jordanblock der Größe d), $(w_2^d, (f-\lambda)w_2^d, ..., (f-\lambda)^{d-1}w_2^d)$ (für den zweiten Jordanblock der Größe d), $(w_1^{d-1}, (f-\lambda)w_1^{d-1}, ..., (f-\lambda)^{d-2}w_1^{d-1})$ (für den ersten Jordanblock der Größe $d-1$), u.s.w. bis $w_1^1, ..., w_{s_1^1}$ die einelementigen Basen für die Jordanblöcke der Größe 1 bilden.

Die Elemente aus $W_j \backslash \{0\}$ heißen Hauptvektoren der Stufe j; insbesondere sind die Hauptvektoren erster Stufe gerade die Eigenvektoren von f. Die Haupträume dieser Eigenvektoren sind gerade die von den Basen in (3) aufgespannten Unterräume von W bzw. V. ◊

Korollar 5.11. (Anzahl der Jordanblöcke zu einem Eigenwert)

Die Anzahl s_j der Jordanblöcke der Größe j, $j = 1, ..., d$, zum Eigenwert λ beträgt

$$s_j = 2\dim U_j - \dim U_{j-1} - \dim U_{j+1} = \dim(U_j/U_{j-1}) - \dim(U_{j+1}/U_j).$$

Speziell sind $s_0 = \dim \mathrm{Kern}(\mathrm{Id}) = 0$ und $s_1 = \dim \mathrm{Kern}(f - \lambda) =$ die geometrische Vielfachheit von λ.

Beweis. (Per Rückwärts-Induktion)

1. Induktionsanfang: Für $j = d$ gilt

$$\begin{aligned} s_d &= \dim U_d - \dim U_{d-1} \\ &= 2\dim U_d - \dim U_{d-1} - \dim U_{d+1}, \end{aligned}$$

denn $\dim U_d = \dim U_{d+1} = \dim W$.

2. Gelte die Behauptung also für $j+1, ..., d$. Die gewählten Basen von $W_1, ..., W_d$ sind

$$\begin{array}{lccccccc} \mathfrak{W}_d &= & \left(w_1^d, \right. & ..., & \left. w_{s_d}^d\right) & & & \\ \mathfrak{W}_{d-1} &= & \left((f-\lambda)w_1^d, \right. & ..., & \left. (f-\lambda)w_{s_d}^d, \right. & \left. w_1^{d-1}, \right. & ..., & \left. w_{s_{d-1}}^{d-1}\right) \\ \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ \mathfrak{W}_1 &= & \left((f-\lambda)^{d-1}w_1^d, ..., \right. & & (f-\lambda)^{d-1}w_{s_d}^d, & (f-\lambda)^{d-2}w_1^{d-1}, ..., & & \left.(f-\lambda)^{d-2}w_{s_{d-1}}^{d-1}, ..., w_1^1, ..., w_{s_1}^1\right). \end{array}$$

Somit gilt nach Induktionsannahme:

$$\begin{aligned} s_j &= \sharp\mathfrak{W}_j - s_{j-1} - \cdots - s_d \\ &= \sharp\mathfrak{W}_j - (\dim U_{j+1} - \dim U_j). \end{aligned}$$

Gleichzeitig ist

$$\begin{aligned} \sharp\mathfrak{W}_j &= \dim U_d - \dim U_{j-1} - (\sharp\mathfrak{W}_{j+1} + \cdots + \sharp\mathfrak{W}_d) \\ &= \dim U_d - \dim U_{j-1} - (\dim U_d - \dim U_j) \\ &= \dim U_j - \dim U_{j-1}, \end{aligned}$$

also

$$s_j = 2\dim U_j - \dim U_{j-1} - \dim U_{j+1}.$$

Die Darstellung über Quotientenräume folgt aus der Dimensionsformel, da $U_j \subseteq U_{j+1}$. □

Beispiel 5.12. (zur Berechnung der Jordanschen Normalform)

Wir transformieren die folgende Matrix auf Jordansche Normalform:

$$A = \begin{pmatrix} 25 & -16 & 30 & -44 & -12 \\ 13 & -7 & 18 & -26 & -6 \\ -18 & 12 & -21 & 36 & 12 \\ -9 & 6 & -12 & 21 & 6 \\ 11 & -8 & 15 & -22 & -3 \end{pmatrix}.$$

1. Dazu berechnen wir zunächst das charakteristische Polynom $P_f(X) = (X-3)^5$. Somit ist $\lambda = 3$ der einzige Eigenwert von A.

2. Die Kerne $U_0 \subsetneq U_1 \subsetneq \cdots \subsetneq U_{d-1} \subsetneq U_d$ sind

$$U_0 = \{0\}, \qquad U_1 = \langle(-2;1;2;0;0),\ (2;0;0;1;0),\ (2;2;0;0;1)\rangle, \qquad U_2 = \mathbb{C}^5,$$

denn via Gaußelimination erhalten wir

$$A-\lambda = \begin{pmatrix} 25-\lambda & -16 & 30 & -44 & -12 \\ 13 & -7-\lambda & 18 & -26 & -6 \\ -18 & 12 & -21-\lambda & 36 & 12 \\ -9 & 6 & -12 & 21-\lambda & 6 \\ 11 & -8 & 15 & -22 & -3-\lambda \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 & -2 & -2 \\ 0 & 1 & -\frac{1}{2} & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

woraus sich sofort eine Basis von U_1 ablesen lässt, und es ist $(A-\lambda)^2 = 0$. Somit ist $d = 2$ und es sind $\dim \mathrm{Kern}(U_0) = 0$, $\dim \mathrm{Kern}(U_1) = 3$, $\dim \mathrm{Kern}(U_2) = 5$.

3. Damit sind $s_0 = 0$, $s_1 = 1$, $s_2 = 2$, d.h. die Jordansche Normalform von A ist

$$J = \mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(A) = \begin{pmatrix} 3 & 0 & & & \\ 1 & 3 & & & \\ & & 3 & 0 & \\ & & 1 & 3 & \\ & & & & 3 \end{pmatrix}$$

bezüglich einer geeigneten Basis $\mathfrak{V}$.

4. Wir berechnen eine Basis von $U_1^{\oplus}$, etwa $\mathfrak{W}_2 = (w_1^1, w_2^1)$ mit

$$\mathfrak{w}_1 = w_1^2 = (1;0;0;0;0),$$
$$\mathfrak{w}_3 = w_2^2 = (0;1;0;0;0).$$

Dies sind die beiden Hauptvektoren erster Stufe.

5. Die zugehörigen Hauptvektoren erster Stufe, d.h. Eigenvektoren, sind folglich die ersten beiden Spalten von $A-\lambda$:

$$\mathfrak{w}_2 = (A-\lambda)w_1^2 = (22;12;-18;-9;11),$$
$$\mathfrak{w}_4 = (A-\lambda)w_2^2 = (-16;-10;12;6;-8).$$

Wir ergänzen $\{\mathfrak{w}_2, \mathfrak{w}_4\}$ zu einer Basis $\mathfrak{W}_1 = (\mathfrak{w}_2, \mathfrak{w}_4, \mathfrak{w}_5)$ von U_1, etwa durch

$$\mathfrak{w}_5 = w_1^1 = (2;0;0;1;0).$$

6. Dann ist $\mathfrak{V} = (\mathfrak{w}_1, \mathfrak{w}_2, \mathfrak{w}_3, \mathfrak{w}_4, \mathfrak{w}_5)$ eine Basis von $\mathbb{C}^5$, so dass mit der Basistransformationsmatrix $P = \mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{V}}(\mathrm{Id})$ gilt: $A = PJP^{-1}$ bzw.

$$A = \mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{E}}(A) = \mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{V}}(\mathrm{Id})\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(A)\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{E}}(\mathrm{Id})$$

Dabei sind die Spalten von P gegeben also $\mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{V}}(\mathrm{Id})^{(i)} = \Psi_{\mathfrak{V}}^{-1}(e_i) = \mathfrak{w}_i$:

$$P = \mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{V}}(\mathrm{Id}) = \begin{pmatrix} 22 & 1 & -16 & 0 & 2 \\ 13 & 0 & -10 & 1 & 0 \\ -18 & 0 & 12 & 0 & 0 \\ -9 & 0 & 6 & 0 & 1 \\ 11 & 0 & -8 & 0 & 0 \end{pmatrix}.$$

Damit ist alles berechnet. ◊

Bemerkung 5.13.

1. Die Jordansche Normalform ist bis auf Vertauschen der Jordanblöcke eindeutig bestimmt.
2. Wir haben gesehen, dass A ähnlich zu einer Matrix in Jordanscher Normalform ist, wenn χ_A in Linearfaktoren zerfällt. Es gilt auch die Umkehrung: Lässt sich eine Matrix jordanisieren, zu faktorisiert ihr charakteristisches Polynom. Das charakteristische Polynom ist nämlich invariant unter Ähnlichkeitstransformationen, d.h. $\chi_A(X) = \chi_J(X) = (\lambda_1 - X)^{\nu_1} \cdots (\lambda_s - X)^{\nu_s}$. ◊

5.3. Normalform über $\mathbb{R}$

Bemerkung 5.14.

Seien jetzt $K = \mathbb{R}$ und $f \in \mathrm{End}_{\mathbb{R}}(V)$. Dann zerfällt P_f über $\mathbb{C}$ in Linearfaktoren, d.h. es gibt eine Jordanmatrix $J \in \mathbb{C}^{n\times n}$ mit $A = \mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{V}}(f) \approx J$ über $\mathbb{C}$, d.h. $A = PJP^{-1}$ für eine komplexe Transformationsmatrix P, wobei $\mathfrak{V}$ eine beliebige Basis von V ist. Wir konstruieren aus J eine Normalform von f über $\mathbb{R}$.

Sei $\lambda \in \mathbb{C}$ ein Eigenwert von f. Dann ist auch $\overline{\lambda}$ ein Eigenwert von f:

$$PJP^{-1} = A = \overline{A} = \overline{PJP^{-1}} = (\overline{P})\overline{J}(\overline{P})^{-1} \qquad \Longrightarrow \qquad A \approx \overline{J} \text{ über } \mathbb{C}.$$

Im Falle $\lambda \in \mathbb{R}$ ist auch der zugehörige Jordanblock $\mathrm{Jor}(\lambda, \cdot, j) \in \mathbb{R}^{j\times j}$ reell. Andernfalls ist $\lambda = \alpha + i\beta$ mit $\beta \neq 0$ und somit gilt $(X-\lambda)(X-\overline{\lambda}) \mid P_f$ über $\mathbb{C}$. Außerdem ist $(X-\lambda)(X-\overline{\lambda}) = (X-\alpha^2)+\beta^2 \in \mathbb{R}[X]$ und $\sharp\mathrm{Jor}(\lambda,\cdot,j) = \mathrm{Jor}(\overline{\lambda},\cdot,j)$.

Satz 5.15. (Struktursatz für endlichdimensionale $\mathbb{R}$-Vektorräume)

Sei $\lambda = \alpha + i\beta$ mit $\alpha, \beta \in \mathbb{R}$, $\beta \neq 0$. Dann gilt:

$$\underbrace{\left(\begin{array}{cccc|cccc} \lambda & & & & & & & \\ 1 & \lambda & & & & & & \\ & \ddots & \ddots & & & & & \\ & & 1 & \lambda & & & & \\ \hline & & & & \overline{\lambda} & & & \\ & & & & 1 & \overline{\lambda} & & \\ & & & & & \ddots & \ddots & \\ & & & & & & 1 & \overline{\lambda} \end{array}\right)}_{=\mathrm{diag}(\mathrm{Jor}(\lambda,1,j),\mathrm{Jor}(\overline{\lambda},1,j))\in\mathbb{C}^{2j\times 2j}} \overset{\text{über } \mathbb{C}}{\approx} \underbrace{\left(\begin{array}{cc|cc|c} \alpha & -\beta & & & \\ \beta & \alpha & & & \\ \hline 1 & 0 & \alpha & -\beta & \\ 0 & 1 & \beta & \alpha & \\ \hline & & \ddots & \ddots & \\ & & & \ddots & \ddots \\ \hline & & & 1 & 0 \quad \alpha \quad -\beta \\ & & & 0 & 1 \quad \beta \quad \alpha \end{array}\right)}_{=:\mathrm{Jor}((\alpha,\beta),1,j)\in\mathbb{R}^{2j\times 2j}},$$

d.h. $\mathrm{diag}(\mathrm{Jor}(\alpha+i\beta,1,j),\mathrm{Jor}(\alpha-i\beta,1,j))$ ist die Jordansche Normalform zu $\mathrm{Jor}((\alpha,\beta),1,j)$.

Insbesondere ist $\mathrm{Jor}((\alpha,\beta),1,j)$ genau dann diagonalisierbar, wenn $j = 1$ gilt.

Beweis.

Setze $B = \mathrm{Jor}((\alpha,\beta),1,j)$. Dann ist $\det(B) = ((X-\alpha)^2+\beta^2)^j = (X-\lambda)^j(X-\overline{\lambda})^j$. Wir definieren $C = B - \lambda$ und $D = B - \overline{\lambda}$ und setzen $v = (i;1;0;\cdots;0)$, $w = (1;i;0;0;\cdots;0;0)$. Außerdem wählen wir $J = \mathrm{Jor}((\alpha,\beta),1,j)$. Dann gelten:

$$\begin{aligned} v &= (i;1;0;0;0;0;\cdots;0;0), & w &= (1;i;0;0;0;0;\cdots;0;0),\\ Cv &= (0;0;i;1;0;0;\cdots;0;0), & Dw &= (0;0;1;i;0;0;\cdots;0;0),\\ C^{j-1}v &= (0;0;\cdots;0;0;0;0;i;1), & D^{j-1}w &= (0;0;\cdots;0;0;0;0;1;i),\\ C^{j}v &= (0;0;\cdots;0;0;0;0;0;0), & D^{j}w &= (0;0;\cdots;0;0;0;0;0;0). \end{aligned}$$

Folglich sind $\mathrm{Ann}(v) = K[X](X-\lambda)^j$ und $\mathrm{Ann}(w) = K[X](X-\overline{\lambda})^j$, außerdem sind $\mathfrak{V} = (v, Cv, ..., C^{j-1}v)$ eine $\mathbb{C}$-Basis von $K[X]v$ und $\mathfrak{W} = (w, Dw, ..., D^{j-1}w)$ eine $\mathbb{C}$-Basis von $K[X]w$. Damit ist

$$\mathbb{C}^{2j} = K[X]v \oplus K[X]w = \bigoplus_{i=0}^{j-1} \mathbb{C}C^i v \oplus \bigoplus_{i=0}^{j-1} \mathbb{C}D^j w.$$

Konkret gilt mit der $\mathbb{C}$-Basis $\mathfrak{V}' = (\mathfrak{V}, \mathfrak{W})$ des $\mathbb{C}^{2j}$, und der Basistransformation zwischen $\mathfrak{E}$ und $\mathfrak{V}'$, $Q = \mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{V}'}(\mathrm{Id}) = (\mathfrak{v}_1, ..., \mathfrak{v}_j, \mathfrak{w}_1, ..., \mathfrak{w}_j) \in \mathbb{C}^{2j\times 2j}$, dass $QJQ^{-1} = B$. □

Bemerkung 5.16.

