Inhaltsverzeichnis

Grundlagen

1 Grundlegende Begriffe

Wir erinnern an einige wohlvertraute Notationen der Mengenlehre. Hier einige Beispiele:

$N := \{a, b, c\}$ (Menge mit den Elementen a,b und c)

$M := \{x \in \mathbb{N} \mid x \geq 2\}$ (Die natürlichen Zahlen größer gleich 2)

$\emptyset = \{\}$ (die leere Menge)

Das kartesische Produkt zweier Mengen M und N ist die Menge aller Paare

$$M \times N = \{(x, y) \mid x \in M, y \in N\} .$$

Ist B eine endliche Menge, dann bezeichne $\#B$ die Anzahl ihrer Elemente.

Seien M und N zwei nichtleere Mengen. Eine Abbildung f von M nach N ist eine Vorschrift, welche jedem Element $m \in M$ ein eindeutig bestimmtes Bildelement $f(m) \in N$ zuordnet. Wir schreiben dies in der Form

$$f: \begin{array}{ccc} M & \to & N \\ m & \mapsto & f(m) \end{array} .$$

Zwei Abbildungen $f : M \to N$ und $g : L \to M$ lassen sich zusammensetzen. Die Komposition

$$f \circ g : L \to N$$

der Abbildungen g und f ist die Abbildung, welche $l \in L$ auf das Element $f(g(l))$ von M abbildet

$$\begin{tikzcd} L \arrow[r, "g"] \arrow[rr, bend right, "f\circ g"'] & M \arrow[r, "f"] & N \end{tikzcd}$$

Eine Abbildung $f : M \to N$ heißt injektiv, falls $f(m) = f(m')$ die Gleichheit der Elemente m und m' impliziert. Mit anderen Worten: wenn für jedes Element $n \in N$ die Menge seiner Urbilder höchstens einelementig ist.

Das Bild einer Abbildung $f : M \to N$, d.h. die Menge der Bildpunkte, ist eine Teilmenge von N. Wir bezeichnen sie mit $f(M)$. Eine Abbildung $f : M \to N$ heißt surjektiv, wenn

$$f(M) = N$$

gilt; mit anderen Worten, wenn jedes Element aus N Bildelement eines Elements aus M ist.

Ist eine Abbildung $f : M \to N$ sowohl injektiv als auch surjektiv, nennt man die Abbildung bijektiv. Eine bijektive Abbildung besitzt offensichtlich eine eindeutig bestimmte Umkehrabbildung $g : N \to M$.

2 Die Gruppenaxiome

Definition: *Eine Gruppe $(G, \circ)$ ist ein Paar bestehend aus einer Menge G und einer Abbildung*

$$\begin{array}{ccc} G \times G & \rightarrow & G \\ (g,h) & \mapsto & g \circ h \end{array}$$

mit den folgenden Eigenschaften:

(E) Es existiert ein Element $e \in G$ mit der Eigenschaft:

$$e \circ g = g$$

für alle $g \in G$.

(I) Für jedes Element $g \in G$ existiert ein Element $g^{-1} \in G$ mit der Eigenschaft

$$g^{-1} \circ g = e.$$

(A) Assoziativgesetz: Für alle $g, g', g'' \in G$ gilt

$$(g \circ g') \circ g'' = g \circ (g' \circ g'')$$

Bemerkung: Ein Element e der Gruppe mit der Eigenschaft (E) heißt neutrales (oder etwas genauer linksneutrales) Element. Elemente mit der Eigenschaft von (I) nennt man inverse (genauer linksinverse) Elemente von g.

Aus den Gruppenaxiomen folgt unmittelbar die Eindeutigkeit des neutralen Elements e sowie die Eindeutigkeit von inversen Elementen. Genauer gilt

Lemma:

(1) Linksinverse Elemente g^{-1} sind auch rechtsinvers, d.h es gilt auch

$$g \circ g^{-1} = e\ .$$

(2) Ein neutrales Element e ist auch rechtsneutral, d.h. es gilt

$$g \circ e = g$$

(3) *Inverse Element sind eindeutig: Für jedes $g \in G$ existiert genau ein Element $\tilde{g} \in G$ mit der Eigenschaft*

$$\tilde{g} \circ g = e.$$

(4) *Das neutrale Element e der Gruppe ist eindeutig bestimmt. Das heißt, es existiert genau ein Element $e \in G$ mit der Eigenschaft*

$$e \circ g = g$$

für alle $g \in G$.

(5) *Es gilt $(g^{-1})^{-1} = g$.*

Beweise:

zu (1):

$$\begin{aligned} g \circ g^{-1} &= e \circ (g \circ g^{-1}) && \text{wegen (E)} \\ &= [(g^{-1})^{-1} \circ g^{-1}] \circ (g \circ g^{-1}) && \text{wegen (I)} \\ &= (g^{-1})^{-1} \circ [g^{-1} \circ (g \circ g^{-1})] && \text{wegen (A)} \\ &= (g^{-1})^{-1} \circ [(g^{-1} \circ g) \circ g^{-1}] && \text{wegen (A)} \\ &= (g^{-1})^{-1} \circ (e \circ g^{-1}) && \text{wegen (I)} \\ &= (g^{-1})^{-1} \circ g^{-1} && \text{wegen (E)} \\ &= e. && \text{wegen (I)} \end{aligned}$$

zu (2):

$$\begin{aligned} g \circ e &= g \circ (g^{-1} \circ g) && \text{wegen (I)} \\ &= (g \circ g^{-1}) \circ g && \text{wegen (A)} \\ &= e \circ g && \text{wegen Bem.(1)} \\ &= g. && \text{wegen (E)} \end{aligned}$$

zu (3): Sei $\tilde{g}$ ein weiteres inverses Element von g. Dann gilt

$$\begin{aligned} \tilde{g} \circ g &= e && | \circ g^{-1} \\ (\tilde{g} \circ g) \circ g^{-1} &= e \circ g^{-1} && \\ \tilde{g} \circ (g \circ g^{-1}) &= g^{-1} && \text{wegen (A) und (E)} \\ \tilde{g} \circ e &= g^{-1} && \text{wegen Bem.(1)} \\ \tilde{g} &= g^{-1} && \text{wegen Bem.(2)} \end{aligned}$$

Die zwei inversen Elemente g^{-1} und $\tilde{g}$ von g sind daher gleich.

zu (4): Sei $\tilde{e}$ neben e ein weiteres neutrales Element der Gruppe. Dann gilt $\tilde{e} \circ g = g$ für alle $g \in G$. Speziell für $g = e$ erhalten wir daher

$$\tilde{e} \circ e = e.$$

Nun wenden wir Bem.(2) an für $g = \tilde{e}$, und erhalten $\tilde{e} \circ e = \tilde{e}$ sowie daraus

$$\tilde{e} = e.$$

zu (5): Nach Bem.(1) und der Definition von $(g^{-1})^{-1}$ gilt

$$\begin{aligned} g \circ g^{-1} &= e \\ (g^{-1})^{-1} \circ g^{-1} &= e \quad . \end{aligned}$$

Somit sind g und $(g^{-1})^{-1}$ inverse Elemente von g^{-1} und stimmen nach Bem.(3) überein.

Definition: Eine Gruppe $(G, \circ)$ heißt kommutativ oder abelsch, wenn zusätzlich zu den Gruppenaxiomen noch das folgende Axiom gilt:

(K) Für alle $g, g' \in G$ gilt $\quad g \circ g' = g' \circ g$.

Beispiele für kommutative Gruppen:

(a) Wählt man für G die Menge der ganzen Zahlen $\mathbb{Z}$ und für $\circ$ die Addition "$+$", dann erhält man eine kommutative Gruppe. Das neutrale Element ist 0 und das inverse Element von $g \in \mathbb{Z}$ ist $-g$.

(b) $G := \{x \in \mathbb{Q} \mid x \neq 0\}$ mit der Multiplikation " $\cdot$ " als Verknüpfung $\circ$ definiert eine kommutative Gruppe. Ihr neutrales Element ist 1 und das inverse Element $g^{-1} = \frac{1}{g}$ ist invers zu $g \in \mathbb{Q}$.

Ein Beispiel für eine nicht kommutative Gruppe: M sei eine beliebige nichtleere Menge. Es bezeichne

$$Bij(M)$$

die Menge aller bijektiven Abbildungen f von M nach M.

$$Bij(M) := \{f : M \to M \mid f \text{ ist bijektiv}\} \ .$$

Behauptung: *Bezüglich der Komposition* $\circ$ *von Abbildungen ist* $(Bij(M), \circ)$ *eine Gruppe.*

Beweis: Wir wollen zeigen, daß die Gruppenaxiome (E),(I), und (A) erfüllt sind. Beachte, daß eine Komposition von Bijektionen wieder eine Bijektion liefert.

Axiom (E) Die Existenz eines neutralen Elements: Wir wählen

$$e = id_M \ ,$$

die identische Abbildung $id_M(m) := m$ der Menge M in sich. Offensichtlich gilt dann

$$(id_M \circ f)(m) = id_M(f(m)) = f(m) \ .$$

Axiom (I) Es sei $f : M \to M$ eine gegebene Bijektion. Sei f^{-1} die Umkehrabbildung von f, welche jedem Punkt m' aus M den (eindeutig bestimmten) Urbildpunkt m zuordnet, für den gilt $f(m) = m'$. Dann gilt offensichtlich

$$(f^{-1} \circ f)(m) = f^{-1}(f(m)) = f^{-1}(m') = m \ .$$

Dies liefert also Axiom (I)

$$f^{-1} \circ f = id_M \ .$$

Axiom (A) Die Komposition von Abbildungen ist immer assoziativ:
Seien $f, g, h \in Bij(M)$, dann nämlich gilt

$$[(f \circ g) \circ h](m) = (f \circ g)(h(m)) = f(g(h(m)))$$

Andererseits ist aber

$$[f \circ (g \circ h)](m) = f[(g \circ h)(m)] = f(g(h(m))) \ .$$

Somit ist das Assoziativgesetz erfüllt.

Es wurde gezeigt: $(Bij(M), \circ)$ ist eine Gruppe.

Permutationsgruppen: Wählt man im obigen Beispiel für M speziell die endliche Menge $M = \{1, 2, 3, ..., n-1, n\}$, dann bezeichnet man die eben konstruierte Gruppe als Gruppe der Permutationen von n Zahlen. Eine synonyme Bezeichnung ist der häufig benutzte Begriff symmetrische Gruppe (von n Elementen). Man schreibt auch oft nur kurz S_n für diese Gruppe.

3 Abbildungen zwischen Gruppen

Sei $(G, \circ)$ eine Gruppe und sei H eine nichtleere Teilmenge von G.

Behauptung: *Unter der Annahme*

$$x^{-1} \circ y \in H$$

für alle $x, y \in H$ ist $(H, \circ)$ eine Gruppe.

<u>Beweis</u>: EIGENSCHAFT (E): Wähle ein $x \in H$ und setze $y = x$. Dann gilt nach der Annahme

$$e = x^{-1} \circ x \in H \ .$$

Somit enthält H das neutrale Element e.

EIGENSCHAFT (I): Wähle $x \in H$ beliebig und setze $y = e$. Nach der Annahme gilt dann

$$x^{-1} = x^{-1} \circ e \in H \ .$$

Somit liegt mit $x \in H$ auch $x^{-1} \in H$.

EIGENSCHAFT (A): Weil sich das Assoziativgesetz offensichtlich vererbt, ist H eine Gruppe. Es muß allerdings noch gezeigt werden, daß für $x, y \in H$ auch $x \circ y$ in H liegt: Aber nach (I) impliziert $x \in H$ bereits $x^{-1} \in H$. Somit gilt unter Benutzung der Annahme $(x^{-1})^{-1} \circ y \in H$. Daraus folgt wegen $x = (x^{-1})^{-1}$

$$x \circ y \in H \ .$$

H ist also abgeschlossen unter der Verknüpfung: $x, y \in H$ impliziert $x \circ y \in H$.

Definition: Eine nichtleere Teilmenge H einer Gruppe $(G, \circ)$ mit der Eigenschaft

$$x, y \in H \ \Rightarrow \ x^{-1} \circ y \in H$$

nennt man eine <u>Untergruppe</u> von $(G, \circ)$.

<u>Beispiel</u>: $H = \{1, -1\}$ ist Untergruppe von $(\mathbb{Q} \setminus \{0\}, \cdot)$.

Seien nun $(G, \circ)$ und $(H, \diamond)$ zwei gegebene Gruppen.

Definition: Eine Abbildung $\varphi : G \to H$ heißt dann Homomorphismus oder Gruppenhomomorphismus, wenn für alle $x, y \in G$ gilt

$$\varphi(x \circ y) = \varphi(x) \diamond \varphi(y) \ .$$

In Worten: Es ist egal, ob man zuerst verknüpft und dann abbildet oder zuerst abbildet und dann verknüpft.