Sind $A, B \in \mathbb{R}^{n\times n}$ ähnlich über $\mathbb{C}$, dann auch über $\mathbb{R}$. Damit erhalten wir: ◊

Satz 5.17. (Normalform für Darstellungsmatrizen über $\mathbb{R}$)

Es existieren $\lambda_1, ..., \lambda_r$ und $\alpha_1, ..., \alpha_s, \beta_1, ..., \beta_s \in \mathbb{R}$, alle $\beta_j \neq 0$, so dass alle $\lambda_j, \alpha_j + i\beta_j, \alpha_j - i\beta j$ Eigenwerte von f sind; weiter existieren Basen $\mathfrak{V}_1, ..., \mathfrak{V}_r, \mathfrak{W}_1, ..., \mathfrak{W}_s$ zu f-invarianten Unterräumen $V_1, ..., V_r, W_1, ..., W_s$ von V, so dass sich V zerlegen lässt in $V = V_1 \oplus \cdots \oplus V_r \oplus W_1 \oplus \cdots \oplus W_s$ und f bzgl. der Basis $\mathfrak{V} = (\mathfrak{V}_1, ..., \mathfrak{V}_r, \mathfrak{W}_1, ..., \mathfrak{W}_s)$ die Darstellung

$$\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f) = \mathrm{diag}(\mathrm{Mat}^{\mathfrak{V}_1}_{\mathfrak{V}_1}(f|_{V_1}), ..., \mathrm{Mat}^{\mathfrak{V}_r}_{\mathfrak{V}_r}(f|_{V_r}), \mathrm{Mat}^{\mathfrak{W}_1}_{\mathfrak{W}_1}(f|_{W_1}), ..., \mathrm{Mat}^{\mathfrak{W}_r}_{\mathfrak{W}_r}(f|_{W_r}))$$

besitzt, wobei alle $\mathrm{Mat}^{\mathfrak{V}_j}_{\mathfrak{V}_j}(f|_{V_j}) = \mathrm{Jor}(\lambda_j, \cdot, \cdot)$ und $\mathrm{Mat}^{\mathfrak{W}_j}_{\mathfrak{W}_j}(f|_{W_j}) = \mathrm{Jor}((\alpha_j, \beta_j), \cdot, \cdot)$ Jordanblöcke sind. $\mathrm{Mat}^{\mathfrak{V}}_{\mathfrak{V}}(f)$ heißt die reelle Normalform von f.

Beispiel 5.18. (zur Berechnung der reellen Normalform)

Wir wollen die Matrix

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

auf reelle Normalform transformieren. Wir gehen dabei in zwei Schritten vor: Zuerst transformieren wir über $\mathbb{C}$ auf Jordansche Normalform, dann fassen wir Jordankästchen zueinander konjugierter Eigenwerte zu reellen Blöcken zusammen.

1. Das charakteristische Polynom von A ist $\chi_A(X) = (X-3)(X-(1+2i))^2(X-(1-2i))^2$. Somit gibt es zwei Möglichkeiten für die zugehörige Jordansche Normalform über $\mathbb{C}$:

$$\begin{pmatrix} 3 & & & & \\ & \lambda & & & \\ & 1 & \lambda & & \\ & & & \overline{\lambda} & \\ & & & 1 & \overline{\lambda} \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} 3 & & & & \\ & \lambda & & & \\ & & \overline{\lambda} & & \\ & & & \lambda & \\ & & & & \overline{\lambda} \end{pmatrix},$$

   wobei $\lambda = \alpha + i\beta = 1 + i2$, und somit auch zwei Möglichkeiten für die reelle Normalform:

$$\begin{pmatrix} 3 & & & & \\ & 1 & -2 & & \\ & 2 & 1 & & \\ & 1 & 0 & 1 & -2 \\ & 0 & 1 & 2 & 1 \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} 3 & & & & \\ & 1 & -2 & & \\ & 2 & 1 & & \\ & & & 1 & -2 \\ & & & 2 & 1 \end{pmatrix},$$

   d.h. es gelten $r = 1$ und entweder $s = 1$ oder $s = 2$.

2. Wir berechnen den eindimensionalen Eigenraum $\mathrm{Kern}(A-3)$ zum Eigenwert 3:

$$A - 3 \mapsto \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \Longrightarrow \quad \mathrm{Kern}(A-3) = \langle 1; 1; 0; 0; 0 \rangle.$$

   Somit sind $\mathfrak{v}_1 = (1; 1; 0; 0; 0)$ und $\mathfrak{V}_1 = (\mathfrak{v}_1)$.

3. Wir berechnen die Kerne zu $(A-\lambda)^j$, $j = 1, ..., d$, so dass d minimal mit $\mathrm{Kern}(A-\lambda)^{d-1} = \mathrm{Kern}(A-\lambda)^d$:

$$A - \lambda \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & i \\ 0 & 0 & 0 & 1 & -i \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \Longrightarrow \quad \mathrm{Kern}(A-\lambda) = \langle (0; 0; -i; i; 1) \rangle.$$

Wegen $\dim \mathrm{Kern}(A-\lambda) = 1$ ist A nicht über $\mathbb{C}$ diagonalisierbar, d.h. die erste der beiden in (1) vorgestellten Varianten trifft zu und es gelten $s = 1$, $d = 2$.

$$(A-\lambda)^2 \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -i & 0 & 1 \\ 0 & 0 & 0 & 1 & -i \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \Longrightarrow \qquad \mathrm{Kern}(A-\lambda)^2 = \left\langle \begin{matrix} (0;i;1;0;0), \\ (0;-1;0;i;1) \end{matrix} \right\rangle .$$

Wir ergänzen $((0;0;-i;i;1))$ durch $\mathfrak{w}_1 = (0;i;1;0;0)$ zu einer Basis von $\mathrm{Kern}(A-\lambda)^2$ und berechnen $\mathfrak{w}_2 = (A-\lambda)\mathfrak{w}_2 = (0;0;-i;i;1)$..

4. Die entsprechenden Basisvektoren zum konjugierten Problem lauten $\mathfrak{w}_3 = \overline{\mathfrak{w}}_1 = (0;-i;1;0;0)$ und $\mathfrak{w}_4 = \overline{\mathfrak{w}}_2 = (0;0;i;-i;1)$. Speziell ist $\mathfrak{W}_1 = (\mathfrak{w}_1, \mathfrak{w}_2, \mathfrak{w}_3, \mathfrak{w}_4)$.
5. Wir wählen die Basis $\mathfrak{V} = (\mathfrak{V}_1, \mathfrak{W}_1)$ mit Transformation $P = \mathrm{Mat}_{\mathfrak{E}}^{\mathfrak{V}}(\mathrm{Id}) = (\mathfrak{v}_1, \mathfrak{w}_1, \mathfrak{w}_2, \mathfrak{w}_3, \mathfrak{w}_4)$, dann ist
$$P^{-1}AP = J = \begin{pmatrix} 3 & & & & \\ & \lambda & & & \\ & 1 & \lambda & & \\ & & & \overline{\lambda} & \\ & & & 1 & \overline{\lambda} \end{pmatrix}.$$
die Jordansche Normalform von A.
6. Zur Bestimmung der Transformationsmatrix auf reelle Normalform sind keine weiteren Berechnungen nötig; es ist
$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & i & 0 & 1 & 0 \\ 0 & 1 & 0 & i & 0 \\ 0 & 0 & i & 0 & 1 \\ 0 & 0 & 1 & 0 & i \end{pmatrix} \qquad \Longrightarrow \qquad QJQ^{-1} = \begin{pmatrix} 3 & & & & \\ & 1 & -2 & & \\ & 2 & 1 & & \\ & 1 & 0 & 1 & -2 \\ & 0 & 1 & 2 & 1 \end{pmatrix}.$$
Somit besitzt $Q(P^{-1}AP)Q^{-1}$ reelle Normalform; die zugehörige Basis dieser Darstellung entspricht den Spalten von QP^{-1}.

6. Unendlichdimensionale Vektorräume und lineare Operatoren

6.1. Dualraum und Bidualraum

Bemerkung 6.1.

Ab jetzt seien stets K ein Körper und V ein K-Vektorraum. Dieser muss nicht mehr notwendig endlichdimensional sein. $\mathbb{K}$ bezeichne einen der Körper $\mathbb{R}, \mathbb{C}$. ◊

Definition 6.2.

$V^* = \mathrm{Hom}_K(V, K)$ bezeichne die Menge der Linearformen über K, d.h. der K-linearen Abbildungen von V nach K.

V^* ist ein K-Vektorraum und wird als der algebraische Dualraum zu V bezeichnet. $V^{**} = (V^*)^*$ heißt der algebraische Bidualraum zu V.

Bemerkung 6.3.

Addition und Multiplikation in V^* sind wie üblich definiert durch

$$(f+g)(v) = f(v) + g(v), \qquad (\alpha f)(v) = \alpha f(v)$$

für alle $f, g \in V^*$ und alle $\alpha \in K$, $v \in V$. ◊

Satz 6.4.

Ist $\dim_K(V) = n$, dann auch $\dim_K(V^*) = n$.

Beweis.

Sei $\mathfrak{V} = (v_1, ..., v_n)$ eine Basis von V. Wir setzen $v_i^* : V \to K$, $v_i^*(v_j) = \delta_{ij}$, d.h. $v_i^*(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_i$.

1. $\{v_1^*, ..., v_n^*\}$ ist ein Erzeugendensystem von V^*: Sei $f \in V^*$ beliebig, dann

$$f = f(v_1)v_1^* + \cdots + f(v_n)v_n^* \qquad \Longrightarrow \qquad f \in \langle v_1^*, ..., v_n^* \rangle.$$

2. $\{v_1^*, ..., v_n^*\}$ ist linear unabhängig: Seien $\beta_1, ..., \beta_n \in K$ und $j \in \{1, ..., n\}$, dann gilt

$$0 = \beta_1 v_1^* + \cdots + \beta_n v_n^* \qquad \Longrightarrow \qquad 0 = 0(v_j) = \beta_j v_j^*(v_j) = \beta_j.$$

Also ist $\mathfrak{V}^* = (v_1^*, ..., v_n^*)$ eine K-Basis von V^*. □

Bemerkung 6.5.

1. Die duale Paarung $\sigma : V \times V^* \to K$ mit $\sigma(v, f) = f(v)$ ist bilinear, d.h. linear in beiden Komponenten, denn für alle $f, g \in V^*$ und alle $\alpha, \beta \in K$ gelten

$$\begin{aligned} \sigma(\alpha v + \beta w, f) &= f(\alpha v + \beta w) = \alpha f(v) + \beta f(w) = \alpha\sigma(v, f) + \beta\sigma(w, f), \\ \sigma(v, \alpha f + \beta g) &= (\alpha f + \beta g)(v) = \alpha f(v) + \beta g(v) = \alpha\sigma(v, f) + \beta\sigma(v, g). \end{aligned}$$

2. Sei nun $v_0 \in V$ fest. Dann ist der Auswertungsoperator $\sigma_0 : V^* \to K$, $\sigma_0(f) = f(v_0)$ eine K-lineare Abbildung, also $\sigma_0 \in V^{**}$. Es gilt: $\sigma_0(\cdot) = \sigma(v_0, \cdot)$. ◊

Satz 6.6. (Kanonische Einbettung in den algebraischen Bidual)

V lässt sich in seinen Bidualraum einbetten via $\Phi : V \to V^{**}$, $\Phi(v_0) = \sigma_0$.

Beweis.

1. Φ ist K-linear: Seien $v, w \in V$, $\alpha, \beta \in K$. Dann gilt für alle $f \in V^*$:

$$\begin{aligned} \Phi(\alpha v + \beta w)(f) &= f(\alpha v + \beta w) \\ &= \alpha f(v) + \beta f(w) \\ &= \alpha\Phi(v)(f) + \beta\Phi(w)(f) \\ &= (\alpha\Phi(v) + \beta\Phi(w))(f). \end{aligned}$$

2. Φ ist injektiv: Sei $v \neq 0$. Wir zeigen, dass v nicht im Kern von Φ liegt, d.h. dass ein $f \in V^*$ existiert mit $\Phi(v)(f) = f(v) = 0$. Ergänze nämlich (v_0) durch $\mathfrak{V}' \subseteq V$ zu einer Basis $\mathfrak{V}$ von V und wähle $f \in V^*$ mit $f(v_0) = 1$ und $f(v') = 0$ für alle $v' \in \mathfrak{V}'$. □

Bemerkung 6.7.

1. Ist V endlichdimensional, dann folgt aus $\dim V = \dim V^* = \dim V^{**}$, dass Φ ein Isomorphismus ist.
2. Umgekehrt ist Φ niemals surjektiv, wenn V ein unendlichdimensionaler Vektorraum ist:

   Sei $\mathfrak{V}$ eine Basis von V, dann ist $\mathfrak{V}^*$ keine Basis von V^*, da $\varphi \in V^*$ mit $\varphi(v) = 1$ für alle $v \in \mathfrak{V}$ nicht als Linarkombination von Elementen aus $\mathfrak{V}^*$ dargestellt werden kann. Wir ergänzen $\mathfrak{V}^* \cup (\varphi)$ zu einer Basis $\mathfrak{V}'$ von V^*. Dann liegt $\phi \in V^{**}$ mit $\phi(\varphi) = 1$ und $\phi(v^*) = 0$ für alle $v^* \in \mathfrak{V}' \backslash \{\varphi\}$ nicht in $\Phi(V)$.
3. Wir fassen V als Teilmenge von V^{**} auf via der Identifikation von v mit $\Phi(v)$. ◊

6.2. Normierte Vektorräume

Definition 6.8.

$(V, ||\cdot||)$ heißt ein normierter Vektorraum, falls $||\cdot||$ eine Norm auf V definiert, d.h. falls für alle $x, y \in V$ und alle $\alpha \in \mathbb{K}$ gelten:

1. $||x|| \geq 0$ und $||x|| = 0 \Rightarrow x = 0$, (positive Definitheit)
2. $||\alpha x|| = |\alpha|\,||x||$, (Multiplikativität)
3. $||x + y|| \leq ||x|| + ||y||$. (Dreiecksungleichung)

Seien nun $(V, ||\cdot||_V)$ und $(W, ||\cdot||_W)$ normierte Vektorräume. $f \in \mathrm{Hom}(V, W)$ heißt beschränkt, falls eine Konstante $c > 0$ existiert, so dass für alle $x \in V$ gilt: $||fx||_W \leq c||x||_V$.

Bemerkung 6.9.

Sei $(V, ||\cdot||)$ ein normierter $\mathbb{K}$-Vektorraum. Dann gelten:

1. Die Addition $\mathrm{Add} : V \times V \to V$, $(v, w) \mapsto v + w$, ist gleichmäßig stetig.

   Dabei ist der Produktraum $V \times V$ mit der kanonischen Produktnorm $||(x, y)||_{V\times V} = ||x||_V + ||y||_V$ versehen.

   Seien nämlich $z = (x, y) \in V \times V$ und $c = (a, b) \in V \times V$. Weiter sei $\epsilon > 0$ gegeben. Wir müssen ein $\delta > 0$ finden, so dass aus $||z - c|| < \delta$ folgt: $||\mathrm{Add}(z) - \mathrm{Add}(c)|| < \epsilon$. Wir wählen $\delta = \epsilon$, d.h. $||x-a|| + ||y-b|| < \epsilon$. Dann ist auch $||(x+y)-(a+b)|| = ||(x-a)+(y-b)|| \leq ||x-a|| + ||y-b|| < \epsilon$. Da δ unabhängig von z, c ist, folgt daraus die gleichmäßige Stetigkeit von Add.

2. Die Skalarmultiplikation $\mathrm{Mul} : \mathbb{K} \times V \to V$, $(\alpha, v) \mapsto \alpha v$ ist stetig.

   Dabei ist die Produkttopoligie von $\mathbb{K} \times V$ gegeben durch $||(\alpha, v)|| = |\alpha| + ||v||$.

   Seien $\alpha, \beta \in \mathbb{K}$, $x, y \in V$ und $\epsilon > 0$. Wir müssen ein $\delta > 0$ finden, so dass für $|\alpha - \beta| + ||x - y|| < \delta$ gilt: $||\alpha x - \beta y|| < \epsilon$. Allgemein ist $||\alpha x - \beta y|| = ||\alpha x - \beta x + \beta x - \beta y|| \leq |\alpha - \beta|\,||x|| + |\beta|\,||x - y||$. Der Fall $x = 0$ und $\beta = 0$ ist trivial. Andernfalls wähle $\delta = \frac{\epsilon}{2}\min(\frac{1}{||x||}, \frac{1}{|\beta|})$.

3. Sei $\alpha \in \mathbb{K}^\times$. Dann ist die Skalierung $\mathrm{Scl} : V \to V$ mit $v \mapsto \alpha v$ gleichmäßig stetig.

   Seien nämlich $x, y \in V$, dann ist $||\alpha x - \alpha x|| \leq |\alpha|\,||x - y|| < \epsilon$, falls $||x - y|| < \frac{\epsilon}{|\alpha|}$.

4. Die Normabbildung $||\cdot|| : V \to \mathbb{K}$ ist gleichmäßig stetig:

   Seien $x, y \in V$ und $\delta = \epsilon$, dann ist $|\,||x|| - ||y||\,| \leq ||x - y|| < \epsilon$, falls $||x - y|| < \delta$. ◊

Lemma 6.10.

Seien V, W normierte $\mathbb{K}$-Vektorräume und $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. Dann sind äquivalent:

1. f ist stetig in 0, d.h. für alle $v \in V$, $\epsilon > 0$ gibt es ein $\delta > 0$ mit $||v|| < \delta \Rightarrow ||fv|| < \epsilon$.
2. f ist stetig auf V, d.h. für alle $v, w \in V$, $\epsilon > 0$ gibt es ein $\delta > 0$ mit $||v - w|| < \delta \Rightarrow ||fv - fw|| < \epsilon$.
3. f ist gleichmäßig stetig auf V, d.h. für alle $\epsilon > 0$ existiert ein $\delta > 0$, so dass für alle $v, w \in V$ gilt $||v - w|| < \delta \Rightarrow ||fv - fw|| < \epsilon$.
4. f ist beschränkt, d.h. es gibt ein $c > 0$, so dass für alle $v \in V$ gilt: $||fv|| \leq c||v||$.

Beweis.

1. Sei f stetig in 0. Wähle $\epsilon = 1$ und dazu ein passendes δ. Setze $c = \frac{2}{\delta}$ und wähle $v \in V\backslash\{0\}$ beliebig. Dann ist $\frac{v}{c||v||} < \delta$, d.h. $\frac{1}{c||v||}||fv|| = ||f\frac{v}{c||v||}|| < 1$, d.h. $||fv|| < c||v||$. Also ist f beschränkt.
2. Sei f beschränkt mit Schranke $c > 0$. Zu $\epsilon > 0$ setze $\delta = \frac{\epsilon}{c}$. Dann gilt für alle $v, w \in V$ mit $||v-w|| < \delta$, dass $||fv - fw|| = ||f(v - w)|| \leq c||v - w|| < \epsilon$. □

6.3. Der Satz von Hahn-Banach

Definition 6.11.

Eine Abbildung $p : V \to \mathbb{R}$ heißt sublinear, falls für alle $x, y \in V$ und $\alpha \geq 0$ gilt:

$$p(\alpha x) = \alpha p(x), \qquad p(x+y) \leq p(x) + p(y).$$

Bemerkung 6.12.

Speziell ist auf jedem normierten Vektorraum die Norm selbst ein sublineares Funktional. ◊

Satz 6.13. (Hahn-Banach)

Sei p ein sublineares Funktional auf einem normierten $\mathbb{R}$-Vektorraum V. Weiter seien U ein Untervektorraum von V und $f \in U^*$ mit $f(x) \leq p(x)$ für alle $x \in U$. Dann lässt sich f zu einem $F \in V^*$ fortsetzen, so dass gelten $f(x) \leq p(x)$ für alle $x \in V$ und $F(x) = f(x)$ für alle $x \in U$.

Beweis.

Wir definieren die Menge

$$\mathcal{M} = \{(W, g) \mid U \subseteq W \subseteq V \text{ Unterräume}, g \in W^*, \ g|_U = f \text{ und } g(w) \leq p(w) \text{ für alle } w \in W\}.$$

Auf $\mathcal{M}$ definieren wir die partielle Ordnung

$$(W, g) \preccurlyeq (W', g') \qquad :\Longleftrightarrow \qquad W \subseteq W' \text{ und } g'|_W = g.$$

1. $\mathcal{M}$ besitzt ein maximales Element: Sei $((W_i, g_i))_{i \in I}$ eine Kette in $\mathcal{M}$, d.h. für alle $i, j \in I$ gelte $(W_j, g_j) \preccurlyeq (W_i, g_i)$ oder $(W_i, g_i) \preccurlyeq (W_j, g_j)$. Wir setzen $W_I = \bigcup W_i$ und $g_I : W \to \mathbb{R}$ mit $g_I(x) = g_i(x)$ für ein beliebiges $i \in I$ mit $x \in W_i$. Dann ist g_I wohldefiniert und ein Element aus W_I^*, außerdem ist W_I ein Unterraum mit $U \subseteq W_I \subseteq V$. Insbesondere ist $(W_I, g_I) \in \mathcal{M}$ eine obere Schranke der Kette $((W_i, g_i))_{i \in I}$. Mit Zorns Lemma folgt: Es gibt ein maximales Element (W, g) in $\mathcal{M}$.

2. Es gilt $W = V$, dann ist $F = g$ eine Fortsetung von f mit den gewünschten Eigenschaften. Angenommen, $W \subsetneq V$, etwa $z \in V \backslash W$. Setze $W_0 = W \oplus \mathbb{R}z \supsetneq W$ und definiere $g_\lambda \in W_0^*$ durch $g(w + \rho z) = g(w) + \rho\lambda$ mit einem gewissen $\lambda = g_\lambda(z) \in \mathbb{R}$. Wir weisen nach, dass ein $\lambda_0 \in \mathbb{R}$ existiert mit $g_{\lambda_0}(v) \leq p(v)$ für alle $v \in W_0$. In dem Fall wäre nämlich $(W_0, g_{\lambda_0}) \in \mathcal{M}$ mit $(W, g) \preccurlyeq (W_0, g_{\lambda_0})$, was der Maximalität von (W, g) widerspräche.

   Genau dann gilt $g_\lambda(v) \leq p(v)$ für alle $v \in W_0$, wenn $g(w) + \rho\lambda \leq p(w + \rho z)$ für alle $w \in W$ und alle $\rho \in \mathbb{R}^\times$, d.h. wenn $g(\frac{1}{\rho}w) + \lambda \leq p(\frac{1}{\rho}w + z)$ für alle $w \in W$, $\rho > 0$ und $g(-\frac{1}{\rho}w) - \lambda \leq p(-\frac{1}{\rho}w - z)$ für alle $w \in W$, $\rho < 0$. Nun gilt für alle $w_1, w_2 \in W$, dass

   $$g(w_1) + g(w_2) = g(w_1 + w_2) \leq p(w_1 + w_2) = p(w_1 - z + w_2 + z) \leq p(w_1 - z) + p(w_2 + z),$$

   d.h. $g(w_1) - p(w_1 - z) \leq g(w_2) + p(w_2 + z)$. Also gilt für die Mengen $W_1 = \{g(w_1) - p(w_1 - z) \mid w_1 \in W\}$ und $W_2 = \{g(w_2) - p(w_2 - z) \mid w_2 \in W\}$, dass $W_1 \leq W_2$. Wir wählen λ mit $W_1 \leq \lambda \leq W_2$ und setzen $w_1 = -\frac{1}{\rho}w$ und $w_2 = \frac{1}{\rho}w$. Dann folgten $\lambda \leq -g(\frac{1}{\rho}w) + (\frac{1}{\rho}w + z)$ und $g(-\frac{1}{\rho}w) - p(-\frac{1}{\rho}w - z) \leq \lambda$, was den erwarteten Widerspruch induziert. □

Korollar 6.14. (Hahn-Banach)

Seien V ein normierter $\mathbb{K}$-Vektorraum, U ein Untervektorraum von V und $f \in U^*$ beschränkt. Dann lässt sich f zu einem stetigen $F \in V^*$ fortsetzen.

Beweis.

1. Sei $\mathbb{K} = \mathbb{R}$. Sei $c > 0$ mit $f(u) \leq c||u||$ für alle $u \in U$. Dann ist $p : V \to \mathbb{R}$, $p(v) = c||v||$ sublinear. Nach dem Satz von Hahn-Banach gibt es ein $F \in V^*$ mit $F(x) \leq p(x) = c||x||$, d.h. F ist beschränkt.

2. Sei $\mathbb{K} = \mathbb{C}$. Wähle $g, h : U \to \mathbb{R}$ mit $f(u) = g(u) + ih(u)$, d.h. $g = \mathrm{Re}(f)$ und $h = \mathrm{Im}(f)$. Dann sind g, h linear über $\mathbb{R}$, denn für alle $u, v \in V$ und alle $\alpha, \beta \in \mathbb{R}$ ist

$$\alpha(g(u) + ih(u)) + \beta(g(u) + ih(v)) = \alpha f(u) + \beta f(v) = f(\alpha u + \beta v) = g(\alpha u + \beta v) + ih(\alpha u + \beta v),$$

d.h. $g(\alpha u + \beta v) = \alpha g(u) + \beta g(v)$ und $h(\alpha u + \beta v) = \alpha h(u) + \beta h(v)$. Wegen

$$|g(u)| \leq |g(u) + ih(u)| = |f(u)| \leq c||u||$$

ist g stetig. Also lässt sich g zu $G : V \to \mathbb{R}$ fortsetzen, das stetig und $\mathbb{R}$-linear ist. Setze $F : V \to \mathbb{R}$ mit $F(x) = G(x) - iG(ix)$. Dann gelten:

a) F ist $\mathbb{C}$-linear, denn seien $\alpha, \beta \in \mathbb{R}$ und $x, y \in V$. Dann gilt:

$$\begin{aligned} F(x + (\alpha + i\beta)y) &= G(x) + \alpha G(y) + \beta G(iy) - i(G(ix) + \alpha G(ix) - \beta G(y)) \\ &= F(x) + \alpha F(y) + i\beta F(y) = F(y) + (\alpha + i\beta)F(y). \end{aligned}$$

b) F ist eine Fortsetzung von f, denn für alle $u \in U$ gilt:

$$g(iu) + ih(iu) = f(iu) = if(u) = ig(u) - h(u) \qquad \Longrightarrow \qquad g(iu) = -h(u),$$

d.h. $f(u) = g(u) + ih(u) = g(u) + ig(iu)$.

c) F ist stetig, denn zu beliebigem $x \in V$ wähle $\theta \in [0, \pi]$, $r \geq 0$ mit $x = re^{i\theta}$. Wegen $|e^{i\theta}| = 1$ folgt:

$$|F(x)| = r = e^{-i\theta}F(x) = F(e^{-i\theta}x) = G(e^{-i\theta}x) + 0 \leq c_G||e^{-i\theta}x|| = c_G|e^{-i\theta}|\,||x|| = c_G||x||. \quad \square$$

6.4. Stetige Operatoren

Definition 6.15.

Seien V, W normierte $\mathbb{K}$-Vektorräume. $\mathscr{L}(V, W) = \{f \in \mathrm{Hom}_{\mathbb{K}}(V, W) \mid f \text{ stetig}\}$ bezeichne den $\mathbb{K}$-Vektorraum der stetigen, $\mathbb{K}$-linearen Abbildungen von V nach W. Die Elemente von $\mathscr{L}(V, W)$ werden als stetige Operatoren bezeichnet.

$V' = \mathscr{L}(V, \mathbb{K}) = \{f \in V^* \mid f \text{ ist stetig}\} \subseteq V^*$ heißt der topologische Dualraum von V und $V'' = (V')' = \mathscr{L}(V', \mathbb{K})$ der topologische Bidualraum von V.