Ein Gruppenhomomorphismus φ hat die folgenden beiden wichtigen Eigenschaften (H1) und (H2):

(H1) Ein Homomorphismus bildet das neutrale Element e_G von G auf das neutrale Element e_H von H ab

$$\boxed{\varphi(e_G) = e_H} \ .$$

Beweis: Aus der Eigenschaft $e_G \circ e_G = e_G$ des neutralen Elements folgt

$$\varphi(e_G) = \varphi(e_G \circ e_G) = \varphi(e_G) \diamond \varphi(e_G)$$

direkt durch Anwenden von φ. Dies benutzt die Definition des Homomorphismus. Durch Kürzen

$$\varphi(e_G) \diamond \varphi(e_G)^{-1} = [\varphi(e_G) \diamond \varphi(e_G)] \diamond \varphi(e_G)^{-1}.$$

und das Anwenden der uns schon bekannten Rechenregeln in einer Gruppe (Assoziativität, Linksinvers gleich Rechtsinvers, Linksneutral gleich Rechtsneutral) erhält man daraus

$$e_H = \varphi(e_G) \diamond e_H = \varphi(e_G).$$

(H2) Ein Homomorphismus bildet das Inverse eines Elements auf das Inverse des Bildelements ab

$$\boxed{\varphi(x^{-1}) = \varphi(x)^{-1}} \ .$$

Beweis: Offensichtlich gilt $\varphi(x^{-1}) \diamond \varphi(x) = \varphi(x^{-1} \circ x) = \varphi(e_G)$. Wie bereits in (H1) gezeigt gilt $\varphi(e_G) = e_H$. Daraus folgt

$$\varphi(x^{-1}) \diamond \varphi(x) = e_H = \varphi(x)^{-1} \diamond \varphi(x) \ .$$

Aus der Eindeutigkeit des inversen Elements folgt daher $\varphi(x^{-1}) = \varphi(x)^{-1}$.

Satz 1: *Das Bild $\varphi(G)$ eines Gruppenhomomorphismus $\varphi : G \to H$ ist eine Untergruppe von H.*

Beweis: Das Bild enthält wegen (H1) das neutrale Element und ist daher nicht leer. Seien $x, y \in \varphi(G)$. Dann gilt $x = \varphi(\xi), y = \varphi(\eta)$ für gewisse $\xi, \eta \in G$. Es folgt

$$\begin{aligned} x^{-1} \diamond y &= \varphi(\xi)^{-1} \diamond \varphi(\eta) && \\ &= \varphi(\xi^{-1}) \diamond \varphi(\eta) && \text{wegen (H2)} \\ &= \varphi(\xi^{-1} \circ \eta). && \text{Def. Homomorphismus} \end{aligned}$$

Also gilt $x^{-1} \diamond y \in \varphi(G)$. Somit ist $\varphi(G)$ eine Untergruppe von H.

Definition und Satz 2: *Sei $\varphi : G \to H$ ein Gruppenhomomorphismus. Dann ist die Teilmenge*

$$Kern(\varphi) \;=\; \{x \in G \mid \varphi(x) = e_H\}$$

– der sogenannte Kern der Abbildung φ – eine Untergruppe von G.

Beweis: Seien x, y aus $Kern(\varphi)$. Dann gilt

$$\varphi(x^{-1} \circ y) = \varphi(x^{-1}) \diamond \varphi(y) = \varphi(x)^{-1} \diamond \varphi(y) = e_H^{-1} \diamond e_H = e_H \;,$$

und es folgt $x^{-1} \circ y \in Kern(\varphi)$. Dies benutzt die Homomorphieeigenschaft und Eigenschaft (H2). Der Kern enthält wegen (H1) außerdem das neutrale Element, und ist somit nicht leer. Also ist $Kern(\varphi)$ eine Untergruppe von G.

Ein injektiver Homomorphismus φ besitzt notwendigerweise den trivialen Kern $Kern(\varphi) = \{e\}$. Die Umkehrung gilt auch

Satz 3: *Der Kern eines Homomorphismus φ ist trivial genau dann, wenn φ injektiv ist.*

Beweis: Sei $Kern(\varphi) = \{e_G\}$. Aus $\varphi(x) = \varphi(y)$ folgt dann wegen $\varphi(x^{-1}) = \varphi(x)^{-1} = \varphi(y)^{-1}$

$$\varphi(x^{-1} \circ y) = \varphi(x^{-1}) \diamond \varphi(y) = e_H \;.$$

Daraus folgt $x^{-1} \circ y \in Kern(\varphi) = \{e_G\}$, also $x^{-1} \circ y = e_G$. Somit folgt $x = y$ (durch Multiplikation mit x von Links unter Benutzung des Assoziativgesetzes).

4 Der Signum-Homomorphismus

Ziel dieses Abschnittes ist die Konstruktion des Signum-Homomorphismus:

$$sign: \ S_n \to \{1, -1\}$$

Vorbereitungen: Seien M und N beliebige nichtleere Mengen. Gegeben sei eine im Augenblick noch beliebige Abbildung

$$f : M^n = \underbrace{M \times \ldots \times M}_{n-mal} \ \to \ N.$$

Beachte: M^n ist die Menge aller Tupel $(m_1, m_2, \ldots, m_n)$ mit $m_i \in M$.

Operation von S_n: Für jedes $x \in S_n$, definiert man nun eine neue Abbildung

$$xf : M^n \to N$$

durch den Ansatz

$$(xf)\Big(m_1, m_2, \ldots, m_n\Big) := f\Big(m_{x(1)}, m_{x(2)}, \ldots, m_{x(n)}\Big).$$

Aus der Definition folgt unmittelbar: Sei $f : M \to N$ eine beliebige Abbildung und für $x, y \in S_n$ sei

$$z = (x \circ y) \in S_n \ .$$

Lemma: *Dann gilt*

$$\boxed{x(yf) = zf \ .}$$

Beweis: Für $m_1, \ldots, m_n \in M$ gilt einerseits

$$[(x \circ y)f]\Big(m_1, \ldots, m_n\Big) = f\Big(m_{(x \circ y)(1)}, \ldots, m_{(x \circ y)(n)}\Big) \ .$$

Andererseits ist

$$[x(yf)]\Big(m_1, \ldots, m_n\Big) = (yf)\Big(m_{x(1)}, \ldots, m_{x(n)}\Big) \ .$$

Setzt man jetzt $\tilde{m}_j := m_{x(j)}$ für $j = 1, .., n$, dann gilt

$$(yf)\Big(m_{x(1)}, \ldots, m_{x(n)}\Big) = f\Big(\tilde{m}_{y(1)}, \ldots, \tilde{m}_{y(n)}\Big) \ .$$

Wegen $\tilde{m}_{y(j)} = m_{x(y(j))} = m_{(x\circ y)(j)}$ erhält man insgesamt

$$[x(yf)]\Big(m_1, \ldots, m_n\Big) = (x \circ y)f\Big(m_1, \ldots, m_n\Big) \ .$$

Damit ist das Lemma bewiesen.

<u>Die Diskriminante</u>: Wir spezialisieren jetzt auf den Fall $M = N = \mathbb{Z}$ und wählen für f (bis auf ein Vorzeichen) die sogenannte **Diskriminanten-Funktion**

$$\Delta(m_1, \ldots, m_n) \ = \prod_{1\leq i<j\leq n} (m_j - m_i) \ .$$

Das Produkt ist anders geschrieben also

$$(m_n - m_{n-1})(m_n - m_{n-2}) \ldots (m_n - m_1)(m_{n-1} - m_{n-2}) \ldots (m_2 - m_1) \ .$$

Für $x \in S_n$ und die gewählte spezielle Funktion $f = \Delta$ (oder ihr Negatives $-\Delta$) gilt auf Grund des speziellen Verhaltens des definierenden Produkts die bemerkenswerte Identität

$$\boxed{(x\Delta)(m_1, \ldots, m_n) = sign(x) \cdot \Delta(m_1, \ldots, m_n)} \ .$$

Hierbei ist $sign(x) \in \{1, -1\}$ ein Vorzeichen, welches nur von der Permutation $x \in S_n$ aber <u>nicht</u> von $m_1, ..., m_n$ abhängt! Man sagt daher, die Diskriminantenfunktion verhalte sich **alternierend** beim Vertauschen der Variablen.

<u>Anwendung</u>: Diesen Umstand kann man nun auf das Vortrefflichste ausnutzen! Wähle für $m_1, m_2, \ldots, m_n \in \mathbb{Z}$ beliebige, allerdings voneinander <u>paarweise verschiedene</u> ganze Zahlen. Dann gilt $\Delta(m_1, .., m_n) \neq 0$ sowie

$$sign(x) \ = \ \frac{(x\Delta)(m_1, \ldots, m_n)}{\Delta(m_1, \ldots, m_n)} \quad \in \quad \{1, -1\} \ .$$

Die Pointe ist, daß der Wert $sign(x)$ auf der linken Seite dabei vollkommen <u>unabhängig</u> von der Wahl der Zahlen $m_1, \ldots, m_n$ auf der rechten Seite ist! Wir folgern daraus

Satz: *Die Abbildung*

$$sign : S_n \to \{1, -1\}$$

definiert einen Gruppenhomomorpismus.

Beweis:

$$\begin{array}{lll} [(x \circ y)\Delta](m_1, \ldots, m_n) & = sign(x \circ y)\Delta(m_1, \ldots, m_n) & \mid \text{Definition} \\ [x(y\Delta)](m_1, \ldots, m_n) & = sign(x)(y\Delta)(m_1, \ldots, m_n) & \mid \text{Definition und} \\ & & \ \ y\Delta = \pm\Delta \\ & = sign(x)sign(y)\Delta(m_1, \ldots, m_n) & \mid \text{Definition} \end{array}$$

Wegen $\Delta(m_1, \ldots, m_n) \neq 0$ und wegen des zuvor allgemeiner bewiesenen Lemmas $(x \circ y)f = x(yf)$ folgt daraus wie behauptet

$$\boxed{sign(x \circ y) = sign(x) \cdot sign(y)\ .}$$

Damit ist der Satz gezeigt.

Konvention: Wir schreiben im folgenden für das Produkt von Gruppenelementen oft nur xy statt $x \circ y$, falls keine Verwechslung möglich sind.

5 Körper

Definition: Ein Körper $(K, +, \cdot)$ ist eine Menge K, versehen mit zwei Abbildungen

$$K \times K \xrightarrow{+} K$$

$$K \times K \xrightarrow{\cdot} K$$

derart, daß gilt:

(I) $(K, +)$ ist eine kommutative Gruppe. Hierbei bezeichne 0 das neutrales Element und $-x$ das inverses Element zu x bezüglich der additiven Gruppenstruktur $+$.

(II) $(K \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe. Hierbei bezeichne 1 das neutrale Element bezüglich $\cdot$ und $\frac{1}{x}$ das inverses Element zu x bezüglich $\cdot$..

(III) Distributivgesetze: Für alle $x, y, z \in K$ soll gelten

$$(x + y) \cdot z = x \cdot z + y \cdot z \ .$$

Wegen des Kommutativgesetzes gilt dann auch $z \cdot (x + y) = z \cdot x + z \cdot y$.

Hierbei haben wir in Formel (III) die Konvention 'Punkt vor Strich' benutzt. Das heißt wir schreiben $x \cdot z + y \cdot z$ und meinen damit eigentlich $(x \cdot z) + (y \cdot z)$.

Notation: Die multiplikative Gruppe des Körpers bezeichnen wir mit $K^* := K \setminus \{0\}$.

Bemerkung 1: Ein Körper K hat mindestens zwei verschiedene Elemente, nämlich 0 und 1.

Bemerkung 2:

$$\boxed{x \cdot 0 = 0 = 0 \cdot x}$$

Insbesondere besitzt also das Nullelement 0 kein multiplikativ inverses Element!

<u>Beweis</u>: Offensichtlich ist $x \cdot 0 = x \cdot (0+0) = x \cdot 0 + x \cdot 0$. Addiert man auf beiden Seiten $-(x \cdot 0)$, so erhält man

$$\begin{aligned}(x \cdot 0) + (-(x \cdot 0)) &= (x \cdot 0) + (x \cdot 0) + (-(x \cdot 0)) \\ 0 &= (x \cdot 0) + 0 \\ 0 &= x \cdot 0\end{aligned}$$

Aus dem Kommutativgesetz folgt $0 \cdot x = 0$.

<u>Bemerkung 3</u>: Aus Bemerkung 2 folgt

$$\boxed{(x \cdot y) \cdot z = x \cdot (y \cdot z)}$$

für alle $x, y, z \in K$.

<u>Beweis</u>: Entweder gilt $x \neq 0, y \neq 0, z \neq 0$. Dann folgt die Behauptung aus (II). Wenn $x = 0$ oder $y = 0$ oder $z = 0$ gilt, dann sind wegen Bemerkung 2 beide Seiten 0, also gleich.

<u>Bemerkung 4</u>: Die kleinste natürliche Zahl $n \in N$ mit der Eigenschaft $\underbrace{1 + \ldots + 1}_{n-\text{mal}} = 0$ – falls sie denn existiert – nennt man die <u>Charakteristik</u> des Körpers. Ansonsten sagt man K hat die Charakteristik 0.