Bemerkung 6.16.

1. Lineare Operatoren werden von jetzt an mit Großbuchstaben bezeichnet. Weiter schreiben wir Ax für $A(x)$.
2. Ist $A \in \mathscr{L}(V, W)$ stetig, dann setzen wir $||A|| = \{c > 0 \mid \forall x \in V : ||Ax|| \leq c||x||\}$. Speziell ist $||Ax|| \leq ||A||\,||x||$ für alle $x \in V$. Es gilt:

$$||A|| = \sup_{||x|| \leq 1} ||Ax|| = \sup_{||x|| = 1} ||Ax|| = \sup_{x \in V \setminus \{0\}} \frac{||Ax||}{||x||}. \qquad \lozenge$$

Satz 6.17.

$||\cdot||$ definiert eine Norm auf $\mathscr{L}(V, W)$, die als Operatornorm bezeichnet wird.

Beweis.

1. Für alle $A, B \in \mathscr{L}(V, W)$ und beliebiges $x \in V$ gilt

$$\begin{aligned} ||(A + B)x||_W = ||Ax + Bx||_W &\leq ||Ax||_W + ||Bx||_W \\ &\leq ||A||\,||x||_V + ||B||\,||x||_V = (||A|| + ||B||)||x||_V, \end{aligned}$$

d.h. $||A + B|| \leq ||A|| + ||B||$.

2. Für alle $A \in \mathscr{L}(V,W)$, $\lambda \in \mathbb{K}$ und beliebiges $x \in V$ gilt

$$||(\lambda A)x||_W = ||\lambda Ax||_W = |\lambda|\,||Ax||_W \leq |\lambda|\,||A||\,||x||_V,$$

d.h. $||\lambda A|| \leq |\lambda|\,||A||$. Für $\lambda = 0$ gilt offenbar auch die umgekehrte Abschätzung; für $\lambda \neq 0$ ist $||\frac{1}{\lambda}\lambda A|| \leq |\frac{1}{\lambda}|\,||\lambda A||$, d.h. auch $|\lambda|\,||A|| \leq ||\lambda A||$. Insgesamt folgt also $||\lambda A|| = |\lambda|\,||A||$.

3. $||A|| \geq 0$ ist per Definition erfüllt. Sei nun $||A|| = 0$, dann $||A\frac{1}{||x||}x||_W = 0$ für alle $x \neq 0$, d.h. $||Ax||_W = 0$ für alle $x \in V$. Da $||\cdot||_W$ definit ist, erhalten wir daraus $Ax = 0$ für alle $x \in V$, d.h. $A = 0$ in $\mathscr{L}(V,W)$. □

Korollar 6.18. (Hahn-Banach)

Seien V normiert und $x_0 \in V$. Dann gibt es ein $F \in V'$ mit $F(x_0) = ||x_0||$ und $||F|| \leq 1$.

Beweis.

Œ Gelte $x_0 \neq 0$, sonst setze $F = 0$. Sei $U = \mathbb{K}x_0 \subseteq V$. Dann ist $f \in U^*$, gegeben durch $f(\lambda x_0) = \lambda||x_0||$, stetig, denn $|f(\lambda x_0)| = ||\lambda x_0||$, also $f \in U'$ mit $||f|| = 1$. Dann existiert eine Fortsetzung $F \in V'$ von f mit $|F(x)| \leq ||f||\,||x|| = ||x||$ für alle $x \in V$, d.h. $||F|| \leq 1$. □

Bemerkung 6.19.

Es wurde dabei benutzt, dass für die Fortsetzung $F \in V'$ von $f \in U'$ stets gilt: $||F|| \leq ||f||$. ◊

Korollar 6.20. (Normformel)

Seien V ein normierter $\mathbb{K}$-Vektorraum und $x_0 \in V$. Dann gilt:

$$||x_0|| = \max\{|F(x_0)| \mid F \in V' \text{ mit } ||F|| \leq 1\}.$$

Beweis.

Für alle $F \in V'$ mit $||F|| \leq 1$ gilt $|F(x_0)| \leq ||F||\,||x_0|| = ||x_0||$. Speziell für dasjenige $F \in V'$ mit $||F|| = 1$ und $F(x_0) = ||x_0||$, das nach dem letzten Korollar existiert, gilt Gleichheit. □

Satz 6.21. (Kanonische Einbettung in den topologischen Bidual)

Die Abbildung

$$\Phi : V \to V'', \qquad \Phi(v)v' = v'v$$

ist eine isometrische Einbettung, d.h. Φ ist ein stetiger, injektiver $\mathbb{K}$-Vektorraumhomomorphismus mit $||\Phi(v)||_{V''} = ||v||_V$ für alle $v \in V$.

Beweis.

1. Φ ist wohldefiniert, d.h. $\Phi(v) \in V''$: Seien $v \in V$ und $v' \in V'$, dann ist $|\Phi(v)v'| = |v'v| \leq ||v'||\,||v||$, d.h. $\Phi(v)$ ist stetig mit $||\Phi(v)|| \leq ||v||$.

2. Φ ist ein Homomorphismus: Seien $v, w \in V$ und $\alpha, \beta \in \mathbb{K}$, dann ist für beliebiges $v' \in V'$

$$\Phi(\alpha v + \beta w)v' = v'(\alpha v + \beta w) = \alpha v'v + \beta v'w = (\alpha\Phi(v) + \beta\Phi(w))v'.$$

Nach Hahn-Banach existiert zu jedem $v \in V$ ein $v' \in V'$ mit $|\Phi(v)v'| = |v'v| = ||v||$, d.h. $||\Phi(v)|| = ||v||$. Also ist Φ beschränkt mit $||\Phi|| = 1$ und isometrisch.

Damit ist Φ auch injektiv, denn aus $\Phi(v) = 0$ folgt $||\Phi(v)|| = 0$, d.h. $||v|| = 0$ und damit $v = 0$, also $\mathrm{Kern}(\Phi) = \{0\}$. □

Bemerkung 6.22.

1. Ist V endlichdimensional, dann ist Φ wieder ein Isomorphismus.
2. Ist V unendlichdimensional, dann kann Φ surjektiv sein. Das für den algebraischen Dual konstruierte ϕ ist nämlich nicht stetig und somit kein Gegenbeispiel für $\Phi(V) = V''$.
3. Zwischen endlichdimensionalen $\mathbb{K}$-Vektorräumen V, W sind alle $A \in \mathscr{L}(V,W)$ stetig: Seien $\mathfrak{V}, \mathfrak{W}$ beliebige Basen von V, W, dann definieren
$$||A|| = \sum_{i,j} |\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{W}}(A)_{ij}|, \qquad ||A|| = \max_j \sum_i |\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{W}}(A)_{ij}|, \qquad ||A|| = \max_i \sum_j |\mathrm{Mat}_{\mathfrak{V}}^{\mathfrak{W}}(A)_{ij}|$$
Normen auf $\mathscr{L}(V,W)$, die alle zur Operatornorm äquivalent ist (da alle Normen auf $\mathbb{K}^n$ äquivalent sind). ◊

Definition 6.23.

Ist $\Phi : V \to V''$ surjektiv, d.h. ein isometrischer Isomomrphismus, dann heißt V reflexiv.

6.5. Banachräume

Definition 6.24.

Ein normierter Vektorraum $(V, ||\cdot||)$ heißt ein Banachraum, falls V bzgl. $||\cdot||$ vollständig ist, d.h. falls jede Cauchyfolge aus V einen Grenzwert in V besitzt.

Beispiel 6.25.

1. Der Raum der stetigen Funktionen $\mathcal{C}^0([a,b])$, versehen mit der Supremumsnorm $||\cdot||_\infty$.
$$\mathcal{C}^0([a,b]) = \{f \in \mathrm{Abb}([a,b],\mathbb{K}) \mid f \text{ ist stetig}\}, \qquad ||f||_\infty = \sup\{f(x) \mid x \in [a,b]\},$$
ist ein Banachraum. Versehen mit der Integralnorm $||\cdot||_{\mathcal{L}^1}$ ist $\mathcal{C}^0([a,b])$ dagegen nicht vollständig.
2. Ebenso ist der Raum der k-fach differenzierbaren Funktionen $\mathcal{C}^k([a,b])$, versehen mit der Norm $||\cdot||_{\mathcal{C}^k}$,
$$\mathcal{C}^k([a,b]) = \{f \in \mathrm{Abb}([a,b],\mathbb{K}) \mid f, f', ..., f^{(k)} \text{ ist stetig}\}, \qquad ||f||_{\mathcal{C}^k} = ||f||_\infty + \cdots + ||f^{(k)}||_\infty,$$
ein Banachraum.
3. Die Lebesgueschen Integrationsräume $\mathcal{L}^p([a,b])$, versehen mit der Integralnorm $||\cdot||_{\mathcal{L}^p}$,
$$\mathcal{L}^p([a,b]) = \{[f]_\sim \mid f \in \mathrm{Abb}([a,b],\mathbb{K}) \mid ||f||_{\mathcal{L}^p} < \infty\}, \qquad ||f||_{\mathcal{L}^p} = \left(\int_a^b |f(x)|^p \, \mathrm{d}\mu\right)^{\frac{1}{p}},$$
ist für alle $1 \le p < \infty$ ein Banachraum, wobei $f \sim g :\Leftrightarrow \mu\{f \neq g\} = 0$. Ebenso ist $\mathcal{L}^\infty([a,b])$ mit der Supremumsnorm $||\cdot||_\infty$,
$$\mathcal{L}^\infty([a,b]) = \{f \in \mathrm{Abb}([a,b],\mathbb{K}) \mid ||f||_{\mathcal{L}^\infty} < \infty\}, \qquad ||f||_\infty = \operatorname*{ess\,sup}_{x\in[a,b]} |f(x)|$$
ein Banachraum.
4. Auch die Folgenräume ℓ^p, versehen mit der Summennorm $||\cdot||_{\ell^p}$,
$$\ell^p = \{x \in \mathrm{Abb}(\mathbb{N},\mathbb{K}) \mid ||x||_{\ell^p} < \infty\}, \qquad ||x||_{\ell^p} = \left(\sum_{k=1}^\infty |x_k|^p\right)^{\frac{1}{p}},$$
ist für alle $1 \le p < \infty$ ein Banachraum. Ebenso ist ℓ^∞ mit der Supremumsnorm $||\cdot||_\infty$,
$$\ell^\infty = \{x \in \mathrm{Abb}(\mathbb{N},\mathbb{K}) \mid ||x||_{\ell^\infty} < \infty\}, \qquad ||x||_\infty = \sup_{k\in\mathbb{N}} |x_k|$$
ein Banachraum. ◊

Satz 6.26.

Seien V ein normierter Raum und W ein Banachraum. Dann ist auch $\mathscr{L}(V,W)$ ein Banachraum.

Beweis.

Sei $(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{L}(V,W)$ eine Cauchyfolge stetiger Operatoren, d.h. $||A_n - A_m|| \to 0$ für $n,m \to \infty$. Dann folgt für $x \in V$:

$$||A_n x - A_m x|| = ||(A_n - A_m)x|| \leq ||A_n - A_m||\,||x|| \to 0 \text{ für } n,m \to \infty.$$

Also ist $(A_n x)_{n\in\mathbb{N}}$ eine Cauchyfolge in W und da W vollständig ist, existiert $\lim A_n x$ für jedes $x \in V$. Wir definieren $A : V \to W$ durch $Ax = \lim A_n x$. Dann gelten:

1. A ist $\mathbb{K}$-linear, denn für alle $\alpha,\beta \in \mathbb{K}$ und alle $x,y \in V$ gilt:
$$A(\alpha x + \beta y) = \lim_{n\to\infty} A_n(\alpha x + \beta y) = \alpha \lim_{n\to\infty} A_n x + \beta \lim_{n\to\infty} A_n y = Ax + \beta Ay.$$
2. A_n konvergiert in der Operatornorm zu A, d.h. $||A_n - A|| \to 0$ für $n \to \infty$: Sei $\epsilon > 0$, dann gibt es $N \in \mathbb{N}$, so dass für alle $n,m \in \mathbb{N}$ gilt: $||A_n - A_m|| < \epsilon$. Seien $n \geq N$, $m \geq n$ und $x \in V$ mit $||x|| \leq 1$. Dann gilt:
$$\begin{aligned} ||A_n x - Ax|| &\leq ||A_n x - A_m x|| + ||A_m x - Ax|| \\ &\leq ||A_n - A_m|| + ||A_m x - Ax|| \leq \epsilon + ||A_m x - A_x||. \end{aligned}$$
Mit $m \to \infty$ folgt $||A_n x - Ax|| \leq \epsilon$, also $A_n \to A$. Insbesondere ist $||A_N - A|| \leq \epsilon$.
3. Ist $(A_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $\mathscr{L}(V,W)$, dann ist $(||A_n||)_{n\in\mathbb{N}}$ eine Cauchyfolge in $\mathbb{K}$, d.h. es gilt $\lim ||A_n|| = c_0 \in \mathbb{K}$. Also ist A stetig mit $||A|| = c_0$. Damit gilt $A \in \mathscr{L}(V,W)$. □

Bemerkung 6.27.

1. V' und V'' sind Banachräume, da $\mathbb{K}$ vollständig ist.
2. Jeder normierte $\mathbb{K}$-Vektorraum kann (via Φ) in einen vollständigen Vektorraum eingebettet werden.
3. Wir identifizieren wieder V mit seinem Bild $\Phi(V) \subseteq V''$. ◊

Definition 6.28.

$\overline{V}$ bezeichne den Abschluss von V in V'', d.h. den kleinsten vollständigen $\mathbb{K}$-Vektorraum über V.

Bemerkung 6.29.

1. $\overline{V}$ ist vollständig in der Norm von V''.
2. V liegt dicht in $\overline{V}$, denn es gilt $V = \overline{V}$ genau dann, wenn V vollständig ist. ◊

Bemerkung 6.30.

1. ℓ^1 ist ein nicht reflexiver Banachraum mit $(\ell^1)' = \ell^\infty$. Auch ℓ^∞ ist ein nicht reflexiver Banachraum, aber $(\ell^\infty)' \neq \ell^1$.
2. Für $p \in (1,\infty)$ ist $(\ell^p)' = \ell^q$, wobei $\frac{1}{p} + \frac{1}{q} = 1$, und diese Räume sind reflexiv.
3. Seien (X,Σ,μ) ein vollständiger Maßraum und
$$\mathcal{L}^p = \{f \in \mathrm{Abb}(X,\mathbb{K}) \mid f \text{ messbar und } ||f||_p < \infty\}.$$
Für $p \in (1,\infty)$ ist $(\mathcal{L}^p)' = \mathcal{L}^q$, wobei wieder $\frac{1}{p} + \frac{1}{q} = 1$. Diese Räume sind alle reflexiv.
4. Ist (X,Σ,μ) σ-endlich, dann ist auch $(\mathcal{L}^1)' = \mathcal{L}^\infty$, dagegen ist $(\mathcal{L}^\infty)' \neq \mathcal{L}^1$. Die Räume $\mathcal{L}^1$ und $\mathcal{L}^\infty$ sind nicht reflexiv. ◊

6.6. Hilberträume

Definition 6.31.

Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $\langle\cdot,\cdot\rangle : V\times V \to \mathbb{K}$ mit $(x,y)\mapsto\langle x,y\rangle$ heißt ein Skalarprodukt, falls für alle $x,y,z\in V$ und alle $\alpha,\beta\in\mathbb{K}$ gelten:

1. $\langle\alpha x+\beta y,z\rangle = \alpha\langle x,z\rangle+\beta\langle y,z\rangle$ und $\langle x,\alpha y+\beta z\rangle = \overline{\alpha}\langle x,y\rangle+\overline{\beta}\langle x,z\rangle$, (Sesquilinearität)
2. $\langle x,y\rangle = \overline{\langle y,x\rangle}$, (Antisymmetrie)
3. $\langle x,x\rangle\geq 0$ für alle $x\in V$ und $\langle x,x\rangle = 0 \Rightarrow x=0$. (positive Definitheit)

Ein $\mathbb{K}$-Prähilbertraum $(E,\langle\cdot,\cdot\rangle)$ ist ein $\mathbb{K}$-Vektorraum mit einem Skalarprodukt.

$x\mapsto ||x|| = \sqrt{\langle x,x\rangle}$ heißt die von $\langle\cdot,\cdot\rangle$ auf E induzierte Norm.

Ist $(E,\langle\cdot,\cdot\rangle)$ bzgl. $||\cdot||$ ein Banachraum, dann heißt $(E,\langle\cdot,\cdot\rangle)$ ein Hilbertraum.

Bemerkung 6.32.

Sei $(E,\langle\cdot,\cdot\rangle)$ ein Prähilbertraum. Dann gelten:

1. $\forall x,y\in E : |\langle x,y\rangle|\leq ||x||\,||y||$. (Cauchy-Schwarzsche Ungleichung)

   Dabei gilt $|\langle x,y\rangle| = ||x||\,||y||$ genau dann, wenn $\{x,y\}$ $\mathbb{K}$-linear abhängig ist.
2. $\forall x,y\in E : ||x+y||^2+||x-y||^2 = 2||x||^2+2||y||^2$. (Parallelogrammregel)
3. $\forall x,y\in W : \langle x,y\rangle = 0 \Rightarrow ||x+y||^2 = ||x||^2+||y||^2$. (Satz des Pythagoras)

   Wir setzen $x\perp y :\Leftrightarrow \langle x,y\rangle = 0$ und nennen x,y orthogonal
4. $\forall y\in W : \langle\cdot,y\rangle\in E'$ mit $||\langle y,\cdot\rangle|| = ||y||$.

   Wir werden zeigen, dass in einem Hilbertraum zu jedem $A\in E'$ genau ein $y\in E$ existiert mit $Ax=\langle y,x\rangle$ für alle $x\in E$. ◊

Beispiel 6.33.

Die Räume $\mathcal{L}^2$ und ℓ^2, versehen mit den Skalarprodukten

$$\langle f,g\rangle_{\mathcal{L}^2} = \int_X f(x)\overline{g(x)}\,\mathrm{d}\mu \qquad (f,g\in\mathcal{L}^2), \qquad\qquad \langle x,y\rangle_{\ell^2} = \sum_{k=1}^{\infty} x_k\overline{y_k} \qquad (x,y\in\ell^2),$$

sind Hilberträume. ◊

Lemma 6.34.

Sei $(E,\langle\cdot,\cdot\rangle)$ ein Prähilbertraum. Dann ist seine Komplettierung $(\overline{E},\langle\cdot,\cdot\rangle)$ ein Hilbertraum.

Beweis.

E ist bzgl. $x\mapsto||x||:=\sqrt{\langle x,x\rangle}$ ein normierter $\mathbb{K}$-Vektorraum und E lässt sich via Φ in einen Banachraum einbetten: $E\hookrightarrow E''$. Sei wie üblich $\overline{E}$ der Abschluss von E in E''. Wir definieren auf $\overline{E}$ die Form $\langle x,y\rangle = \langle\lim x_n,\lim y_n\rangle = \lim\langle x_n,y_n\rangle$ mit $x_n,y_n\in E$, $x,y\in\overline{E}$. Diese ist wohldefiniert und wir erhalten so ein Skalarprodukt auf $\overline{E}$. ◊

7. Hilbertraumtheorie

7.1. Orthonormalsysteme und Hilbertbasen

Definition 7.1.

Sei $(E,\langle\cdot,\cdot\rangle)$ ein Prähilbertraum. $\{e_i\mid i\in I\}\subseteq E$ heißt ein Orthonormalsystem in E, falls $\langle e_i,e_j\rangle=\delta_{ij}$.

Bemerkung 7.2.

Seien $(E, \langle\cdot,\cdot\rangle)$ ein Prähilbertraum und $\mathfrak{E} = (e_1, ..., e_n) \subseteq E$ ein Orthonormalsystem in E. Dann gelten:

1. $x = \alpha_i e_i \Rightarrow \alpha_i = \langle x, e_i\rangle$: Sei $j \in \{1, ..., n\}$, dann ist $\langle \sum \alpha_i e_i, e_j\rangle = \sum \alpha_i \langle e_i, e_j\rangle = \alpha_j$.
2. $(x - \sum\langle x, e_i\rangle e_i) \perp e_j$ für alle $j \in \{1, ..., n\}$: $\langle x - \sum\langle x, e_i\rangle e_i, e_j\rangle = \langle x, e_j\rangle - \langle x, e_j\rangle\langle e_j, e_j\rangle = 0$.
3. $\sum |\langle x, e_i\rangle|^2 \leq ||x||^2$: Setze $y = \sum\langle x, e_i\rangle e_i$ und $z = x - y$. Wir haben gerade eben gezeigt, dass $y \perp z$, also $||x||^2 = ||y + z||^2 = ||y||^2 + ||z||^2 \geq ||y||^2 = \sum |\langle x, e_i\rangle|^2$ nach dem Satz des Pythagoras. Dies ist die sogenannte Besselsche Ungleichung. ◊

Definition 7.3.

Sei E ein Prähilbertraum. Ein Orthonormalsystem $(e_i)_{i\in I} \subseteq E$ heißt eine Hilbertbasis von E, falls $\langle e_i\rangle_{i\in I}$ dicht in E liegt. Ein Prähilbertraum mit abzählbarer Hilbertbasis heißt separabel.

Bemerkung 7.4.

1. $\overline{\langle e_i\rangle_{i\in I}} = E$ heißt: Zu jedem $x \in E$ und jedem $\epsilon > 0$ existieren ein $n \in \mathbb{N}$ und $\alpha_1, ..., \alpha_n \in \mathbb{K}$ mit $||x - \sum \alpha_j e_j|| < \epsilon$.
2. $\langle e_i\rangle_{i\in I} = \{\sum \alpha_j e_j \mid n \in \mathbb{N},\ \alpha_1, ..., \alpha_n \in \mathbb{K}\}$ muss nicht gleich E sein, d.h. nicht jede Hilbertbasis ist eine Vektorraumbasis. ◊

Satz 7.5. (Parseval-Gleichung)

Seien $(E, \langle\cdot,\cdot\rangle)$ ein Prähilbertraum mit Hilbertbasis $\mathfrak{E} = (e_n)_{n\in\mathbb{N}}$ und $x \in E$. Dann gelten:

$$x = \sum_{i=0}^{\infty} \langle x, e_i\rangle e_i, \qquad ||x||^2 = \sum_{i=0}^{\infty} |\langle x, e_i\rangle|^2.$$

$\langle x, e_i\rangle$ heißt der i-te Fourierkoeffizient und $\sum\langle x, e_i\rangle e_i$ die Fourierreihe von x bzgl. $\mathfrak{E}$.

Beweis.

Seien $(\alpha_i)_{i\in\mathbb{N}}$, $\epsilon > 0$ und $n \in \mathbb{N}$ mit

$$x = \sum_{i=1}^{\infty} \alpha_i e_i, \qquad \left\| x - \sum_{i=1}^{n} \alpha_i e_i \right\| < \epsilon$$

Wir müssen zeigen: $\alpha_i = \langle x, e_i\rangle$ für alle $i \in I$. Für alle $m \geq n$ gilt:

$$\epsilon^2 > \left\| x - \sum_{i=1}^{n} \alpha_i e_i \right\|^2 = \left\| x - \sum_{i=1}^{m} \langle x, e_i\rangle e_i + \sum_{i=1}^{m} (\langle x, e_i\rangle - \alpha_i) e_i \right\|^2,$$

wobei wir setzen $\alpha_i = 0$ für $i \geq n$. Wegen der Orthogonalität

$$x - \sum_{i=1}^{m} \langle x, e_i\rangle e_i \quad \perp \quad \sum_{i=1}^{m} (\langle x, e_i\rangle - \alpha_i) e_i$$

folgt mit dem Satz des Pythagoras, dass

$$\epsilon^2 > \left\| x - \sum_{i=1}^{m} \langle x, e_i\rangle e_i + \sum_{i=1}^{m} (\langle x, e_i\rangle - \alpha_i) e_i \right\|^2 = \left\| x - \sum_{i=1}^{m} \langle x, e_i\rangle e_i \right\|^2 + \left\| \sum_{i=1}^{m} (\langle x, e_i\rangle - \alpha_i) e_i \right\|^2,$$

dies ist das gleiche Argument wie bei der Bessel-Ungleichung, insbesondere

$$\epsilon^2 > \left\| x - \sum_{i=1}^{m} \langle x, e_i\rangle e_i \right\|^2 \qquad \Longrightarrow \qquad x = \lim_{n\to\infty} \sum_{i=1}^{n} \langle x, e_i\rangle e_i = \sum_{i=1}^{\infty} \langle x, e_i\rangle e_i.$$

Außerdem ist nach der umgekehrten Dreiecksungleichung

$$\left| ||x|| - \left\| \sum_{i=1}^{n} \langle x, e_i\rangle e_i \right\| \right| \leq \left\| ||x|| - \sum_{i=1}^{n} \langle x, e_i\rangle e_i \right\| \overset{n\to\infty}{\longrightarrow} 0 \qquad \Longrightarrow \qquad ||x||^2 = \lim_{n\to\infty} \sum_{i=1}^{n} |\langle x, e_i\rangle|^2. \qquad \square$$

Satz 7.6.

Sei $(E, \langle\cdot,\cdot\rangle)$ ein Hilbertraum mit abzählbarer Hilbertbasis $(e_i)_{i\in\mathbb{N}}$. Dann ist E isometrisch zu ℓ^2.

Beweis.

$(e^{(i)})_{i\in\mathbb{N}}$ ist eine Hilbertbasis von ℓ^2. Definiere die Abbildung

$$F : E \to \ell^2, \qquad \sum_{i=1}^{\infty} \langle x, e_i\rangle e_i \mapsto \sum_{i=1}^{\infty} \langle x, e_i\rangle e^{(i)}.$$

1. F ist wohldefiniert, d.h. $F(x) \in \ell^2$ für alle $x \in E$: Nach der Bessel-Ungleichung ist

$$\forall n \in \mathbb{N} : \sum_{i=1}^{n} |\langle x, e_i\rangle|^2 \leq ||x||^2 < \infty \qquad \Longrightarrow \qquad \sum_{i=1}^{\infty} |\langle x, e_i\rangle|^2 < \infty \qquad \Longrightarrow \qquad \sum_{i=1}^{\infty} \langle x, e_i\rangle e_i \in \ell^2.$$

2. F ist $\mathbb{K}$-linear, da $\langle\cdot, e_i\rangle$ für alle $i \in \mathbb{N}$ $\mathbb{K}$-linear ist, d.h. für $x, y \in E$ und $\alpha, \beta \in \mathbb{K}$ gilt:

$$\begin{aligned} F(\alpha x + \beta y) &= F\left(\alpha \sum_{i=1}^{\infty} \langle x, e_i\rangle e_i + \beta \sum_{i=1}^{\infty} \langle y, e_i\rangle e_i\right) = F\left(\sum_{i=1}^{\infty} \langle \alpha x + \beta y, e_i\rangle e_i\right) \\ &= \sum_{i=1}^{\infty} \langle \alpha x + \beta y, e^{(i)}\rangle e^{(i)} = \alpha \sum_{i=1}^{\infty} \langle x, e^{(i)}\rangle e^{(i)} + \beta \sum_{i=1}^{\infty} \langle y, e^{(i)}\rangle e^{(i)} \\ &= \alpha F(x) + \beta F(y). \end{aligned}$$

3. F ist injektiv, denn sei $F(x) = 0$, dann gilt für alle $j \in \mathbb{N}$, dass $0 = \langle F(x), e^{(j)}\rangle = \langle x, e_j\rangle$, d.h. $x = \sum \langle x, e_i\rangle e_i = 0$.

4. F ist surjektiv: Sei $y = \sum \alpha_i e^{(i)} \in \ell^2$ beliebig und seien

$$y_n = \sum_{i=1}^{n} \alpha_i e^{(i)}, \qquad x_n = \sum_{i=1}^{n} \alpha_i e_i \qquad \Longrightarrow \qquad F(x_n) = y_n \text{ und } \langle x_n, x_m\rangle = \langle y_n, y_m\rangle.$$

   Wegen $y = \lim y_n$ ist $(y_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in ℓ^2, d.h. es gilt $||x_n - x_m|| = ||y_n - y_m|| \to 0$ für $n, m \to \infty$. Also ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in E. Da E ein Hilbertraum ist, existiert deren Grenzwert $x = \lim x_n$ in E und x hat die Darstellung $x = \sum \alpha_i e_i$, d.h. für alle $n \geq i$ gilt

$$\alpha_i = \langle x, e_i\rangle = \langle x_n, e_i\rangle = \langle y_n, e^{(i)}\rangle = \langle y, e^{(i)}\rangle \qquad \Longrightarrow \qquad F\left(\sum_{i=1}^{\infty} \langle x, e_i\rangle e_i\right) = \sum_{i=1}^{\infty} \langle y, e^{(i)}\rangle e^{(i)}.$$

   und damit $F(x) = y$.