<u>Übungsaufgabe</u>: n ist eine Primzahl oder ist Null.

Wir verwenden folgende Abkürzungen:

$\frac{a}{b} := b^{-1} \cdot a$ wobei $b \neq 0$

$a - b := (-b) + a$

$z^i = \underbrace{z \cdot \ldots \cdot z}_{i-mal}$ für $i \in \mathbb{N}, z \in K$.

Mit diesen Konventionen gilt in jedem Körper (Übungsaufgabe):

$\frac{a}{b} + \frac{c}{d} = \frac{ad+cb}{bd}$ wobei $b, d \neq 0$

$a \cdot (b - c) = ab - ac$

$(-a)(-b) = ab$

$-(-a) = a$

$-b = (-1) \cdot b$

Beispiele für Körper:

(1) $(\mathbb{Q}, +, \cdot)$ Körper der rationalen Zahlen

(2) $(\mathbb{R}, +, \cdot)$ Körper der reellen Zahlen

(3) $\mathbb{F}_2 := \{0, 1\}$ der Körper mit zwei Elementen, wobei gilt

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

(4) Setze $K = \mathbb{C} = \mathbb{R} \times \mathbb{R} = \{(x, y) | x, y \in \mathbb{R}\}$. Weiter definiere

$$(x, y) + (x', y') := (x + x', y + y')$$
$$(x, y) \cdot (x', y') := (xx' - yy', xy' + yx')$$

Satz: *Die so definierte Menge $\mathbb{C}$ mit den Abbildungen $+$ und $\cdot$ bildet einen Körper, den Körper der komplexen Zahlen.*

Beweis:

(I) $0 = (0, 0)$ und $-(x, y) = (-x, -y)$

$$\begin{aligned} [(x, y) + (x', y')] + (x'', y'') &= (x + x', y + y') + (x'', y'') \\ &= ((x + x') + x'', (y + y') + y'') \\ &= (x + (x' + x''), y + (y' + y'')) \\ &= (x, y) + (x' + x'', y' + y'') \\ &= (x, y) + [(x', y') + (x'', y'')] \ . \end{aligned}$$

Das Assoziativ-Gesetz folgt somit aus dem Assoziativ-Gesetz der Körpers $\mathbb{R}$.

(II) $1 = (1,0)$ und für $(x,y) \neq (0,0)$ ist

$$(x,y)^{-1} = \left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2} \right),$$

denn

$$\left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2} \right) \cdot (x,y) = \left(\frac{x^2-(-y^2)}{x^2+y^2}, \frac{xy-yx}{x^2+y^2} \right) = (1,0)$$

Das Assoziativ-Gesetz folgt mittels einer Rechnung, analog zu (I).

(III) Die Distributiv-Gesetze folgen mittels Rechnungen analog zu (I)

<u>Beachte</u>: Beim Nachweis von (II) wurde ganz wesentlich benutzt, daß $x^2+y^2 \neq 0$ für alle $(x,y) \neq (0,0)$ gilt. Dies beruht auf der Existenz der Ordnungsstruktur des Körpers der reellen Zahlen. Alle anderen Rechnungen wären auch möglich mit einem anderen Körper K anstelle von $\mathbb{R}$.

Die herkömmliche Schreibweise der komplexen Zahlen erhält man auf folgende Weise: Setzt man $i := (0,1)$, dann gilt:

$$\begin{aligned} (x,y) &= (x+0, 0+y) \\ &= (x,0)+(0,y) \\ &= (x,0)+[(y,0)\cdot(0,1)] \\ &= (x,0)+(y,0)\cdot i \\ &= x+y\cdot i \end{aligned}$$

Hierbei wurde die Teilmenge $\mathbb{R} \times \{0\} = \{(x,0) | x \in \mathbb{R}\}$ mit dem Körper $\mathbb{R}$ identifiziert.

$$\begin{array}{ccccc} \mathbb{R} & \to & \mathbb{C} & & \\ x & \mapsto & (x,0) & & \\ x+x' & \mapsto & (x+x',0) & = & (x,0)+(x',0) \\ x\cdot x' & \mapsto & (x\cdot x',0) & = & (x,0)\cdot(x',0) \end{array}$$

Wir schreiben abkürzend einfach x statt $(x,0)$.

Das Element $i \in \mathbb{C}$ hat die bemerkenswerte Eigenschaft:

$$i \cdot i = (0,1)\cdot(0,1) = (0-1,0) = -1.$$

Insbesondere ist für jedes $x \in \mathbb{R}$ die Gleichung $z^2 = x$ in $\mathbb{C}$ lösbar. Eine mögliche Lösung ist

$$z = \begin{cases} \sqrt{x} & \text{falls } x \geq 0 \\ \\ \sqrt{|x|} \cdot i & \text{falls } x < 0 \end{cases}$$

Eine tieferliegende Analyse des Körpers der komplexen Zahlen zeigt, daß ganz allgemein jede Polynomgleichung eine Lösung in $\mathbb{C}$ besitzt.

Fundamentalsatz der Algebra: *Jedes Polynom*

$$P(X) = X^n + a_1 X^{n-1} + \ldots + a_{n-1} X + a_n$$

mit $n \geq 1$ *und Koeffizienten* $a_1, ..., a_n \in \mathbb{C}$ *besitzt eine komplexe Nullstelle, das heißt, es existiert eine Zahl* $z \in \mathbb{C}$ *mit der Eigenschaft:*

$$z^n + a_1 \cdot z^{n-1} + \ldots + a_{n-1} \cdot z + a_n = 0.$$

(Hier ohne Beweis)

Vektorräume

6 Vektorräume

Sei im folgenden K ein festgewählter Körper.

Definition: Ein K-Vektorraum V ist eine kommutative Gruppe $(V, +)$ mit einer zusätzlichen Struktur, der sogenannten Skalarmultiplikation:

$$\begin{array}{ccc} K \times V & \to & V \\ (\lambda, v) & \mapsto & \lambda \cdot v \end{array}$$

derart, daß die folgenden Verträglichkeitseigenschaften gelten:

(I) $(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$ für alle $v \in V$ und für alle $\lambda, \mu \in K$

(II) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$ für alle $v \in V$ und für alle $\lambda, \mu \in K$.

$\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$ für alle $v, w \in V$ und für alle $\lambda \in K$.

(III) $1 \cdot v = v$ für alle $v \in V$.

Bezeichnungen: Elemente aus V heißen Vektoren und Elemente aus K heißen Skalare oder Koeffizienten. Wir schreiben häufig auch nur λv anstelle von $\lambda \cdot v$.

Grundlegende Gesetze:

(1) $\left(\sum_{i=1}^{n} \lambda_i\right) \cdot v = \sum_{i=1}^{n} \lambda_i v$

(2) $\lambda \cdot \left(\sum_{i=1}^{n} v_i\right) = \sum_{i=1}^{n} \lambda v_i$

Die Eigenschaften (1) und (2) werden durch vollständige Induktion bewiesen.
Zum Beweis von (2) sei E(n) die Aussage

$$\text{E(n)}: \qquad \lambda \cdot \left(\sum_{i=1}^{n} v_i\right) = \sum_{i=1}^{n} \lambda v_i$$

Der Induktionsanfang E(1) ist trivialerweise wahr, denn $\lambda \cdot v_1 = \lambda \cdot v_1$.
Nun der Induktionsschluß: Sei E(n) wahr. Dann gilt

$$\begin{aligned}
\lambda \cdot \left(\sum_{i=1}^{n+1} v_i\right) &= \lambda \cdot \left(\sum_{i=1}^{n} v_i + v_{n+1}\right) \\
&\overset{(II)}{=} \lambda \cdot \left(\sum_{i=1}^{n} v_i\right) + \lambda v_{n+1} \\
&\overset{E(n)}{=} \sum_{i=1}^{n} \lambda v_i + \lambda v_{n+1} \\
&= \sum_{i=1}^{n+1} \lambda v_i
\end{aligned}$$

Somit ist mit "E(n) wahr" auch "E(n+1) wahr". Da E(1) wahr ist, ist somit E(n) wahr für alle $n \in \mathbb{N}$.

Der Nachweis von Gesetz(1) geht analog.

<u>1. Wichtigstes Beispiel eines Vektorraums</u>: Sei n eine positive ganze Zahl. Dann wird das n-fache kartesische Produkt $V = K^n$ mittels der Verknüpfungen

$$\begin{aligned}
(x_1, \ldots, x_n) + (y_1, \ldots, y_n) &:= (x_1 + y_1, \ldots, x_n + y_n) \\
\lambda \cdot (x_1, \ldots, x_n) &:= (\lambda \cdot x_1, \ldots, \lambda \cdot x_n)
\end{aligned}$$

zu einem K-Vektorraum.

Bemerkungen zu den Gruppengesetzen bezüglich der additiven Struktur des Vektorraums K^n:
$(0, \ldots, 0)$ ist das neutrale Element der Addition, und $(-x_1, \ldots, -x_n)$ ist das inverse Element des Vektors $(x_1, \ldots, x_n)$. Das additive Assoziativgesetz des Vektorraums ist eine unmittelbare Folge des Assoziativgesetzes des Körpers K. Nun zu den Verträglichkeitseigenschaften:

(I) $(\lambda\mu)(x_1, \ldots, x_n) = (\lambda\mu x_1, \ldots, \lambda\mu x_n) = \lambda(\mu x_1, \ldots, \mu x_n) = \lambda(\mu(x_1, \ldots, x_n))$

(II) folgt unmittelbar aus dem Distributivgesetz des Körpers K.

(III) $1 \cdot (x_1, \ldots, x_n) = (1 \cdot x_1, \ldots, 1 \cdot x_n) = (x_1, \ldots, x_n)$

Bemerkung: Wir haben bisher die Vektoren des K^n als Zeilenvektoren geschrieben. Genauso gut könnten wir die Vektoren in Form von Spaltenvektoren schreiben. Im Moment wollen wir hierbei aus Platzgründen keinen Unterschied machen. Später, wenn wir mit Matrizen rechnen, wird dies einen großen Unterschied machen und wir werden dann ab Kapitel III den Vektorraum K^n – anders als jetzt – ausschließlich als den Raum der Spaltenvektoren auffassen.

2. Beispiel für Vektorräume: Sei M eine beliebige Menge und sei K ein Körper. Sei weiterhin $V = \{f : M \to K\}$. Definiert Addition und Skalarmultiplikation durch

$$\begin{aligned} (f_1 + f_2)(m) &:= f_1(m) + f_1(m) \\ (\lambda \cdot f_1)(m) &:= \lambda \cdot f_1(m) \qquad \forall f_1, f_2 \in V \text{ und } \forall \lambda \in K \end{aligned}$$

wird V zu einem K-Vektorraum. Im Spezialfall $M = \{1, 2, \dots, n-1, n\}$ reproduziert dies das erste Beispiel. Benutze dazu die Zuordnung

$$V \ni \; f \mapsto (f(1), f(2), \dots, f(n)) \; \in K^n \; .$$

Lemma: *Für $v \in V$ und $\lambda \in K$ sind äquivalent:*

(a) $\lambda = 0$ *oder* $v = 0$

(b) $\lambda \cdot v = 0$

Beweis:
(a)⇒(b) Sei zuerst $\lambda = 0$. Dann gilt

$$0 \cdot v \;=\; (0 + 0) \cdot v \;\overset{(II)}{=}\; 0 \cdot v + 0 \cdot v.$$

Also ist $0 \cdot v = 0$ das neutrale Element, wie man durch Addition von $-(0 \cdot v)$ auf beiden Seiten sieht. Aus $\lambda = 0$ folgt also $\lambda \cdot v = 0$.

Sei nun $v = 0$. Dann gilt

$$\lambda \cdot 0 = \lambda \cdot (0 + 0) \overset{(II)}{=} \lambda \cdot 0 + \lambda \cdot 0.$$

Durch Addition von $-(\lambda \cdot 0)$ auf beiden Seiten folgt daraus wiederum $\lambda \cdot 0 = 0$.

(b)⇒(a) Sei nunmehr umgekehrt $\lambda \cdot v = 0$. Dann ist entweder $\lambda = 0$, oder es gilt $\lambda \in K^*$ mit $\lambda^{-1} \in K$. Im letzteren Fall folgt

$$\lambda^{-1} \cdot (\lambda \cdot v) = \lambda^{-1} \cdot 0 = 0$$

unter Benutzung der bereits bewiesenen Aussage (a)⇒(b). Andererseits ist nach den Axiomen einen Vektorraums die linke Seite gleich

$$(\lambda^{-1} \cdot \lambda) \cdot v = 1 \cdot v = v \ .$$

Somit ergibt sich $v = 0$.

Aus $\lambda \cdot v = 0$ folgt also wie behauptet entweder $\lambda = 0$ oder $v = 0$.

7 Erzeugendensysteme, lineare Unabhängigkeit

Sei V ein K-Vektorraum.