5. F ist isometrisch: Seien $x, y \in E$, dann gilt F ist isometrisch: Betrachte

$$\begin{aligned} \langle F(x), F(y)\rangle &= \left\langle \sum_{i=1}^{\infty} \langle x, e_i\rangle e^{(i)}, \sum_{j=1}^{\infty} \langle y, e_j\rangle e^{(j)} \right\rangle = \lim_{n,m\to\infty} \left\langle \sum_{i=1}^{n} \langle x, e_i\rangle, \sum_{j=1}^{m} \langle y, e_j\rangle e^{(j)} \right\rangle \\ &= \lim_{n,m\to\infty} \left\langle \sum_{i=1}^{n} \langle x, e_i\rangle e_i, \sum_{j=1}^{m} \langle y, e_j\rangle e_j \right\rangle = \left\langle \sum_{i=1}^{\infty} \langle x, e_i\rangle e_i, \sum_{j=1}^{\infty} \langle y, e_j\rangle e_j \right\rangle \\ &= \langle x, y\rangle. \end{aligned}$$

$\square$

Beispiel 7.7.

$\mathcal{C}^0([0,2\pi],\mathbb{R})$, versehen mit der Integralform $\langle\cdot,\cdot\rangle_{\mathcal{L}^2}$, ist ein Prähilbertraum, der nicht vollständig ist. Für jedes $g \in \mathcal{C}^0([0,2\pi],\mathbb{R})$ existieren nämlich Koeffizientenfolgen $(a_n)_{n\in\mathbb{N}_0}, (b_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ mit

$$g(x) = \frac{a_0}{\sqrt{2\pi}} + \sum_{n=1}^{\infty} a_n \frac{\cos(nx)}{\sqrt{\pi}} + b_n \frac{\sin(nx)}{\sqrt{\pi}}.$$

Wir wählen eine Hilbertraumbasis aus periodischen Basisfunktionen, den sogenannten trigonometrischen Polynomen, zu $\mathcal{C}^0([0,2\pi],\mathbb{R})$. Die Funktionen

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos(x)}{\sqrt{\pi}}, \frac{\sin(x)}{\sqrt{\pi}}, \frac{\cos(2x)}{\sqrt{\pi}}, \frac{\sin(2x)}{\sqrt{\pi}}, \frac{\cos(3x)}{\sqrt{\pi}}, \frac{\sin(3x)}{\sqrt{\pi}}, \ldots$$

bilden nämlich bzgl. $\langle\cdot,\cdot\rangle$ ein Orthonormalsystem:

$$\begin{aligned}
\left\|\frac{1}{\sqrt{2\pi}}\right\|^2 &= \int_0^{2\pi} \frac{1}{2\pi}\,\mathrm{d}x &&= \frac{1}{2\pi}\int_0^{2\pi} 1\,\mathrm{d}x &&= 1;\\
\left\|\frac{\cos(x)}{\sqrt{\pi}}\right\|^2 &= \int_0^{2\pi} \frac{\cos^2(x)}{\pi}\,\mathrm{d}x &&= \frac{1}{\pi}\int_0^{2\pi} 1-\cos^2(x)\,\mathrm{d}x &&= 1;\\
\left\langle\frac{1}{\sqrt{2\pi}}, \frac{\cos(x)}{\sqrt{\pi}}\right\rangle &= \int_0^{2\pi} \frac{\cos x}{\sqrt{2\pi}}\,\mathrm{d}x &&= \frac{1}{\sqrt{2\pi}}\sin(x)\Big|_0^{2\pi} &&= 0;\\
\left\|\frac{\sin(x)}{\sqrt{\pi}}\right\|^2 &= \int_0^{2\pi} \frac{\sin^2(x)}{\pi}\,\mathrm{d}x &&= \frac{1}{\pi}\int_0^{2\pi} 1-\sin^2(x)\,\mathrm{d}x &&= 1;\\
\left\langle\frac{1}{\sqrt{2\pi}}, \frac{\sin(x)}{\sqrt{\pi}}\right\rangle &= \int_0^{2\pi} \frac{\sin x}{\sqrt{2\pi}}\,\mathrm{d}x &&= \frac{1}{\sqrt{2\pi}}(-\cos(x))\Big|_0^{2\pi} &&= 0;\\
\left\langle\frac{\cos(x)}{\sqrt{\pi}}, \frac{\sin(x)}{\sqrt{\pi}}\right\rangle &= \int_0^{2\pi} \frac{\sin(x)\cos(x)}{\pi}\,\mathrm{d}x &&= -\frac{\cos^2(x)}{\pi}\Big|_0^{2\pi} - \int_0^{2\pi} \frac{\cos(x)\sin(x)}{\pi}\,\mathrm{d}x &&= 0;\\
&\vdots &&\vdots &&\vdots
\end{aligned}$$

Die Fourierkoeffizienten a_n, b_n sind gegeben durch

$$a_0 = \frac{1}{\sqrt{2\pi}}\int_0^{2\pi} g(x)\,\mathrm{d}x; \qquad a_n = \frac{1}{\sqrt{\pi}}\int_0^{2\pi} g(x)\cos(nx)\,\mathrm{d}x; \qquad b_n = \frac{1}{\sqrt{\pi}}\int_0^{2\pi} g(x)\sin(nx)\,\mathrm{d}x. \qquad \lozenge$$

7.2. Orthogonalräume und orthogonale Summen

Satz 7.8. (vom Minimalabstand)

Seien E ein Hilbertraum, F ein abgeschlossener Untervektorraum von E und $x_0 \in E\backslash F$.

1. Dann existiert ein $x^\star \in F$ mit $\mathrm{dist}(x_0, F) = ||x_0 - x^\star||$.
2. Es gilt $(x_0 - x^\star)\perp F$, d.h. $(x_0 - x^\star)\perp x$ für alle $x \in F$.

Beweis.

1. Sei $(x_n)_{n\in\mathbb{N}}$ eine Minimalfolge in F, so dass mit $\delta = \lim||x_0 - x_n||$ gilt: $||x_0 - x_n|| \le \delta + \frac{1}{n}$ für alle $n \in \mathbb{N}$. Dann ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge, denn mit der Parallelogrammregel gilt:

$$||x_n - x_m||^2 = 2\underbrace{||x_n - x_0||^2}_{\xrightarrow{n\to\infty}\delta^2} + 2\underbrace{||x_m - x_0||^2}_{\xrightarrow{m\to\infty}\delta^2} - 4\underbrace{||\tfrac{1}{2}(x_n + x_m) - x_0||^2}_{\le\delta^2,\ \text{da } \frac{1}{2}(x_n - x_m)\in F} \xrightarrow{n,m\to\infty} 0$$

Setze $x^\star = \lim x_n$ in E, dann ist $x^\star \in F$, da F abgeschlossen ist. Damit gilt:

$$||x_0 - x^\star|| = ||x_0 - \lim x_n|| = ||\lim(x_0 - x_n)|| = \lim ||x_0 - x_n|| = \delta.$$

2. Wähle $x^\star \in F$ mit $||x_0 - x^\star|| \leq ||x_0 - x||$ für alle $x \in F$ und $z = x_0 - x^\star$; insbesondere $||z|| \leq ||z + \alpha x||$ für alle $x \in F$ und $\alpha \in \mathbb{K}$. Es gilt

$$||z||^2 = \langle z, z\rangle \leq \langle z, z\rangle + \alpha\langle x, z\rangle + \overline{\alpha}\langle z, x\rangle + \alpha\overline{\alpha}\langle x, x\rangle,$$

d.h. $0 \leq \alpha\langle x, z\rangle + \overline{\alpha}\langle z, x\rangle + \alpha\overline{\alpha}\langle x, x\rangle$. Setze $\alpha = \beta\langle z, x\rangle$ mit $\beta \in \mathbb{R}$. Angenommen, $\langle z, y\rangle \neq 0$. Dann

$$0 \leq \beta|\langle x, z\rangle|^2 + \beta|\langle z, x\rangle|^2 + \beta^2|\langle z, x\rangle|^2\langle x, x\rangle \qquad \Longrightarrow \qquad 0 \leq 2\beta + \beta^2\langle x, x\rangle,$$

d.h. $-2\beta \leq \beta^2||x||^2$. Wähle $0 < \beta < \frac{2}{||x||^2}$, dann $2 \leq (-\beta)||x||^2$, also $\frac{2}{||x||^2} \leq -\beta$, ein Widerspruch. Also ist $z \perp x$ für alle $x \in F$, d.h. $x_0 - x^\star \in F^\perp$. □

Korollar 7.9.

Jeder Hilbertraum besitzt eine Hilbertbasis.

Beweis.

Wähle mit Zorns Lemma ein maximales Orthonormalsystem $(e_i)_{i\in I}$ in E. Dann liegt $\langle e_i\rangle_{i\in I}$ dicht in seinem Abschluss $F = \overline{\langle e_i\rangle_{i\in I}}$ und F ist ein abgeschlossener Untervektorraum von E. Wäre nun $F \neq E$, dann gäbe es ein Œ normiertes $z \in E\backslash\{0\}$ mit $z \in F^\perp$, d.h. $z \perp e_i$ für alle $i \in I$. Dies ist ein Widerspruch zur Maximalität von $(e_i)_{i\in I}$. □

Bemerkung 7.10.

1. Sei F ein Unterraum von E, dann ist $F^\perp$ ein Unterraum von E mit $F \subseteq (F^\perp)^\perp$.
2. $F^\perp$ ist abgeschlossen, denn da $\langle\cdot,\cdot\rangle$ stetig ist, folgt: $\langle x, x_n\rangle = 0$ für alle $n \in \mathbb{N} \Rightarrow \langle x, \lim x_n\rangle = 0$. ◊

Satz 7.11. (Orthogonalraumzerlegung)

Seien E ein Hilbertraum und F ein Untervektorraum von E. Dann gelten:

1. Ist F abgeschlossen, dann lässt sich E zerlegen in $E = F \oplus F^\perp$.
2. Genau dann ist F abgeschlossen, wenn $(F^\perp)^\perp = F$.

Beweis.

1. a) Die Summe ist direkt, denn sei $x \in F \cap F^\perp$, dann $x \perp x$, d.h. $\langle x, x\rangle = 0$ und damit $x = 0$.

   b) Da $F, F^\perp$ abgeschlossen, ist auch $F + F^\perp$ abgeschlossen: Sei $(z_n) = (x_n + y_n)$ eine Cauchyfolge in $F + F^\perp$, dann gilt mit dem Satz des Pythagoras:

   $$||(y_n + z_n) - (y_m + z_m)|| = ||y_n - y_m||^2 + ||z_n - z_m||^2 \xrightarrow{n,m\to\infty} 0.$$

   Also sind $(x_n), (y_n)$ Cauchyfolgen in $F, F^\perp$, d.h. $\lim x_n \in F$ und $\lim y_n \in F^\perp$ existieren und damit auch $\lim x_n + y_n = \lim x_n + \lim y_n \in F + F^\perp$.

   c) Angenommen, $F + F^\perp \subsetneq E$. Dann gäbe es ein $z \in E\backslash\{0\}$ mit $z \perp (F + F^\perp)$, d.h. $z \perp F$ und $z \perp F^\perp$, also $z \in F^\perp \cap F = \{0\}$, ein Widerspruch.

2. a) Sei $F = (F^\perp)^\perp$, dann ist F abgeschlossen, da Orthogonalräume abgeschlossen sind.

   b) Sei F abgeschlossen. Offensichtlich ist $F \subseteq (F^\perp)^\perp$. Sei jetzt also $z \in (F^\perp)^\perp$. Wir zeigen: $z \in F$. Da F abgeschlossen, ist $E = F \oplus F^\perp$, also gibt es $x \in F, y \in F^\perp$ mit $z = x + y$. Dann gilt $0 = \langle z, y\rangle = \langle x, y\rangle + \langle y, y\rangle = ||y||^2$, d.h. $y = 0$ und damit $z = x \in F$. □

7.3. Adjungierte Operatoren und Satz von Riesz

Satz 7.12. (Riesz)

Seien E ein Hilbertraum und $x \in E$. Dann definiert $\varkappa : E \to E'$ mit $\varkappa x = \langle \cdot, x\rangle$ einen antilinearer, isometrischer Isomorphismus.

Beweis.

1. $\varkappa x : E \to \mathbb{K}$, $(\varkappa x)y = \langle y, x\rangle$ ist $\mathbb{K}$-linear und wegen $|(\varkappa x)y| = |\langle y, x\rangle| \leq ||y||\,||x||$, d.h. $\varkappa x$ ist beschränkt mit $||\varkappa x|| \leq ||x||$, d.h. $\varkappa x \in E'$ für alle $x \in E$.
2. Außerdem ist $||x||^2 = |\langle x, x\rangle| = |(\varkappa x)x| \leq ||\varkappa x||\,||x||$, d.h. $||\varkappa x|| = ||x||$ für alle $x \in E$, d.h. $\varkappa$ ist isometrisch und speziell injektiv.
3. $\varkappa$ ist antilinear, d.h. für alle $x, y \in E$ und alle $\alpha, \beta \in \mathbb{K}$ gilt $\varkappa(\alpha x + \beta y) = \overline{\alpha}\varkappa x + \overline{\beta}\varkappa y$, denn sei $z \in E$ beliebig, dann $\varkappa(\alpha x + \beta y)(z) = \langle z, \alpha x + \beta y\rangle = \overline{\alpha}\langle z, x\rangle + \overline{\beta}\langle z, y\rangle = (\overline{\alpha}\varkappa(x) + \overline{\beta}\varkappa(y))(z)$. Insbesondere ist $\varkappa$ ein Gruppenhomomorphismus bzgl. $+$.
4. $\varkappa$ ist surjektiv: Sei $f \in E'$. Zu zeigen: Es gibt $x \in E$ mit $f = \varkappa x$. Sei $F = \mathrm{Kern}(f)$. Dann ist F ein abgeschlossener Untervektorraum von E, denn f ist stetig und $\{0\}$ ist abgeschlossen, also auch $F = f^{-1}(\{0\})$. Sei Œ $F \neq E$. Dann gibt es $z \in F^{\perp}\backslash\{0\}$. Wir zeigen: Es gibt ein $\alpha \in \mathbb{K}$ mit $f = \langle \cdot, \alpha z\rangle = \varkappa x$ mit $x = \alpha z$. Wähle $\overline{\alpha} = f(z)\langle z, z\rangle^{-1}$. Dann gilt für alle $y \in E$:

$$\langle y, x\rangle = \overline{\alpha}\langle y, z\rangle = \frac{f(z)}{\langle z, z\rangle}\left\langle \underbrace{y - \frac{f(y)}{f(z)}z}_{\in F} + \underbrace{\frac{f(y)}{f(z)}z}_{\in F^{\perp}}, z\right\rangle = \frac{f(z)}{\langle z, z\rangle}\left\langle \frac{f(y)}{f(z)}z, z\right\rangle = f(y). \qquad \square$$

Lemma 7.13.

Seien E ein Hilbertraum, $x \in E$, $A \in \mathscr{L}(E, E)$. Dann ist $\varkappa(x) \circ A$ mit $y \mapsto (\varkappa x)(Ay) = \langle Ax, y\rangle$ in E'.

Beweis.

1. $\varkappa(x) \circ A$ ist linear: Seien $\alpha, \beta \in \mathbb{K}$ und $y, z \in E$, dann gilt

$$(\varkappa(x) \circ A)(\alpha y + \beta z) = \langle A(\alpha y + \beta z, x)\rangle = \alpha\langle Ay, x\rangle + \beta\langle Az, x\rangle = \alpha(\varkappa(x) \circ A)y + \beta(\varkappa(x) \circ A)z.$$

2. $\varkappa(x) \circ A$ ist stetig: Sei $y \in E$, dann gilt

$$|(\varkappa(x) \circ A)(y)| = |\langle Ay, x\rangle| \leq ||Ay||\,||x|| \leq ||A||\,||y||\,||x|| = (||A||\,||x||)||y||,$$

d.h. $\varkappa(x) \circ A$ ist beschränkt mit $||\varkappa(x) \circ A|| \leq ||A||\,||x||$. $\square$

Bemerkung 7.14.

1. Nach dem Satz von Riesz existiert dann zu jedem $x \in E$ genau ein $x^* \in E$ mit $\varkappa(x) \circ A = \varkappa x^*$, d.h. für alle $y \in E$ ist $\langle Ax, y\rangle = \langle y, x^*\rangle$.
2. Wir definieren $A^* : E \to E$ mit $A^*x = x^*$.
3. Es gilt $A^* \in \mathscr{L}(E, E)$, wobei für alle $x \in E$ gilt

$$||A^*x||^2 = \langle AA^*x, x\rangle \leq ||A^*x||\,||A||\,||x||,$$

d.h. $||A^*|| \leq ||A||$. A^* heißt der zu A adjungierte Operator.

4. Im Fall $A = A^*$ heißt A selbstadjungiert.
5. Ist $\mathbb{K} = \mathbb{R}$, dann nennen wir A in dem Fall auch symmetrisch, im Fall $\mathbb{K} = \mathbb{C}$ Hermitesch.
6. Die Zuordnung $\mathrm{Adj} : \mathscr{L}(E, E) \to \mathscr{L}(E, E)$ mit $\mathrm{Adj}(A) = A^*$ nennen wir Adjunktion.

7. A^* ist eindeutig bestimmt durch die Gleichung $\langle Ax, y\rangle = \langle x, A^*y\rangle$ für alle $x, y \in E$:

   Gelte $\langle x, A^*y\rangle = \langle Ax, y\rangle = \langle x, A'y\rangle$ für $A', A^* \in \mathscr{L}(E,E)$. Dann ist $\langle x, (A^* - A')y\rangle = 0$ für alle $x, y \in E$, d.h. speziell für $x = (A^* - A')y$: $||(A^* - A')y||^2 = 0$ für alle $y \in E$, d.h. $(A^* - A')y = 0$ für alle $y \in E$ und damit $A^* - A' = 0$, d.h. $A^* = A'$. ◊

Satz 7.15.

Sei E ein Hilbertraum. Für Adj gelten die folgenden Rechengesetze:

1. $(\alpha A + \beta B)^* = \overline{\alpha}A^* + \overline{\beta}B^*$ für alle $\alpha, \beta \in \mathbb{K}$ und $A, B \in \mathscr{L}(E,E)$;
2. $(AB)^* = B^*A^*$ für alle $A, B \in \mathscr{L}(E,E)$;
3. $(A^*)^* = A$ für alle $A \in \mathscr{L}(E,E)$;
4. $||A^*|| = ||A||$ und $||A^*A|| = ||A||^2$ für alle $A \in \mathscr{L}(E,E)$.

Beweis.

1. Seien $x, y \in E$, $\alpha, \beta \in \mathbb{K}$ und $A, B \in \mathscr{L}(E,E)$. Dann gilt:

$$\begin{aligned}\langle(\alpha A + \beta B)x, y\rangle &= \langle \alpha Ax + \beta Bx, y\rangle &= \alpha\langle Ax, y\rangle + \beta\langle Bx, y\rangle &= \alpha\langle x, A^*y\rangle + \beta\langle x, B^*y\rangle \\ &= \langle x, \overline{\alpha}A^*x + \overline{\beta}B^*y\rangle &= \langle x, (\overline{\alpha}A^* + \overline{\beta}B^*)y\rangle &= \langle x, (\overline{\alpha}A + \overline{\beta}B)^*y\rangle.\end{aligned}$$

   Also ist $(\alpha A + \beta B)^* = \overline{\alpha}A^* + \overline{\beta}B^*$.

2. Seien $x, y \in E$ und $A, B \in \mathscr{L}(E,E)$. Dann gilt:

$$\langle x, (AB)^*y\rangle = \langle ABx, y\rangle = \langle Bx, A^*y\rangle = \langle x, B^*A^*y\rangle,$$

   also $(AB)^* = B^*A^*$.

3. Seien $x, y \in E$ und $A \in \mathscr{L}(E,E)$, dann ist $\langle y, Ax\rangle = \langle A^*y, x\rangle = \langle y, (A^*)^*x\rangle$, also $(A^*)^* = A$.

4. Sei $A \in \mathscr{L}(E,E)$. Schon gezeigt: $||A^*|| \le ||A||$, also auch $||A|| = ||(A^*)^*|| \le ||A^*||$, d.h. $||A^*|| = ||A||$.

   Weiter gelten:

$$||A^*Ax|| \le ||A^*||\,||Ax|| \le ||A^*||\,||A||\,||x|| \qquad \Longrightarrow \qquad ||A^*A|| \le ||A^*||\,||A|| = ||A||^2.$$

   Umgekehrt folgt mit der Cauchy-Schwarzschen Ungleichung:

$$||Ax||^2 = \langle Ax, Ax\rangle = \langle A^*Ax, x\rangle \le ||A^*A||\,||x||^2 \qquad \Longrightarrow \qquad ||Ax|| \le \sqrt{||A^*A||}||x||,$$

   Also $||A||^2 \le ||A^*A||$. Zusammen erhalten wir $||A||^2 = ||A^*A||$. □

7.4. Spektraltheorie in Hilberträumen

Lemma 7.16.

Seien $A \in \mathscr{L}(E,E)$ und $c > 0$ derart, dass $|\langle Ax, x\rangle| \le c||x||^2$ für alle $x \in E$. Dann gilt für alle $x, y \in E$:

$$|\langle Ax, y\rangle| + |\langle Ay, x\rangle| \le 2c||x||\,||y||.$$

Speziell für $\mathbb{K} = \mathbb{R}$ und $A = A^*$ gilt: $|\langle Ax, y\rangle| \le c||x||\,||y||$.

Beweis.

Seien $x, y \in E$. Dann gilt unter Benutzung der Parallelogrammregel:

$$\begin{aligned}2|\langle Ax, y\rangle + \langle Ay, x\rangle| &= |\langle Ax, x\rangle + \langle Ay, x\rangle + \langle Ax, y\rangle + \langle Ay, y\rangle - \langle Ax, x\rangle + \langle Ay, y\rangle + \langle Ay, x\rangle - \langle Ay, y\rangle| \\ &= |\langle Ax + Ay, x\rangle + \langle Ax + Ay, y\rangle - \langle Ax - Ay, x\rangle + \langle Ax - Ay, y\rangle|\end{aligned}$$

$$= |\langle Ax + Ay, x + y\rangle - \langle Ax - Ay, x - y\rangle| \leq c(||x + y||^2 + ||x - y||^2)$$
$$= 2c(||x||^2 + ||y||^2),$$

d.h. $|\langle Ax, y\rangle + \langle Ay, x\rangle| \leq c(||x||^2 + ||y||^2)$. Wir ersetzen x nun durch tx und y durch $\frac{1}{t}y$ für ein $t > 0$. Dann gilt

$$|\langle Ax, y\rangle| + |\langle Ay, x\rangle| = |\langle A(tx), \tfrac{1}{t}y\rangle| + |\langle A(\tfrac{1}{t}y), x\rangle| \leq c(t^2||x||^2 + \tfrac{1}{t^2}||y||^2).$$

Dies gilt speziell für $t^2 = \frac{||x||}{||y||}$ (Fall $y = 0$ trivial). Wir erhalten: $|\langle Ax, y\rangle| + |\langle Ay, x\rangle| \leq 2c||x||\,||y||$. □

Satz 7.17. (Normformel für Hermitesche Operatoren)

Sei $A \in \mathscr{L}(E, E)$ Hermitesch. Dann gilt:

$$||A|| = \inf\{c > 0 \mid \langle Ax, x\rangle \leq c||x||^2 \text{ für alle } x \in E\} = \sup\{|\langle Ax, x\rangle| \mid ||x|| = 1\}.$$

Beweis.

1. Wegen $A = A^*$ gilt nach dem letzten Lemma $|\langle Ax, y\rangle| \leq ||x||^2$, d.h. $|\langle Ax, y\rangle| \leq c||x||\,||y||$ für alle $x, y \in E$. Insbesondere ist $|\langle Ax, y\rangle| \leq \iota||x||\,||y||$ mit $\iota = \inf\{c > 0 \mid |\langle Ax, x\rangle| \leq c||x||^2 \text{ für alle } x \in E\}$. Mit $\varkappa y \in E'$, $x \mapsto \langle x, y\rangle$ folgt: $|\lambda_{Ay}(x)| = |\langle x, Ay\rangle| = |\langle Ax, y\rangle| \leq \iota||x||\,||y||$. Damit erhalten wir die Abschätzungen $||Ay|| = ||\varkappa(Ay)|| \leq \iota||y||$, d.h. $||A|| \leq \iota$, und $|\langle Ax, x\rangle| \leq ||A||\,||x||^2$, d.h. $\iota \leq ||A||$. Also ist $\iota = ||A||$.
2. Wegen $|\langle Ax, x\rangle| \leq ||A||\,||x||^2$ für alle $x \in E$ ist $\sup\{|\langle Ax, x\rangle| \mid ||x|| = 1\} \leq ||A||$. Weiter gilt:

$$|\langle Ax, x\rangle| = \left|\left\langle A\frac{x}{||x||}, \frac{x}{||x||}\right\rangle\right| ||x||^2 \leq \sup_{||z||=1}\{|\langle Az, z\rangle|\} \quad \Longrightarrow \quad ||A|| = \iota \leq \sup_{||z||=1}|\langle Az, z\rangle|. \quad \square$$

Definition 7.18.

Sei E ein $\mathbb{K}$-Vektorraum. $\lambda \in \mathbb{K}$ heißt ein Eigenwert von A, falls es ein $x \in E\backslash\{0\}$ gibt mit $Ax = \lambda x$. x heißt dann ein Eigenvektor zu c und $\mathrm{Eig}(\lambda) = \mathrm{Kern}(A - \lambda\mathrm{Id}) \subseteq E$ der Eigenraum zu λ. Die Menge $\sigma(A) = \{\lambda \in \mathbb{C} \mid A - \lambda\mathrm{Id} \text{ ist nicht bijektiv in } \mathscr{L}(E, E)\}$ heißt das Spektrum von A.

Bemerkung 7.19.

1. Eigenräume sind Hilberträume, denn A ist stetig und $\{0\}$ abgeschlossen, also $\mathrm{Eig}(\lambda) = (A-\lambda\mathrm{Id})^{-1}(\{0\})$ als Urbild einer abgeschlossenen Menge unter einer stetigen Abbildung abgeschlossen.
2. Ist $\dim_{\mathbb{K}}(E) < \infty$, dann ist $\sigma(A)$ die Menge der Eigenwerte in $\mathbb{C}$, denn $A - \lambda\mathrm{Id}$ ist genau dann nicht invertierbar, wenn $A - \lambda\mathrm{Id}$ nicht injektiv ist: Injektivität, Surjektivität und Bijektivität linearer Operatoren zwischen endlichdimensionalen Räumen sind äquivalent.
3. Ist λ ein Eigenwert von A, dann liegt λ offenbar in $\sigma(A)$. Allerdings kann ein $\lambda \in \mathbb{C}$ existieren, so dass $A - \lambda\mathrm{Id}$ zwar injektiv, aber nicht surjektiv ist.
4. Das Spektrum von A zerlegt sich disjunkt in

$$\sigma_p(A) = \{\lambda \in \mathbb{C} \mid A - \lambda\mathrm{Id} \text{ nicht injektiv}\} = \{\lambda \in \mathbb{C} \mid \lambda \text{ Eigenwert von } A\}, \quad \text{(Punktspektrum)}$$
$$\sigma_c(A) = \{\lambda \in \mathbb{C} \mid A - \lambda\mathrm{Id} \text{ injektiv, nicht surjektiv und } \overline{\mathrm{Bild}(A - \lambda\mathrm{Id})} = E\}, \quad \text{(kont. Spektrum)}$$
$$\sigma_r(A) = \{\lambda \in \mathbb{C} \mid A - \lambda\mathrm{Id} \text{ injektiv, nicht surjektiv und } \overline{\mathrm{Bild}(A - \lambda\mathrm{Id})} \subsetneq E\}. \quad \text{(Restspektrum)}$$

$\varrho(A) = \{\lambda \in \mathbb{C} \mid A - \lambda\mathrm{Id} \text{ bijektiv}\} = \mathbb{C}\backslash\sigma(A)$ heißt die Resolventenmenge von A. ◊

Beispiel 7.20.