Definition: Sei $\mathcal{G} \subset V$ eine Menge von Vektoren. Ein Vektor $v \in V$ heißt linear abhängig von den Vektoren aus $\mathcal{G}$, falls es endlich viele Skalare $\lambda_i \in K$ und geeignete $v_i \in \mathcal{G}$ gibt, so daß gilt

$$v = \sum_{i=1}^{n} \lambda_i v_i \ .$$

Definition: $\mathcal{G}$ heißt Erzeugendensystem von V, falls jeder Vektor $v \in V$ linear abhängig von $\mathcal{G}$ ist.

Endlich viele Vektoren $v_1, \ldots, v_n$ bilden also ein Erzeugendensystem $\{v_1, \ldots, v_n\}$ von V, wenn sich jeder Vektor $v \in V$ auf mindestens eine Weise in der Gestalt

$$v = \sum_{i=1}^{n} \lambda_i v_i$$

mit $\lambda_i \in K$ schreiben läßt.

Definition: Ein Vektorraum heißt endlich dimensional, wenn es eine endliche Menge von Vektoren gibt, welche ein Erzeugendensystem bildet.

Definition: Endlich viele Vektoren $v_1, \ldots, v_n$ heißen linear unabhängig (über K), wenn für alle $\lambda_i \in K$ gilt

$$\sum_{i=1}^{n} \lambda_i v_i = 0 \quad \Rightarrow \quad \text{alle } \lambda_i \text{ sind gleich null .}$$

In Worten: Die Vektoren $v_1, ..., v_n$ sind linear unabhängig, wenn der Nullvektor sich nur auf triviale Weise aus $v_1, \ldots, v_n$ kombinieren lässt.

Definition: Eine unendliche Menge von Vektoren eines K-Vektorraumes heißt linear unabhängig, wenn jede endliche Teilmenge linear unabhängig ist.

Definition: $v_1, \dots, v_n$ heißen linear unabhängig, falls die Vektoren $v_1, \dots, v_n$ nicht linear abhängig sind.

Lemma: Vektoren $v_1, \dots, v_n$ eines K-Vektorraums V sind genau dann linear unabhängig, wenn jeder Vektor v des Vektorraumes V sich auf höchstens eine Weise als Linearkombination

$$v = \sum_{i=1}^{n} \lambda_i v_i$$

schreiben lässt.

Beweis:

$\boxed{\Rightarrow:}$ $v_1, \dots, v_n$ seien linear unabhängig. Für $v \in V$ mit

$$v = \sum_{i=1}^{n} \lambda_i v_i = \sum_{i=1}^{n} \mu_i v_i.$$

folgt dann

$$\sum_{i=1}^{n} \lambda_i v_i - \sum_{i=1}^{n} \mu_i v_i = \sum_{i=1}^{n} (\lambda_i - \mu_i) v_i = 0.$$

Wegen der linearen Unabhängigkeit der Vektoren folgt $\lambda_i - \mu_i = 0$ für alle Koeffizienten. Das heißt $\lambda_i = \mu_i$ für $i = 1, \dots, n$.

$\boxed{\Leftarrow:}$ Offensichlich gilt $\sum_{i=1}^{n} 0 \cdot v = 0$. Läßt sich also jeder Vektor $v \in V$ auf höchstens eine Weise als Linearkombination (der Elemente $v_i \in V$) schreiben, dann insbesondere $v = 0$. Für jede andere Darstellung

$$\sum_{i=1}^{n} \lambda_i v_i = 0$$

des Nullvektors folgt somit $\lambda_i = 0$ für alle λ_i wegen der Eindeutigkeit der Darstellung. Also sind $v_1, \dots, v_n$ linear unabhängig. Q.e.d.

Einige Beispiele:

1) Sei $\mathcal{G} = \{(2,3,0), (-1,-2,1)\} \subset \mathbb{R}^3$.

   a) $v := (1,1,1)$ ist linear abhängig von $\mathcal{G}$, denn $v = 1 \cdot (2,3,0) + 1 \cdot (-1,-2,1) = (1,1,1)$.

   b) $v := (3,5,-1)$ ist linear abhängig von $\mathcal{G}$, denn $v = 1 \cdot (2,3,0) + (-1) \cdot (-1,-2,1) = (3,5,-1)$.

2) $\mathcal{G} := \{(1,0,0), (0,1,0), (0,0,1)\}$ bildet ein Erzeugendensystem des $\mathbb{R}^3$.

   Aber auch $\mathcal{G} := \{(1,1,1), (0,2,1), (1,1,2)\}$ bildet ein Erzeugendensystem des $\mathbb{R}^3$.

Bemerkung: Ein einzelner Vektor v ist genau dann linear unabhängig, wenn $v \neq 0$ ist (benutze das Lemma in Abschnitt 6).

8 Der Begriff der Basis

Definition: Vektoren $v_1, \dots, v_n$ eines Vektorraums V bilden eine Basis von V, wenn sie linear unabhängig sind und $\{v_1, \dots, v_n\}$ gleichzeitig ein Erzeugendensystem von V ist.

Lemma: Eine Menge $\mathcal{G}$ von Vektoren $v_1, \dots, v_n$ eines Vektorraumes V bildet eine Basis von V genau dann, wenn sich jeder Vektor v des Vektorraumes V auf genau eine Weise als Linearkombination

$$v = \sum_{i=1}^{n} \lambda_i v_i \qquad (\lambda_i \in K) \tag{$*$}$$

schreiben läßt.

Beweis: Folgt unmittelbar aus der Definition eines Erzeugendensystems und dem Lemma in §6.

Bezeichnung: Die Zahlen $\lambda_i \in K$ in der Formel $(*)$ heißen die Koordinaten des Vektors bezüglich der Basis $v_1, \dots, v_n$.

Prinzip I: (Elimination) Seien $v_1, \dots, v_r$ Vektoren, so daß gilt: $v_1, \dots, v_{r-1}$ seien linear unabhängig, $v_1, \dots v_{r-1}, v_r$ seien nicht linear unabhängig, dann ist v_r linear abhängig von $v_1, \dots, v_{r-1}$.

Beweis: Es gibt Zahlen $\lambda_1, \dots, \lambda_r$ (nicht alle null), so daß gilt

$$\sum_{i=1}^{r} \lambda_i v_i = 0.$$

Da $v_1, \dots v_{r-1}$ linear unabhängig sind, gilt außerdem $\lambda_r \neq 0$. Daraus folgt

$$v_r = \sum_{i=1}^{r-1} (-\lambda_r^{-1} \cdot \lambda_i) \cdot v_i.$$

Also ist v_r linear abhängig von $v_1, \dots v_{r-1}$.

Prinzip II: (Transitivität) Seien $\{v_1,\ldots,v_r\}$ und $\{w_1,\ldots,w_s\}$ Mengen von Vektoren in einem Vektorraum V. Sei $v \in V$ linear abhängig von $\{v_1,\ldots,v_r\}$ und sei jedes v_i $(i = 1,\ldots,r)$ linear abhängig von $\{w_1,\ldots,w_s\}$. Dann ist v linear abhängig von $\{w_1,\ldots,w_s\}$.

Beweis: Es gilt $v = \sum\limits_{i=1}^{r} \lambda_i v_i$ und für jedes v_i gilt $v_i = \sum\limits_{j=1}^{s} \mu_{ij} w_j$ für geeignete $\lambda_i, \mu_{ij} \in K$. Einsetzen liefert

$$v = \sum_{i=1}^{r} \lambda_i \left(\sum_{j=1}^{s} \mu_{ij} w_j \right) = \sum_{i=1}^{r} \sum_{j=1}^{s} (\lambda_i \mu_{ij} w_j) = \sum_{j=1}^{s} \left(\sum_{i=1}^{r} \lambda_i \mu_{ij} \right) w_j = \sum_{i=1}^{r} \nu_j w_j .$$

Also ist v linear abhängig von $\{w_1,\ldots,w_s\}$.

Sprechweise: Zwei Mengen $\mathcal{G} = \{v_1,\ldots,v_r\}$ und $\mathcal{G}' = \{w_1,\ldots,w_s\}$ von Vektoren in V heißen linear äquivalent, falls alle v_i $(i = 1,\ldots,r)$ linear abhängig von $\{w_1,\ldots,w_s\}$ sind und alle w_j $(j = 1,\ldots,s)$ linear abhängig von $\{v_1,\ldots,v_r\}$ sind.

Prinzip III: Sei $\{w_1,\ldots,w_s\}$ linear äquivalent zu $\{v_1,\ldots,v_r\}$ und sei $\{v_1,\ldots,v_r\}$ linear äquivalent zu $\{u_1,\ldots,u_t\}$. Dann ist $\{w_1,\ldots,w_s\}$ linear äquivalent zu $\{u_1,\ldots,u_t\}$.

Beweis: Wähle $v = u_i$. Aus Prinzip II folgt: u_i ist linear abhängig von $\{w_1,\ldots,w_s\}$. Wähle umgekehrt $v = w_j$. Aus Prinzip II folgt w_j ist linear abhängig von $\{u_1,\ldots,u_t\}$.

Steinitz'scher Austauschsatz: *Gegeben seien linear unabhängige Vektoren $v_1,\ldots,v_s$ eines Vektorraumes V. Weiterhin gegeben seien Vektoren $w_1,..,w_r$ aus V so daß jeder Vektor v_i $(i = 1,\ldots,s)$ linear abhängig von $\{w_1,\ldots,w_r\}$ ist. Dann existieren s Vektoren in $\{w_1,\ldots,w_r\}$, so daß nach Ersetzen dieser Vektoren durch die Vektoren $v_1,\ldots,v_s$ die neue entstandene Menge von Vektoren linear äquivalent zur ursprünglichen Menge $\{w_1,\ldots,w_r\}$ ist. Insbesondere gilt $r \geq s$.*

Insbesondere gilt: Ist $w_1,..,w_r$ ein Erzeugendensystem von V, dann ist auch die durch Austausch entstandene Menge von Vektoren ein Erzeugendensystem von V. (Diese Eigenschaft der linearen Äquivalenz folgt aus Prinzip II).

<u>Beweis</u>: (mittels vollständiger Induktion und Prinzip II und III).

Der Induktionsanfang $s = 0$ ist trivial. Mittels Induktion können wir weiterhin annehmen, daß für $s-1$ Vektoren $v_1, \ldots, v_{s-1}$ die Aussage bereits richtig ist.

Da eine Teilmenge linear unabhängiger Vektoren linear unabhängig ist, existieren $s-1$ Vektoren in $\{w_1, \ldots, w_r\}$, so daß

$$T := \{\cancel{w}_1, \ldots, \cancel{w}_i, \ldots, \cancel{w}_r\} \cup \{v_1, \ldots, v_{s-1}\}$$

linear äquivalent zu $\{w_1, \ldots, w_r\}$ ist.

v_s ist nach Annahme linear abhängig von der Menge $\{w_1, \ldots, w_r\}$, welche ihrerseits linear äquivalent zu T ist. Aus Prinzip II folgt dann

$$v_s = \sum_{w_i \in T} \mu_i w_i + \sum_{i=1}^{s-1} \lambda_i v_i \qquad (*)$$

Mindestens einer der Koeffizienten μ_{i_0} ist nicht null, da sonst eine nichttriviale Relation

$$v_s = \sum_{i=1}^{s-1} \lambda_i v_i \qquad \text{nicht alle } \lambda_\mathrm{i} = 0$$

bestünde, was ein Widerspruch zur linearen Unabhängigkeit der Vektoren $v_1, \ldots, v_s$ wäre.

Man erhält dann durch Auflösen nach w_{i_0}

$$w_{i_0} = \mu_{i_0}^{-1} v_s - \sum_{\substack{w_i \in T \\ i \neq i_0}} \mu_{i_0}^{-1} \mu_i w_i + \sum_{i=1}^{s-1} \mu_{i_0}^{-1} \lambda_i v_i \qquad (**).$$

Zum Beweis des Induktionsschlusses genügt es zu zeigen, daß $(T \setminus \{w_{i_0}\}) \cup \{v_s\}$ linear äquivalent zu T ist und damit auch zu $\{w_1, \ldots, w_r\}$ nach Prinzip III.

w_{i_0} ist, wie in $(**)$ gezeigt, linear abhängig von $(T \setminus \{w_{i_0}\}) \cup \{v_s\}$. Außerdem sind alle Vektoren in $T \neq w_{i_0}$ enthalten in $(T \setminus \{w_{i_0}\}) \cup \{v_s\}$. Also ist T linear abhängig von $(T \setminus \{w_{i_0}\}) \cup \{v_s\}$.

Umgekehrt ist v_s linear abhängig von T wegen $(*)$. Alle Vektoren $\neq v_s$ in $(T \setminus \{w_{i_0}\}) \cup \{v_s\}$ sind enthalten in T. Also ist $(T \setminus \{w_{i_0}\}) \cup \{v_s\}$ linear abhängig von T.

Basisergänzungssatz: *Eine beliebige Menge linear unabhängige Vektoren $v_1, \ldots, v_r$ eines endlich dimensionalen Vektorraumes V kann zu einer Basis von V ergänzt werden. Das heißt, V besitzt eine Basis $w_1, \ldots, w_n$ mit der Eigenschaft $\{v_1, \ldots, v_r\} \subseteq \{w_1, \ldots, w_n\}$.*

Zusatz: *Jeder endlich dimensionale Vektorraum $V \neq \{0\}$ besitzt eine Basis.*

Beweis: Da V endlich dimensional ist, gibt es per Definition ein endliches Erzeugendensystem $\{w_1, \ldots, w_n\}$ von V.

Sei B eine maximale Teilmenge linear unabhängiger Vektoren von $\{w_1, \ldots, w_n\}$. Wegen $V \neq \{0\}$ ist B nicht die leere Menge (benutzte die letzte Bemerkung von 7!).

Aus der Maximalität von B und aus dem Eliminationsprinzip I folgt, daß jeder der Vektoren w_i linear abhängig von B ist. Also sind $\{w_1, \ldots, w_n\}$ und B linear äquivalent. Insbesondere ist B auch ein Erzeugendensystem (Prinzip II). B ist daher eine Basis von V. Damit ist der Zusatz gezeigt.

Der eigentliche Ergänzungssatz folgt aus der obigen Konstruktion der Basis B, indem man die Erzeuger $w_1, ..., w_n$ geeignet wählt. Ersetze dazu s der Vektoren $w_1, \ldots, w_n$ durch $v_1, \ldots, v_s$ (Steinitzscher Austauschsatz). Diese neue Menge von Vektoren ist auch ein endliches Erzeugendensystem, welches nunmehr die Vektoren $v_1, .., v_s$ enthält. Nun wählt man die maximale Teilmenge B so, daß sie $\{v_1, \ldots, v_s\}$ enthält. B ist die gesuchte Basisergänzung!

9 Der Dimensionsbegriff

Wir wollen die Dimension $dim_K(V)$ eines K-Vektorraumes V definieren. Wir setzen dazu $dim_K(V) := \infty$, wenn V kein endlich dimensionaler Vektorraum ist. Ist $V = \{0\}$, dann setzen wir $dim_K(V) := 0$. Ist $V \neq \{0\}$ endlich dimensional, dann besitzt V eine endliche Basis B wie in §7 gezeigt. Wir setzen

$$dim_K(V) := \#B.$$

Dies ist wohldefiniert wegen

Lemma: *Die Größe $dim_K(V)$ ist wohldefiniert, hängt also nur von dem K-Vektorraum V ab und heißt die Dimension von V.*

Beweis: Sei obdA $V \neq \{0\}$ ein endlich dimensionaler Vektorraum. Es genügt zu zeigen, daß dann je zwei Basen B und B' von V gleich viele Elemente haben. Aus Symmetriegründen genügt dazu bereits $\#B \leq \#B'$.

Nun sind die Vektoren der Basis $B = \{v_1, \ldots, v_s\}$ linear unabhängig. Die Vektoren der Basis $B' = \{w_1, \ldots, w_r\}$ bilden ein Erzeugendensystem. Aus dem Steinitzschen Austauschsatz folgt daher wie gewünscht

$$s = \#B \leq \#B' = r.$$

Lemma:

$$\boxed{dim_K(K^n) = n}$$

Beweis: Wir betrachten die Standardvektoren

$$e_i := (0, 0, \ldots, 0, 1, 0, \ldots, 0, 0)$$

(mit 1 an der $i-$ten Stelle und sonst nur Nullen) für $i = 1, \ldots, n$. Jeder Vektor $v = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ aus K^n lässt sich auf eindeutige Weise als Linearkombination der Vektoren $e_1, \ldots e_n$ schreiben

$$v = (\lambda_1, \ldots, \lambda_n) = \sum_{i=1}^{n} \lambda_i(0, \ldots, 1, \ldots, 0) = \sum_{i=1}^{n} \lambda_i e_i.$$

Also bilden die Vektoren $e_1, \ldots, e_n$ ein linear unabhängiges Erzeugendensystem von K^n, das heißt eine Basis. Es folgt $dim_K(K^n) = n$.

Die obige Basis nennen wir **Standardbasis** des K-Vektorrraums K^n.

Achtung: $V := \mathbb{C}$ hat als $\mathbb{C}-$Vektorraum aufgefaßt die Dimension $dim_{\mathbb{C}}(\mathbb{C}) = 1$, aber als $\mathbb{R}-$Vektorraum aufgefaßt die Dimension

$$dim_{\mathbb{R}}(\mathbb{C}) = 2.$$

10 K-lineare Abbildungen

Definition: Seien V, W K-Vektorräume. Eine Abbildung $\varphi : V \to W$ heißt K-linear oder auch (K-Vektorraum)-Homomorphismus, wenn für alle $v, w \in V$ und für alle $\lambda \in K$ gilt

a) $\varphi(v + w) = \varphi(v) + \varphi(w)$

b) $\varphi(\lambda v) = \lambda(\varphi(v))$.

Bemerkung: Äquivalent dazu ist die Bedingung

$$\varphi(\lambda v + \mu w) = \lambda\varphi(v) + \mu\varphi(w)$$

für alle $v, w \in V$ und für alle $\lambda, \mu \in K$.

Eine K-lineare bijektive Abbildung $\varphi : V \to W$ heißt Isomorphismus. Gibt es solch einen Isomorphismus, dann heißen V und W isomorph.

Ist $V = W$, dann nennt man K-lineare Abbildungen auch Endomorphismen und K-lineare Isomorphismen auch Automorphismen.

Lemma 1: *Die Komposition $\varphi \circ \psi$ zweier K-linearer Abbildungen $\varphi : V \to W$ und $\psi : U \to V$ ist eine K-lineare Abbildung.*

Beweis:

$$\begin{aligned}
(\varphi \circ \psi)(\lambda v + \mu w) &= \varphi(\psi(\lambda v + \mu w)) \\
&= \varphi(\lambda\psi(v) + \mu\psi(w)) \\
&= \lambda\varphi(\psi(v)) + \mu\varphi(\psi(w)) \\
&= \lambda(\varphi \circ \psi)(v) + \mu(\varphi \circ \psi)(w)
\end{aligned}$$

Lemma 2: (ohne Beweis) *Die Umkehrabbildung $\varphi^{-1} : W \to V$ eines Isomorphismus $\varphi : V \to W$ ist wieder ein Isomorphismus.*

Folgerung: *Die Automorphismen $\mathcal{G}\ell(V)$ eines Vektorraumes V bilden eine Gruppe bezüglich der Komposition der Abbildungen, eine Untergruppe der Gruppe $Bij(V)$.*

Beweis: Seien $\varphi : V \to V$ und $\psi : V \to V$ Isomorphismen. Dann ist $\varphi^{-1} : V \to V$ nach Lemma 2 wieder ein Isomorphismus. Also ist $\varphi^{-1} \circ \psi : V \to V$ ein Isomorphismus wegen Lemma 1. Die identische Abbildung $id_V : V \to V$ ist K-linear, und definiert das neutrale Element von $\mathcal{G}\ell(V)$.

Bezeichnung: Die Menge der K-linearen Abbildungen $\varphi : V \to W$ wird mit $Hom(V,W)$ oder $Hom_K(V,W)$ bezeichnet.

Behauptung: $Hom_K(V,W)$ trägt die Struktur eines K-Vektorraumes. Für $\varphi, \psi \in Hom_K(V,W)$ und $\lambda \in K$ setzen wir dazu

$$\begin{array}{rcl} (\varphi+\psi)(v) & := & \varphi(v)+\psi(v) \\ (\lambda\varphi)(v) & := & \lambda\varphi(v) \\ \\ \text{in } Hom_K(V,W) & & \text{in } W \end{array}$$

Wie man leicht nachprüfen kann erfüllen die Addition $+$ und die skalare Multiplikation auf $Hom_K(V,W)$ die Axiome eine K-Vektorraumes.

Beispielsweise muß man zeigen, daß $\varphi+\psi, \lambda\varphi \in Hom_K(V,W)$, falls $\varphi, \psi \in Hom_K(V,W)$.

$$\begin{array}{rcl} (\varphi+\psi)(\lambda v+\mu w) & = & \varphi(\lambda v+\mu w)+\psi(\lambda v+\mu w) \\ & = & \lambda\varphi(v)+\mu\varphi(w)+\lambda\psi(v)+\mu\psi(w) \\ & = & \lambda\varphi(v)+\psi(v)+\mu\varphi(w)+\psi(w) \\ & = & \lambda(\varphi+\psi)(v)+\mu(\varphi+\psi)(v) \end{array}$$

Also ist $\varphi+\psi$ wieder K-linear. Analog zeigt man dies für $\lambda\varphi$.

Definition: Im Spezialfall $W = K$ definiert der Vektorraum $Hom_K(V,W) = Hom_K(V,K)$ den sogenannten Dualraum V^* von V.

Für $\varphi \in V^*$ und $v \in V$. schreiben wir oft

$$< \varphi, v > = \varphi(v) \in K \ .$$

Dies definiert eine kanonische Paarung zwischen V^* und V.

11 Existenzsätze für Homomorphismen

Bevor wir uns mit den Struktursätzen im nächsten Abschnitt beschäftigen, machen wir hier erst einige Vorbemerkungen über Basen, Erzeugendensysteme und lineare Abbildungen:

Bemerkung (I) *Sei $\mathcal{G} \subset V$ ein Erzeugendensystem von V. Ein Homomorphismus $\varphi \in Hom_K(V,W)$ ist durch seine Werte $\varphi(v)$, $v \in \mathcal{G}$ vollständig festgelegt.*

Beweis: Für $v \in V$ existieren $\lambda_i \in K$ so daß gilt $v = \sum_{endl.} \lambda_i v_i$, mit $v_i \in \mathcal{G}$. Dies legt den Funktionswert $\varphi(v)$ durch die Werte $\varphi(v_i), v_i \in \mathcal{G}$ eindeutig fest

$$\varphi(v) = \sum_{endl.} \lambda_i \varphi(v_i) \ .$$

Bemerkung (II) *Sei V endlich dimensional, $B = \{b_1, \ldots, b_n\}$ sei eine Basis von V und W sei ein weiterer K-Vektorraum (nicht notwendigerweise endlichdimensional). Dann existiert für gegebene Vektoren $w_1, \ldots, w_n \in W$ eine K-lineare Abbildung $\varphi \in Hom(V,W)$ mit der Eigenschaft*

$$\varphi(b_i) = w_i$$

für $i = 1, \ldots, n$.

Zusatz: *Wegen Bemerkung (I) ist φ dadurch eindeutig bestimmt.*

Beweis: Jeder Vektor $v \in V$ schreibt sich eindeutig in der Form

$$v = \sum_{i=1}^{n} \lambda_i b_i$$

mit den Koordinaten $\lambda_i \in K$. Wir definieren dann φ durch

$$\varphi(v) := \sum_{i=1}^{n} \lambda_i w_i \ .$$

Diese Abbildungsvorschrift ist 'wohldefiniert' - und offensichtlich gilt $\varphi(b_i) = w_i$.

Es bleibt nur zu zeigen: φ ist K-linear. Dazu beachte

$$\begin{aligned}
\varphi(\lambda v + \lambda' v') &= \varphi(\lambda \sum_{i=1}^{n} \lambda_i b_i + \lambda' \sum_{i=1}^{n} \lambda_i' b_i) = \varphi(\sum_{i=1}^{n} (\lambda\lambda_i + \lambda'\lambda_i') b_i) \\
&=_{Def} \sum_{i=1}^{n} (\lambda\lambda_i + \lambda'\lambda_i') w_i = \lambda \sum_{i=1}^{n} \lambda_i w_i + \lambda' \sum_{i=1}^{n} \lambda_i' w_i \\
&= \lambda\varphi(v) + \lambda'\varphi(v')
\end{aligned}$$

Bemerkung (III) *Sei* $dim_K(V) < \infty$ *und* W *ein beliebiger* K*-Vektorraum. Seien* $b_1, \ldots, b_r$ *linear unabhängige Vektoren von* V *und* $w_1, \ldots, w_r \in W$ *seien beliebig. Dann existiert mindestens eine* K*-lineare Abbildung* $\varphi \in Hom_K(V, W)$ *mit* $\varphi(v_i) = w_i$ *für* $i = 1, \ldots, r$.

Beweis: Ergänze $b_1, \ldots, b_r$ zu einer Basis $b_1, \ldots, b_n$ von V. Wähle $w_{r+1}, \ldots, w_n$ in W beliebig. Nach Bemerkung (II) existiert dann eine K-lineare Abbildung $\varphi \in Hom_K(V, W)$ mit $\varphi(b_i) = w_i$ für $i = 1, \ldots, n$.

Nach diesen Vorbemerkungen wollen wir Struktursätze für endlichdimensionale K-Vektorräume beweisen.