Seien $E = \ell^2$ und $A : E \to E$ der lineare Operator $A((\alpha_j)_{j\in\mathbb{N}}) = (j^{-1}\alpha_j)_{j\in\mathbb{N}}$. Dann ist A injektiv, denn es gilt $(j^{-1}a_j)_{j\in\mathbb{N}} = 0$ gdw. $(\alpha_j)_{j\in\mathbb{N}} = (0)_{j\in\mathbb{N}}$. Aber A ist nicht surjektiv, denn $(j^{-2})_{j\in\mathbb{N}}$ liegt in ℓ^2, aber nicht im Bild von A: Das Urbild von $(j^{-2})_{j\in\mathbb{N}}$ unter E wäre $(j^{-1})_{j\in\mathbb{N}} \notin \ell^2$. Damit ist $A - 0\mathrm{Id}$ nicht bijektiv, d.h. $0 \in \sigma(A)$. Allerdings ist 0 kein Eigenwert von A. ◊

7.5. Spektralsatz kompakter, Hermitescher Operatoren

Satz 7.21.

Seien E ein Hilbertraum und $A \in \mathscr{L}(E,E)$ Hermitesch. Dann sind alle Eigenwerte von A reell und für $\lambda_1 \neq \lambda_2$ gilt $\mathrm{Eig}(\lambda_1) \perp \mathrm{Eig}(\lambda_2)$, d.h. Eingenvektoren zu verschiedenen Eigenwerten sind orthogonal.

Beweis.

Gelte $Ax = \lambda x$ für ein $\lambda \in \mathbb{C}$ und ein $x \in E\backslash\{0\}$. Dann gilt:

$$\lambda ||x||^2 = \langle \lambda x, x\rangle = \langle Ax, x\rangle = \langle x, Ax\rangle = \langle x, \lambda x\rangle = \overline{\lambda}||x||^2,$$

d.h. $\lambda = \overline{\lambda}$ und damit $\lambda \in \mathbb{R}$. Seien nun $Ax = \lambda_1 x$ und $Ay = \lambda_2 y$ mit $\lambda_1 \neq \lambda_2$. Dann gilt:

$$\lambda_1\langle x,y\rangle = \langle Ax,y\rangle = \langle x,Ay\rangle = \lambda_2\langle x,y\rangle \qquad \Longrightarrow \qquad (\lambda_1-\lambda_2)\langle x,y\rangle = 0,$$

d.h. $\langle x,y\rangle = 0$ und damit $x \perp y$. □

Definition 7.22.

$A \in \mathscr{L}(E,E)$ heißt kompakt, falls es zu jeder beschränkten Folge $(x_n)_{n\in\mathbb{N}}$ aus E eine konvergente Teilfolge von $(Ax_n)_{n\in\mathbb{N}}$ gibt.

Bemerkung 7.23.

Ist $\dim_{\mathbb{K}}(E) < \infty$, dann ist jeder Operator $A \in \mathscr{L}(E,E)$ kompakt: Sei $(x_n)_{n\in\mathbb{N}} \subseteq E$ beschränkt durch ein $c > 0$. Dann ist die Kugel $\{x \in E \mid ||x|| \leq c\}$ kompakt in E und A ist stetig, also hat $(Ax_n)_{n\in\mathbb{N}}$ einen Häufungspunkt in E. ◊

Lemma 7.24.

Seien $E \neq \{0\}$ ein Hilbertraum und $A \in \mathscr{L}(E,E)$ Hermitesch und kompakt. Dann ist $||A||$ oder $-||A||$ ein Eigenwert von A.

Beweis.

Nach der Normformel für Hermitesche Operatoren gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$ mit $||x_n|| = 1$ für alle $n \in \mathbb{N}$ und $\lim|\langle Ax_n, x_n\rangle| = ||A||$. Dann besitzt $(x_n)_{n\in\mathbb{N}}$ eine Teilfolge, Œ $(x_n)_{n\in\mathbb{N}}$ selbst, mit $\langle Ax_n, x_n\rangle \to c$, wobei $c = ||A||$ oder $c = -||A||$. $(x_n)_{n\in\mathbb{N}}$ ist beschränkt durch 1 und A ist kompakt, d.h. es gibt eine Teilfolge, Œ wieder $(x_n)_{n\in\mathbb{N}}$, so dass $(Ax_n)_{n\in\mathbb{N}}$ konvergiert, etwa $Ax_n \to y \in E$. Der Fall $c = 0$ ist trivial, da dann $A = 0$. Sei also $c > 0$, dann gilt wegen

$$\begin{aligned} 0 \leq ||Ax_n - cx_n||^2 &= ||Ax_n||^2 - 2c\langle Ax_n, x_n\rangle + c^2||x_n||^2 \\ &\leq \underbrace{||A||^2}_{=c^2}\underbrace{||x_n||^2}_{=1} - 2c\underbrace{\langle Ax_n, x_n\rangle}_{\xrightarrow{n\to\infty} c} + c^2\underbrace{||x_n||^2}_{=1} \xrightarrow{n\to\infty} 0, \end{aligned}$$

dass $||y - cx_n|| \leq ||Ax_n - cx_n|| + ||y - Ax_n|| \xrightarrow{n\to\infty} 0$, d.h. $(x_n)_{n\in\mathbb{N}}$ konvergiert in E gegen $x = \frac{1}{c}y$. Dann ist $Ax = \lim Ax_n = \lim cx_n = cx$, d.h. x ist ein Eigenvektor von A zum Eigenwert c. □

Satz 7.25. (Spektralsatz für kompakte, Hermitesche Operatoren)

Seien E ein Hilbertraum, $\dim_{\mathbb{K}} E = \infty$, und $A \in \mathscr{L}(E,E)$ ein kompakter, Hermitescher Operator.

Dann besitzt $(\mathrm{Kern}(A))^\perp$ eine (endliche oder unendliche) Hilbertbasis aus Eigenvektoren $(e_n)_{n\in\mathbb{M}}$ mit zugehörigen Eigenwerten $(\lambda_n)_{n\in\mathbb{M}}$, $\lim\lambda_n = 0$ für $\sharp\mathbb{M} = \infty$, und $\sigma(A) \subseteq \{0\} \cup \{\lambda_n \mid n < \sharp\mathbb{M}\}$.

Für alle $x \in (\mathrm{Kern}(A))^\perp$ gelten damit $x = \sum\langle x, e_n\rangle e_n$ und $Ax = \sum\lambda_n\langle x, e_n\rangle e_n$.

Insbesondere gilt $(\mathrm{Kern}(A))^\perp \cong \ell^2$, falls $\sharp\mathbb{M} = \infty$, und $(\mathrm{Kern}(A))^\perp \cong \mathbb{K}^{\sharp\mathbb{M}}$ für $\sharp\mathbb{M} < \infty$.

Beweis.

1. Mit $E_0 = \mathrm{Kern}(A)$ gilt $E = E_0 \oplus E_0^\perp$, wobei $E_0, E_0^\perp$ A-invariant sind:

   E_0 ist abgeschlossen, d.h. E lässt sich in E_0 und sein Komplement zerlegen, und für $x \in E_0^\perp, y \in E_0$ gilt $\langle Ax, y\rangle = \langle x, Ay\rangle = \langle x, 0\rangle = 0$, d.h. $Ax \perp E_0$ und damit $Ax \in E_0^\perp$, d.h. $AE_0^\perp \subseteq E_0^\perp$.

   Wir können daher Œ annehmen, dass $E_0 = \{0\}$, d.h. dass A injektiv ist und somit 0 kein Eigenwert von A ist. Außerdem nehmen wir Œ $E \neq \{0\}$ an, d.h. zumindest einer der Werte $\lambda = ||A||$ oder $\lambda = -||A||$ ist ein Eigenwert von A.

2. Zu jedem Eigenraum $\mathrm{Eig}(\lambda)$ wählen wir mit Zorns Lemma ein maximales Orthonormalsystem $\Sigma(\lambda)$ aus Eigenvektoren zu λ. Dann ist $\Sigma = \bigcup \Sigma(\lambda)$ eine Hilbertbasis von E:

   Setze $F = \langle\Sigma\rangle$, dann ist Σ eine Hilbertbasis von $\overline{F}$. Falls $E \neq \overline{F}$, dann zerlegen wir E in $E = \overline{F} \oplus F^\perp$. Dann ist $F^\perp \neq \{0\}$ ein A-invarianter Unterraum: Da $F = \langle\Sigma\rangle$ A-invariant ist, gilt wie eben für alle $x \in F^\perp$, $y \in F$: $\langle Ax, y\rangle = \langle x, Ay\rangle = 0$, d.h. $Ax \perp F$ und damit $AF^\perp \subseteq F^\perp$. $A|_{F^\perp}$ hat nun einen Eigenwert $\lambda \in \{\pm||A|_{F^\perp}||\}$ mit zugehörigem Eigenvektor $e \in \mathrm{Eig}(\lambda)$ und aus $e \in F^\perp \perp \mathrm{Eig}(\lambda)$ folgt $e \perp e$, d.h. $e = 0$, ein Widerspruch. Also ist $\overline{\langle\Sigma\rangle} = E$, d.h. Σ eine Hilbertbasis von E.

3. Zu jedem $\epsilon > 0$ gibt es nur endlich viele $e \in \Sigma$ und $\lambda \in \mathbb{R}$ mit $Ae = \lambda e$ und $|\lambda| > \epsilon$, d.h. es gibt ein $\mathbb{M} \subseteq \mathbb{N}$ mit $\Sigma = \{e_n \mid n \in \mathbb{M}\}$ und $\Lambda = (\lambda_n)_{n\in\mathbb{M}} \subseteq \mathbb{R}$ mit $Ae_n = \lambda_n e_n$ für alle $n \in \mathbb{M}$, alle Eigenräume $\mathrm{Eig}(\lambda_n)$ sind endlichdimensional und $\lim \lambda_n = 0$ für $\sharp\mathbb{M} = \infty$:

   Seien $(e_n)_{n\in\mathbb{N}} \subseteq \Sigma$ und $(\lambda_n)_{n\in\mathbb{N}} \subseteq \Lambda$ mit $Ae_n = \lambda_n e_n$ und $|\lambda_n| > \epsilon$ für alle $n \in \mathbb{N}$. Da A kompakt ist, besitzt $(\lambda_n e_n)_{n\in\mathbb{N}} = (Ae_n)_{n\in\mathbb{N}}$ einen Häufungspunkt $x \in E$. Wähle $m, n \in \mathbb{N}$ mit $||\lambda_m e_m - x|| < \frac{\epsilon}{2}$ und $||\lambda_n e_n - x|| < \frac{\epsilon}{2}$, dann ist $2\epsilon^2 \leq \lambda_m^2 + \lambda_n^2 = ||\lambda_m e_m - \lambda_n e_n||^2 \leq (||\lambda_m e_m - x|| + ||\lambda_n e_n - x||)^2 = \epsilon^2$, ein Widerspruch.

4. Ist $\lambda \neq 0$ kein Eigenwert von A, dann ist $A - \lambda\mathrm{Id}$ surjektiv, d.h. alle $\lambda \in \sigma(A)$ mit $\lambda \neq 0$ sind Eigenwerte von A:

   Sei $y \in E$ beliebig. Dann besitzt y die Darstellung $y = \sum\langle y, e_n\rangle e_n$ und damit gilt

   $$x = \sum_{n\in\mathbb{M}} \frac{\langle y, e_n\rangle}{\lambda_n - \lambda} e_n \qquad \Longrightarrow \qquad (A - \lambda\mathrm{Id})x = \sum_{n\in\mathbb{M}} \frac{\langle y, e_n\rangle}{\lambda_n - \lambda}(Ae_n - \lambda e_n) = \sum_{n\in\mathbb{M}} \langle y, e_n\rangle e_n = y.$$

   Im Fall $\dim_{\mathbb{K}}(E) = \infty$ bleibt zu zeigen, dass x ein Element aus E ist. Nach der Parsevalschen Gleichung ist $\sum |\langle y, e_n\rangle|^2 = ||y||^2 < \infty$ und wegen $|\lambda_n - \lambda|^{-1} \xrightarrow{n\to\infty} |\lambda|^{-1}$ ist $|\lambda_n - \lambda|^{-1}$ beschränkt, also ist die Reihe $\sum |\frac{\langle y, e_n\rangle}{\lambda_n - \lambda}|^2 = ||x||^2$ konvergent, d.h. $x \in E$. □

Bemerkung 7.26.

Ist $A \in \mathscr{L}(E, E)$ selbstadjungiert und surjektiv, dann ist A auch injektiv, denn aus $Ax = 0$ folgt $0 = \langle x, Ay\rangle = \langle Ax, y\rangle$ für alle $y \in E$, d.h. $x \perp \mathrm{Bild}(A) = E$ und damit insbesondere $x \perp x$, d.h. $x = 0$ und folglich $\mathrm{Kern}(A) = 0$. ◊

Index

Literatur

[1] Beutelspacher, A.: *Lineare Algebra – Eine Einführung.* Vierweg+Teubner Verlag, 7th ed., 2010.

[2] Bosch, S.: *Lineare Algebra.* Springer-Verlag, 4th ed., 2008.

[3] Denk, R.: *Funktionalanalysis I.* Vorlesungsskript Universität Konstanz, p. 103, 2004.

[4] Fischer, G.: *Lineare Algebra und analytische Geometrie.* Vierweg+Teubner Verlag, 2011.

[5] Gubisch, M.: *Lineare Algebra I.* Vorlesungsskript Universität Konstanz, p. 88, 2006.

[6] Gubisch, M.: *Funktionalanalysis.* Vorlesungsskript Universität Konstanz, p. 96, 2008.

[7] Gubisch, M.: *Tutorium zur Linearen Algebra I.* Vorlesungsskript Universität Konstanz, p. 44, 2010.

[8] Kowalsky, H.-J.: *Lineare Algebra.* De Gruyter Lehrbuch, 7th ed., 1975.

[9] Scheiderer, C.: *Lineare Algebra I und II.* Vorlesungsskript Universität Duisburg, p. 248, 2001.

[10] Stammbach, U.: *Lineare Algebra.* Vorlesungsskript ETH Zürich, p. 208, 2002.

[11] Timmann, S.: *Repetitorium der Topologie und Funktionalanalysis.* Binomi Verlag, 1st ed., 2004.

[12] Werner, D.: *Funktionalanalysis.* Springer-Verlag, 5th ed., 2004.

[13] Wille, D.: *Repetitorium der Linearen Algebra 1.* Binomi Verlag, 2003.

[14] Wille, D. & Holz, M.: *Repetitorium der Linearen Algebra 2.* Binomi Verlag, 2002.