12 Struktursätze

Das Hauptresultat dieses Abschnitts ist der

1.Struktursatz: *Ein endlich dimensionaler K-Vektorraum V der Dimension $n = dim_K(V)$ ist isomorph zum K-Vektorraum K^n.*

<u>Beweis</u>: Der Fall $V = \{0\}$ ist trivial. Sei daher obdA $b_1, \ldots, b_n$ eine Basis von V. Wir bezeichnen mit $e_1, .., e_n$ die Standardbasis des Vektorraums K^n.

Nach Bemerkung (II) des letzten Abschnitts gibt es dann eine K-lineare Abbildung φ mit der Eigenschaft $\varphi(e_i) = b_i$ für $i = 1, .., n$. Es gilt dann nach Bemerkung (I) des letzten Abschnitts

$$\begin{array}{cccc} \varphi: & K^n & \to & V \\ & (\lambda_1, \ldots, \lambda_n) & \mapsto & \sum\limits_{i=1}^{n} \lambda_i b_i \end{array}$$

Wir wollen zeigen, daß diese Abbildung ein Isomorphismus ist.

Offensichlich ist φ surjektiv, denn zu jedem $v \in V$ existieren $\lambda_1, \ldots, \lambda_n \in K$ mit $v = \sum\limits_{i=1}^{n} \lambda_i b_i$. (die Erzeugendeneigenschaft der b_i). Jedes v ist also von der Form $\varphi((\lambda_1, \ldots, \lambda_n))$ für ein geeignetes Element $(\lambda_1, \ldots \lambda_n) \in K^n$. Dies zeigt die Surjektivität. Andererseits ist φ aber auch injektiv (benutze die lineare Unabhängigkeit der b_i unter Benutzung des Lemmas in Paragraph 7).

2.Struktursatz: *Zwei endlich dimensionale K-Vektorräume V, W sind genau dann isomorph, wenn gilt:*

$$dim_K(V) = dim_K(W) \; .$$

<u>Beweis</u>: Der Fall $V = \{0\}$ ist trivial. Sei daher $\varphi : V \to W$ ein Isomorphismus und obdA $b_1, ..., b_n$ eine Basis von V. Wir behaupten dann – und dies reicht aus zum Nachweis von $dim_K(V) = dim_K(W)$ aus – daß $\varphi(b_1), .., \varphi(b_n)$ eine Basis von W ist.

Wir benutzen dazu die Charakterisierung von Basen aus dem Lemma von Paragraph 8: Vektoren $B_1, ..., b_n$ bilden genau dann ein Basis von V, wenn jeder Vektor $v \in V$ sich auf eindeutige Weise in der Form $v = \sum_i \lambda_i b_i$ schreiben lässt. Somit schreibt sich das Bild $w = \varphi(v)$ unter dem Isomorphismus φ auf eindeutige Weise in der Form $\sum_i \lambda_i \varphi(b_i)$. Da φ ein Isomorphismus ist, kommt außerdem jedes $w \in W$ auf diese Weise vor. Also bilden wegen des selben Kriteriums $\varphi(b_1), .., \varphi(b_n)$ eine Basis von W. Ein alternativer Beweis: Kopiere den Beweis der Behauptung in Paragraph 13 (im Beweis des 1.Dimensionssatzes).

Gilt umgekehrt $dim_K(V) = dim_K(W) = n < \infty$, dann folgt aus dem 1.Struktursatz die Existenz von Isomorphismen φ, ψ :

$$\varphi : \; K^n \longrightarrow V$$

$$\psi : \; K^n \longrightarrow W \; .$$

Dann ist auch φ^{-1} ein Isomorphismus $\varphi^{-1} : V \to K^n$. Der gesuchte Isomorphismus zwischen V und W ist die zusammengesetzte Abbildung

$$\psi \circ \varphi^{-1} : \quad V \xrightarrow{\varphi^{-1}} K^n \xrightarrow{\psi} W.$$

Beachte: $\psi \circ \varphi^{-1}$ ist als Komposition von K-linearen Abbildungen wieder K-linear und als Komposition von bijektiven Abbildungen wieder bijektiv, also ein Isomorphismus.

Bezeichnung: Um anzudeuten, daß eine K-lineare Abbildung $\varphi : V \to W$ ein Isomorphismus ist, schreibt man oft:

$$\varphi : V \simeq W \text{ oder } \varphi : V \xrightarrow{\simeq} W \text{ oder } \varphi : V \cong W.$$

Spielt der spezielle Isomorphismus φ keine besondere Rolle, dann schreiben wir $V \simeq W$ oder $V \cong W$ (lies V ist isomorph zu W).

13 Dimensionssätze für Kern und Bild

Definition: Ein Untervektorraum U eines K-Vektorraumes V ist eine nichtleere Teilmenge $U \subseteq V$ mit der Eigenschaft

$$v, w \in U \quad \Rightarrow \quad \lambda v + \mu w \in U \qquad , \qquad \forall \lambda, \mu \in K.$$

Bemerkung: Für $\lambda = 1, \mu = -1 \in K$ folgt daraus insbesondere $v - w \in U$ wegen $(-1)w = -w$. Also ist ein Untervektorraum U insbesondere eine Untergruppe der additiven Gruppe $(V, +)$ des Vektorraums V. Umgekehrt ist eine Untergruppe U der additiven Gruppe $(V, +)$ ein Untervektorraum, falls gilt

$$v \in U \quad \Rightarrow \quad \lambda v \in U \quad , \quad \forall v \in U, \lambda \in K \ .$$

Versehen mit der Einschränkung der skalaren Multiplikation ist ein Untervektorraum U ein K-Vektorraum. Die Inklusionsabbildung $U \hookrightarrow V$ ist aus offensichtlichen Gründen eine injektive K-lineare Abbildung. Der Basisergänzungssatz liefert

Lemma 1: *Jeder Untervektorraum U eines endlich dimensionalen K-Vektorraums V ist endlich dimensional und es gilt*

$$dim_K(U) \leq dim_K(V) \ .$$

Dimensionsgleichheit gilt genau dann, wenn gilt $U = V$.

Beispiele für Untervektorräume erhält man durch Kern und Bild von K-linearen Abbildungen. Zur Erinnerung: Eine K-lineare Abbildung ist insbesondere ein Gruppenhomomorphismus. Der $Kern(\varphi)$ einer K-linearen Abbildung

$$\varphi : V \to W$$

ist die Untergruppe aller Vektoren $v \in V$ mit $\varphi(v) = 0$.

Lemma 2: *$Kern(\varphi)$ und $Bild(\varphi)$ einer K-linearen Abbildung $\varphi : V \to W$ definieren Untervektorräume von V respektive von W.*

Beweis: Dies ist eine Variante der Sätze 1 und 2 von Paragraph 3. Auf Grund der obigen Bemerkung genügt es daher zu zeigen, daß $Kern(\varphi)$ und $Bild(\varphi)$ unter skalarer Multiplikation abgeschlossen sind. Sei dazu $v \in Kern(\varphi)$. Dann gilt $\varphi(\lambda v) = \lambda\varphi(v) = \lambda \cdot 0 = 0$. Also folgt $\lambda v \in Kern(\varphi)$ für alle $\lambda \in K$. Ist $w \in Bild(\varphi)$. Dann gilt $w = \varphi(v)$ und wie behauptet ist $\lambda w = \lambda\varphi(v) = \varphi(\lambda v)$ auch im Bild von φ.

Als Spezialfall von Satz 3 in Paragraph 3 erhält man

Lemma 3: *Eine K-lineare Abbildung φ ist genau dann injektiv, wenn gilt $Kern(\varphi) = 0$.*

1.Dimensionssatz: *Sei V ein endlichdimensionaler K-Vektorraum und $\varphi : V \to W$ eine K-lineare Abbildung. Dann gilt*

$$\boxed{dim_K(V) = dim_K\left(\varphi(V)\right) + dim_K\left(Kern(\varphi)\right).}$$

Beweis: Im Fall $V = \{0\}$ gilt $\varphi(V) = \{0\}$ sowie $Kern(\varphi) = \{0\}$, und die Aussage ist trivial. Somit können wir annehmen, daß eine Basis $b_1 \ldots, b_n$ von V existiert. Ohne Beschränkung der Allgemeinheit ist dabei $b_1 \ldots, b_r$ mit $r \leq n$ eine Basis von $Kern(\varphi) \subseteq V$, falls $Kern(\varphi) \neq \{0\}$. Dies erhält man aus dem Basisergänzungssatz! Ist $Kern(\varphi) = V$ und damit $Bild(\varphi) = \{0\}$, ist wiederum die Behauptung trivial. Wir können daher annehmen $Bild(\varphi) \neq 0$ und behaupten dann

Behauptung: Die $n - r$ Vektoren $\varphi(b_{r+1}), \ldots, \varphi(b_n)$ bilden eine Basis des Bildraums $\varphi(V)$.

Daraus folgt der Dimensionssatz $dim_K(V) = r + (n - r) = dim_K(Kern(\varphi)) + dim_K(Bild(\varphi))$.

Beweis der Behauptung:

(1) Die Bilder $\varphi(b_{r+1}), \ldots, \varphi(b_n)$ sind ein Erzeugendensystem von $Bild(\varphi)$:
Sei $w \in Bild(\varphi)$. Dann gilt

$$\begin{aligned} w &= \varphi(v) \\ &= \varphi\left(\sum_{i=1}^{n} \lambda_i b_i\right) \\ &= \varphi\left(\sum_{i=1}^{r} \lambda_i b_i + \sum_{i=r+1}^{n} \lambda_i b_i\right) \\ &= \sum_{i=1}^{r} \lambda_i \varphi(b_i) + \sum_{i=r+1}^{n} \lambda_i \varphi(b_i) \\ &= \sum_{i=r+1}^{n} \lambda_i \varphi(b_i). \end{aligned}$$

Also ist $\varphi(b_{r+1}), \ldots, \varphi(b_n)$ ein Erzeugendensystem von $\varphi(V)$.

(2) Die Vektoren $\varphi(b_{r+1}), \ldots, \varphi(b_n)$ sind linear unabhängig:
Wäre $\sum\limits_{i=r+1}^{n} \lambda_i \varphi(b_i) = 0$ eine nichttriviale Relation, dann folgt $\varphi(\sum\limits_{i=r+1}^{n} \lambda_i b_i) = 0$ und es gilt also $\sum\limits_{i=r+1}^{n} \lambda_i b_i \in Kern(\varphi)$. Wir erhalten

$$\sum_{i=r+1}^{n} \lambda_i b_i = \sum_{i=1}^{r} \mu_i b_i.$$

Setzt man $\lambda_i := -\mu_i$ für $i = 1, \ldots, r$, dann folgt eine nichttriviale Relation

$$\sum_{i=1}^{n} \lambda_i b_i = 0.$$

Dies steht im Widerspruch zur Annahme, daß $b_1, \ldots, b_n$ eine Basis von V ist.

Damit ist der Dimensionssatz gezeigt.

Lemma 4: *Sei $\varphi : V \to W$ eine K-lineare Abbildung zwischen endlich dimensionalen K-Vektorräumen V und W. Dann sind die folgenden Aussagen äquivalent:*

(i) φ ist ein Isomorphismus.

(ii) Es gilt $Kern(\varphi) = 0$ und $Bild(\varphi) = W$.

(iii) Es gilt $Kern(\varphi) = 0$ und $dim_K(V) = dim_K(W)$.

(iv) Es gilt $Bild(\varphi) = W$ und $dim_K(V) = dim_K(W)$.

Beweis: (i) $\Rightarrow$ (ii) ist trivial: φ ist bijektiv. Also folgt $Kern(\varphi) = 0$ – denn φ ist injektiv – und $Bild(\varphi) = W$ (denn φ ist surjektiv).

(ii) $\Rightarrow$ (iii) folgt aus dem Dimensionssatz: $dim_K(V) = dim_K(Kern(\varphi)) + dim_K(Bild(\varphi)) = 0 + dim_K(W)$

(iii) $\Rightarrow$ (iv) : Aus $dim_K(V) = dim_K(Kern(\varphi)) + dim_K(Bild(\varphi))$ folgt $dim_K(W) = dim_K(Bild(\varphi))$. Wegen $Bild(\varphi) \subseteq W$ folgt daher $Bild(\varphi) = W$ unter Benutzung von Lemma 1.

(iv) $\Rightarrow$ (i) : Es genügt die Injektivität von φ beziehungsweise $Kern(\varphi) = 0$ zu zeigen.

Aus $dim_K(V) = dim_K(Kern(\varphi)) + dim_K(Bild(\varphi))$ folgt

$$dim_K(W) = dim_K(Kern(\varphi)) + dim_K(W)$$

und damit $dim_K(Kern(\varphi)) = 0$. Also ist $Kern(\varphi) = 0$.

Korollar: *Im Fall $dim_K(V) \overset{!}{=} dim_K(W) < \infty$ sind für $\varphi \in Hom_K(V, W)$ folgende Aussagen äquivalent:*

(i) φ ist ein Isomorphismus.

(ii) φ ist injektiv.

(iii) φ ist surjektiv.

14 Summen von Vektorräumen

Definition: Seien U, V Untervektorräume von W. Dann definiert

$$V + U := \{v + u \mid v \in V, u \in U\}$$

die Summe der Räume V, U. Die Summe $V + U$ ist ein Untervektorraum von W wegen

$$\begin{array}{ccccc} (v+u)+(v'+u') & = & \underset{\in V}{(v+v')} & + & \underset{\in U}{(u+u')} \\ \lambda(v+u) & = & \lambda v & + & \lambda u \end{array}$$

Definition: $V + U$ heißt direkte Summe, wenn jedes Element $w \in V + U$ sich auf eindeutige Weise in der Form

$$w = v + u \quad , \quad v \in V, u \in U$$

schreibt. In diesem Fall schreibt man anstelle von $V + U$ auch

$$V \oplus U.$$

Ist B eine Basis von V und B' eine Basis von U, dann ist die Vereinigung $B \cup B'$ offensichtlich eine Basis von $U \oplus V$. Benutze zum Beweis zum Beispiel das Lemma aus Paragraph 8. Insbesondere folgt daher

Folgerung: $dim_K(U \oplus V) = dim_K(U) + dim_K(V)$

Drei Beispiele:

(I) Im Fall $U = V = W \neq 0$ gilt $U + V = W$. Dies ist aber keine direkte Summe.

(II) Im Fall $W := \mathbb{R}^2$ gilt für die Unterräume $U = \{(\lambda, 0) | \lambda \in \mathbb{R}\}$ und $V = \{(0, \mu) | \mu \in \mathbb{R}\}$ die Zerlegung

$$\mathbb{R}^2 = W = U \oplus V$$

wegen der eindeutigen Zerlegung $(\lambda, \mu) = (\lambda, 0) + (0, \mu)$.

(III) Seien U und V zwei K-Vektorräume, dann definiert das kartesische Produkt $U \times V$ mit komponentenweiser Addition und komponentenweise Skalarmultiplikation wieder einen K-Vektorraum. Die Untervektorräume $U \times \{0\}$ beziehungsweise $\{0\} \times V$ des kartesischen Produkt können mit den Vektorräumen U resp. V identifiziert werden, und man erhält eine Zerlegung der Form

$$U \times V \ = \ (U \times \{0\}) \ \oplus \ (\{0\} \times V) \ .$$

Es folgt

$$dim_K(U \times V) = dim_K(U \times \{0\}) + dim_K(\{0\} \times V) = dim_K(U) + dim_K(V) \ .$$

Beispiel (II) ist natürlich ein Spezialfall von Beispiel (III).

2.Dimensionssatz: *Seien U und V endlich dimensionale K-Untervektorräume eines K-Vektorraums W. Dann gilt*

$$\boxed{dim_K(U+V) = dim_K(U) + dim_K(V) - dim_K(U \cap V).}$$

Bemerkung: Hierbei haben wir benutzt, daß der Durchschnitt $U \cap V$ zweier Untervektorräume wieder ein Untervektorraum ist. Daher ist $dim_K(U \cap V)$ definiert. Zum Beweis können wir obdA annehmen $W = U + V$.

Beweis: Wir betrachten das cartesische Produkt $U \times V$ der Vektorräume U und V und definieren eine K-lineare Abbildung

$$\varphi : \quad U \times V \ \longrightarrow \ W = U + V$$

wie folgt durch

$$\varphi((u,v)) = u + v \ .$$

Wegen der Definition der Summe ist diese Abbildung surjektiv. Weiterhin gilt

$$Kern(\varphi) \ = \ \{(u,v) \mid u = -v, \ u \in U, \ v \in V\}$$

oder $Kern(\varphi) \;=\; \{(u,-u) \mid u \in U \cap V\}$.

$Kern(\varphi)$ ist daher als K-Vektorraum isomorph zu $Kern(\varphi) \;\cong\; U \cap V$ via der Abbildung $(u,-u) \mapsto u$. Der zweite Dimensionssatz folgt somit aus dem ersten Dimensionssatz

$$dim_K(U) + dim_K(V) = dim_K(U \times V) \;=\; dim_K(Bild(\varphi)) + dim_K(Kern(\varphi))$$

wegen $Bild(\varphi) = U + V$ und $Kern(\varphi) \cong U \cap V$.

Lemma: *$U+V$ ist genau dann eine direkte Summe, wenn gilt $U \cap V = \{0\}$.*

Beweis: Wir zeigen die Umkehrung: $U+V$ ist genau dann keine direkte Summe, wenn gilt $U \cap V \neq \{0\}$.

- $\Leftarrow$ Sei $0 \neq w \in U \cap V$, dann ist $U+V$ nicht direkt, denn $w = w+0 = 0+w$ liefert zwei verschiedene Darstellungen von w.
- $\Rightarrow$ Sei $U+V$ nicht direkt, dann existiert ein w mit $w = v+u = v'+u'$ und $v \neq v'$ in V sowie $u \neq u'$ in U. Daraus folgt $v - v' = u' - u = \xi \neq 0$ mit $\xi \in U \cap V$. Also $U \cap V \neq \{0\}$.

Die nächsten drei Abschnitte können beim Lesen übersprungen werden!

15 *Quotientenräume

Gegeben sei ein Untervektorraum U eines K-Vektorraums V. Wir wollen daraus einen neuen K-Vektorraum V/U (den Quotienten von V nach U) konstruieren zusammen mit einer surjektiven K-linearen Abbildung

$$\pi : V \to V/U \ ,$$

deren Kern der Untervektorraum U ist.

Die Elemente von V/U werden abstrakt definiert als Teilmengen von V. Dazu definieren wir eine **Äquivalenzrelation** auf V: Wir nennen Vektoren $v_1, v_2 \in V$ parallel (bezüglich U) und schreiben $v_1 \| v_2$, wenn gilt

$$v_1 - v_2 \in U \iff v_1 \in v_2 + U \ .$$

Diese Relation zwischen den Vektoren von V ist eine Äquivalenzrelation, d.h. es gilt

a) $v_1 \| v_1$

b) $v_1 \| v_2 \Rightarrow v_2 \| v_1$

c) $v_1 \| v_2, v_2 \| v_3 \Rightarrow v_1 \| v_3$.

Die **Äquivalenzklasse** eines Elementes $v \in V$, d.h. die Menge aller zu v äquivalenten Elemente, ist die 'Parallele von U durch v

$$[v] = \{v' \in V \mid \quad v' \| v \text{ bzgl. } U\} \ ,$$

oder anders geschrieben die Menge

$$v + U \ := \ \{v' \in V \mid v' \in v + U\}$$

Behauptung: Seien $v, v' \in V$. Dann gilt entweder $[v] = [v']$ oder $[v] \cap [v'] = \emptyset$.

Beweis: Dies ist eine allgemeine Eigenschaft von Äquivalenzklassen.

Definition: Wir definieren für $U \subseteq V$ eine Vektorraumstruktur auf V/U :

$$[v] + [w] \quad := \quad [v + w] \qquad \forall v, w \in V$$

$$\lambda \cdot [v] \quad := \quad [\lambda \cdot v] \quad \forall v \in V, \lambda \in K$$

Diese Definitionen sind wohldefiniert.

Beweis: Seien $[v] = [v']$ und $[w] = [w']$, dann gilt $v - v' \in U$ und $w - w' \in U$. Zu zeigen ist $[v + w] = [v' + w']$.

$$(v + w) - (v' + w') = (v - v') + (w - w') \in U + U \subseteq U$$

Daraus folgt direkt $[v + w] = [v' + w']$, denn $v' + w' \in v + w + U$.

Analog gilt: $[v] = [v'] \Rightarrow [\lambda v] = [\lambda v']$.

$$v - v' \in U \Rightarrow \lambda(v - v') = \lambda v - \lambda v' \in \lambda U \subseteq U$$

Diese Verknüpfungen erfüllen die Axiome eines Vektorraums. Also ist V/U ein K-Vektorraum.

Behauptung: Die Abbildung

$$\begin{array}{rccc} \pi : & V & \rightarrow & V/U \\ & v & \mapsto & [v] \end{array}$$

ist eine surjektive K-lineare Abbildung.

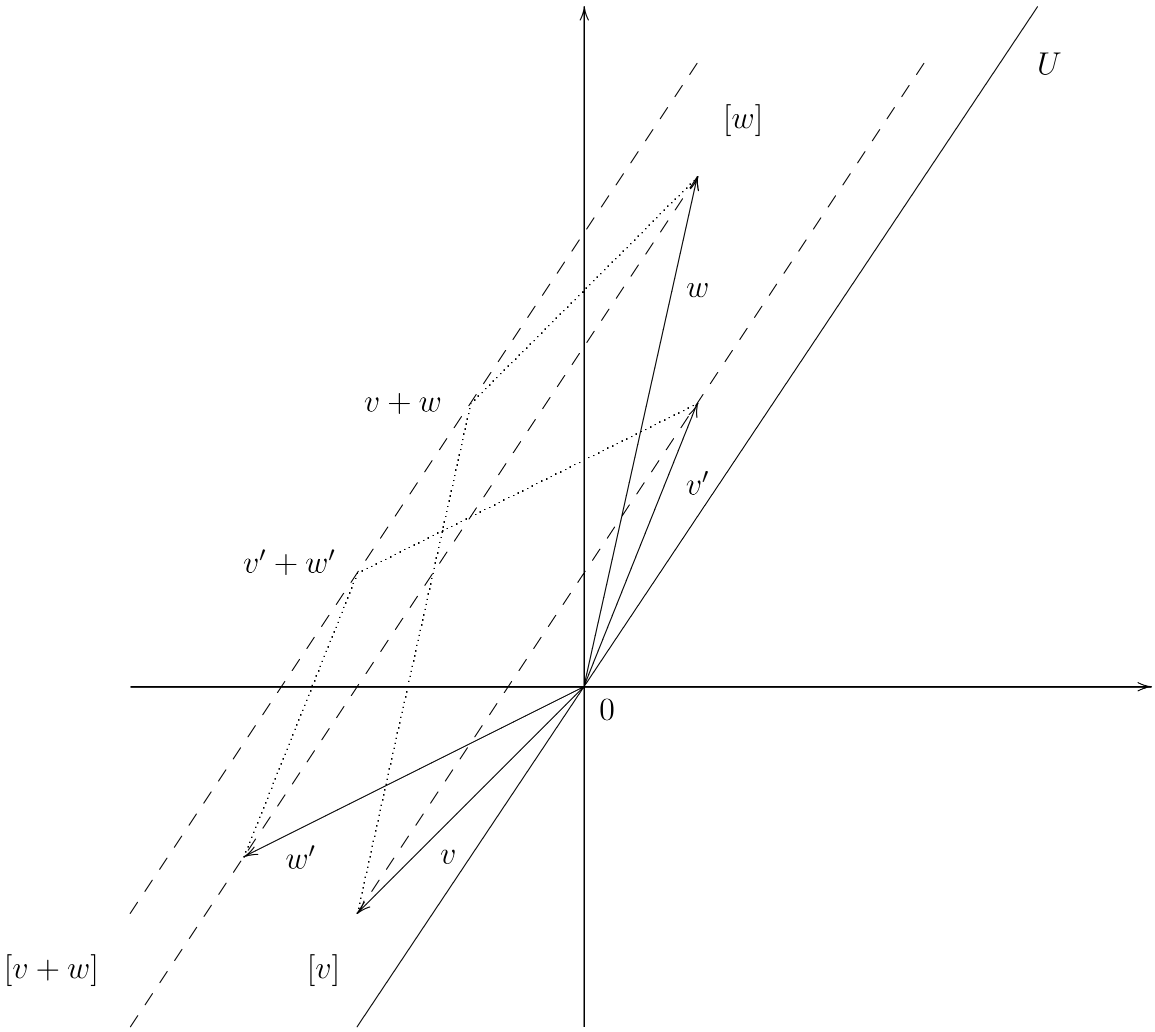
Beweis: $\pi(\lambda v + \mu w) = [\lambda v + \mu w] = [\lambda v] + [\mu w] = \lambda[v] + \mu[w] = \lambda \pi(v) + \mu \pi(w)$
Die Surjektivität ist klar.

Behauptung: $Kern(\pi) = U$.

Beweis: $v \in Kern(\pi) \iff \pi(v) = 0 \iff [v] = [0] \Leftrightarrow v \in 0 + U = U$

Ist V ein $\mathbb{R}$-Vektorraum, also $K = \mathbb{R}$, dann sind die Teilmengen $[v]$ wirklich Parallelen von U durch V. Der Quotientenraum V/U hat also die Bedeutung eines Raumes von Parallelen von U.

Diese Intuition ist natürlich etwas gewöhnungsbedürftig im Falle von Körpern K wie zum Beispiel $K = \mathbb{F}_2$, welche verschieden von $\mathbb{R}$ sind.



16 *Dualitätstheorie und Annulator

Sei $b_1, \ldots, b_n$ eine Basis von V. Im Dualraum $V^* = Hom(V, K)$ gibt es wegen Bemerkung (II) §11 eindeutig bestimmte K-lineare Abbildung

$$b_i^* \in Hom_K(V, K)$$

mit der Eigenschaft

$$b_i^*(b_j) = \delta_{ij} \in K$$

wobei

$$\delta_{ij} := \begin{cases} 0 & \text{wenn } j \neq i \\ 1 & \text{wenn } j = i \end{cases}$$

Satz 1: *Die $b_1^*, \ldots, b_n^*$ bilden eine Basis von V^*, die sogenannte* <u>*Dualbasis*</u> *von $b_1, \ldots, b_n$. Insbesondere gilt $dim_K(V^*) = dim_K(V)$.*

<u>Beweis</u>: Zuerst zeigen wir die lineare Unabhängigkeit der b_i^*. Sei

$$\sum_{i=1}^{n} \lambda_i b_i^* = 0.$$

Einsetzen von b_j liefert

$$0 = \sum_{i=1}^{n} \lambda_i b_i^*(b_j) = \lambda_j \cdot 1 = \lambda_j$$

für alle $j = 1, \ldots, n$. Also sind die b_i^* linear unabhängig.
Nun zeigen wir, daß $b_1^*, \ldots, b_n^*$ ein Erzeugendensystem bilden: Für $\varphi \in V^* = Hom_K(V, K)$ setzen wir $\lambda_i := \varphi(b_i) \in K$ für $i = 1, \ldots, n$ und betrachten

$$\psi := \sum_{i=1}^{n} \lambda_i b_i^* \ .$$

Dann gilt $\psi(b_j^*) = \lambda_j = \varphi(b_j^*)$ für alle $j = 1, \ldots, n$. Wegen §11 Bemerkung (II) folgt daraus $\psi = \varphi$. Also ist $b_1^*, \ldots, b_n^*$ ein Erzeugendensystem, und somit eine Basis von V^*.

Definition: Sei U ein Untervektorraum von V. Setze

$V^* \supseteq U^\perp := \{v^* \in V^* \mid v^*(u) = 0 \quad \forall u \in U\} = \{v^* \in V^* \mid \ < v^*, u > = 0 \quad \forall u \in U\}$.

$U^\perp$ heißt der Annulator von U.

Bemerkung: $U^\perp$ ist ein Untervektorraum von V^*.

Ein allgemeines Prinzip: $v^* \in U^\perp \Leftrightarrow v^*(b_i) = 0$ für alle $b_1, \ldots, b_r$ einer Basis von U.

Beweis: Dies ist klar, da $< v^*, u >$ linear in u ist und jedes u eine Linearkombination der Basisvektoren ist.

Satz 2:

$$\boxed{dim_K(U^\perp) = dim_K(V) - dim_K(U)}$$

Beweis: Sei $b_1, ..., b_r$ Basis von U, $b_1, ..., b_r, ..., b_n$ Basis von V und $b_1^*, ..., b_n^*$ Dualbasis von V^*.

Behauptung: $U^\perp$ hat als Basis die Elemente $b_{r+1}^*, ..., b_n^*$.

Beweis: Offensichtlich sind $b_{r+1}^*, ..., b_n^*$ linear unabhängig. Andererseits gilt

$$b_i^* \in U^\perp \text{ für } i \geq r+1.$$

Dies folgt aus dem allgemeinen Prinzip wegen

$$b_i^*(b_1) = \cdots = b_i^*(b_r) = 0 \text{ für } i \geq r.$$

Zu zeigen bleibt: $b_{r+1}^*, ..., b_n^*$ ist Erzeugendensystem (und somit Basis) von $U^\perp$.

Sei $v^* = \sum\limits_{i=1}^{n} \lambda_i b_i^* \in U^\perp \overset{allg.Prinz.}{\iff} v^*(b_i) = \lambda_i = 0 \quad \forall i = 1, ..., r \iff \lambda_1 = \cdots = \lambda_r = 0$. Es folgt

$$U^\perp \ = \ \{v^* = \sum_{i=r+1}^{n} \lambda_i b_i^* \mid \lambda_i \in K\}$$

und $b_{r+1}^*, ..., b_n^*$ ist Erzeugendensystem von $U^\perp$. Somit gilt $dim_K(V) = r + n - r = dim_K(U) + dim_K(U^\perp)$.

q.e.d.

17 *Die duale Abbildung

Definition: Sei $\varphi : V \to W$ eine K-lineare Abbildung. Dann definiert

$$\begin{array}{l} \varphi^* : W^* \to V^* \\ \varphi^* : \; w^* \; \mapsto w^* \circ \varphi \end{array}$$

eine K-lineare Abbildung von W^* nach V^*. φ^* heißt duale oder transponierte Abbildung von φ.

Zur Erläuterung betrachte das folgende Diagram

$$\begin{array}{ccc} V & \overset{\varphi}{\to} & W \\ & {\scriptstyle \varphi^*(w^*)}\searrow & \downarrow_{w^*} \\ & & K \end{array}$$

Wir wollen die K-Linearität von φ^* beweisen: Für alle $v \in V$ gilt:

$$\begin{aligned} \varphi^*(\lambda w_1^* + \mu w_2^*)(v) &= (\lambda w_1^* + \mu w_2^*)\varphi(v) \\ &= (\lambda w_1^* + \mu w_2^*)(\varphi(v)) \\ &= \lambda w_1^*(\varphi(v)) + \mu w_2^*(\varphi(v)) \\ &= \lambda(\varphi(w_1^*))(v) + \mu(\varphi(w_2^*))(v) \end{aligned}$$

Also gilt $\varphi(\lambda w_1^* + \mu w_2^*) = \lambda\varphi(w_1^*) + \mu\varphi(w_2^*)$ in V^*.

Zur Erinnerung: $\varphi^*(v) :=< \varphi^*, v > \quad \in K \qquad$ für $\varphi \in V^*, v \in V$

Lemma:

$$\boxed{< w^*, \varphi(v) > = < \varphi^*(w^*), v >}$$

Beweis: $< w^*, \varphi(v) > = w^*(\varphi(v)) = (w^* \circ \varphi)(v) = \varphi^*(w^*)(v) = < \varphi^*(w^*), v >$

Satz 1: $Kern(\varphi^*) = (Bild(\varphi))^\perp$ (in W^*)

Beweis: Sei $\varphi^* : W^* \to V^*$. Dann gilt

$$w^* \in Kern(\varphi^*) \iff \varphi^*(w^*)(v) = 0 \qquad \forall v \in V$$

$$\begin{aligned}
&\iff \quad < \varphi^*(w^*), v > = 0 \qquad \forall v \in V \\
&\iff \quad < w^*, \varphi(v) > = 0 \qquad \forall v \in V \\
&\iff \quad w^* \in (Bild(\varphi))^\perp
\end{aligned}$$

Definition: Sei $\varphi : V \to W$ eine K-lineare Abbildung zwischen endlich dimensionalen Vektorräumen. Dann nennen wir $dim_K\,(Bild(\varphi))$ den Rang $r(\varphi)$ der Abbildung φ.

Es sei für den Rest des Abschnitts V endlich dimensionaler K-Vektorraum.

Satz 2: $r(\varphi) = r(\varphi^*)$

Beweis:

$$\begin{aligned}
r(\varphi^*) &= dim_K(Bild(\varphi^*)) \\
&= dim_K(W^*) - dim_K(Kern(\varphi^*)) \\
&\overset{Satz\ 1}{=} dim_K(W^*) - dim_K\left((Bild(\varphi))^\perp\right) \\
&= dim_K(W) - dim_K\left((Bild(\varphi))^\perp\right) \\
&\overset{\S 16}{=} dim_K(W) - (dim_K(W) - dim_K(Bild(\varphi))) \\
&= dim_K(Bild(\varphi)) \\
&= r(\varphi)
\end{aligned}$$

Satz 3: $Bild(\varphi^*) = Kern(\varphi)^\perp$ (in V^*)

Beweis: Sei $\varphi^* : W^* \to V^*$. Dann gilt

$$\begin{aligned}
v^* \in Bild(\varphi^*) &\Leftrightarrow \exists w^* \in W^* \text{ mit } v^* = \varphi^*(w^*) \\
&\Leftrightarrow \exists w^* \text{ mit } v^*(v) = \varphi^*(w^*)(v) = w^*(\varphi(v)) \qquad \forall v \in V.
\end{aligned}$$

Für $v^* \in Bild(\varphi^*)$ folgt somit für alle $v \in Kern(\varphi)$ $v^*(v) = w^*(\varphi(v)) = w^*(0) = 0$. Man erhält also $Bild(\varphi^*) \subseteq Kern(\varphi)^\perp$.

Um die Gleichheit zu zeigen genügt nach §13 Lemma 1 die Gleichheit der Dimensionen. Es gilt:

$$\begin{aligned}
dim_K(Kern(\varphi)^\perp) &= dim_K(V) - dim_K(Kern(\varphi)) && \text{§16} \\
&= dim_K(V) - (dim_K(V) - dim_K(Bild(\varphi))) && \text{Dimensionssatz} \\
&= dim_K(Bild(\varphi)) \\
&= dim_K(Bild(\varphi^*)) && \text{Satz 2}
\end{aligned}$$

Wie wir bereits gesehen haben, gilt $dim_K(V^*) = dim_K(V)$. Aus dem Struktursatz in §12 folgt also

$$V^* \cong V.$$

Leider gibt es keine ausgezeichnete Wahl eines solchen Isomorphismus. Dies wird uns später zu der Untersuchung der Bilinearformen führen.

Anders ist die Situation beim Bidual $(V^*)^*$.

Satz 4: *Es gibt einen kanonischen Isomorphismus* $\kappa : V \xrightarrow{\simeq} (V^*)^*$.

Beweis: Wir betrachten die Abbildung

$$\begin{aligned} \kappa : V &\to (V^*)^* \\ v &\mapsto \Big(v^* \mapsto v^*(v)\Big) \end{aligned}$$

Zur Erläuterung sei wiederholt, daß $(V^*)^* = Hom_K(V^*, K)$ ist.
κ ist linear, denn es gilt:

$$\begin{aligned} \kappa(\mu v_1 + \lambda v_2) &= \Big(v^* \mapsto v^*(\mu v_1 + \lambda v_2)\Big) \\ &= \Big(v^* \mapsto \mu v^*(v_1) + \lambda v^*(v_2)\Big) \\ &= \mu\Big(v^* \mapsto v^*(v_1)\Big) + \lambda\Big(v^* \mapsto v^*(v_2)\Big) \\ &= \mu\kappa(v_1) + \lambda\kappa(v_2) \end{aligned}$$

κ ist auch injektiv: Sei $v \in Kern(\kappa)$, dann gilt

$$v^*(v) = 0$$

für alle $v^* \in V^*$. Man erhält $v = 0$. Denn wäre $v \neq 0$, dann ergänze $v = b_1$ zu einer Basis und setze $v^* := b_1^*$. Aus $v = 0$ folgt $Kern(\kappa) = 0$.
Aus $dim_K((V^*)^*) = dim_K(V^*) = dim_K(V)$ (nach §16) folgt dann, daß κ automatisch bijektiv ist. (Siehe §13.)

Lineare Gleichungssysteme und Matrizen

18 Spalten- und Zeilenvektoren

Wir betrachten im folgenden den Vektorraum K^n. Dieser Vektorraum war definiert als Vektorraum aller Zeilenvektoren $(\lambda_1, \ldots, \lambda_n)$ mit den Koordinaten $\lambda_i \in K$. Wir hätten diesen Vektorraum auch in Form von Spaltenvektoren der Form

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

schreiben können. Addition und Skalarmultiplikation sind dann wie folgt definiert:

$$\begin{aligned} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} &= \begin{pmatrix} \lambda_1 + \mu_1 \\ \vdots \\ \lambda_n + \mu_n \end{pmatrix} \\ \lambda \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} &= \begin{pmatrix} \lambda \cdot \lambda_1 \\ \vdots \\ \lambda \cdot \lambda_n \end{pmatrix} \end{aligned}$$

<u>Konvention</u>: Von nun an bezeichnet K^n den K–Vektorraum der Spaltenvektoren der Länge n und $(K^n)^*$ den K–Vektorraum der Zeilenvektoren der Länge n. <u>Standardbasen</u> von K^n und $(K^n)^*$ sind

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \quad e_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \cdots \quad e_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

respektive

$$e_1^* = (1, 0, \ldots, 0) \qquad e_2^* = (0, 1, 0, \ldots, 0) \qquad \cdots \qquad , e_n^* = (0, \ldots, 0, 1)$$

Für $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ und $y = (y_1, \ldots, y_n) \in (K^n)^*$ definieren wir eine <u>kanonische Paarung</u> zwischen $(K^n)^*$ und K^n (das <u>Skalarprodukt</u>)

$$< x, y > := \sum_{i=1}^{n} y_i x_i = y_1 x_1 + \cdots + y_n x_n \qquad \in K.$$

19 Matrizen und Abbildungen (I)

Wir betrachten in diesem Abschnitt K–lineare Abbildungen

$$\varphi : V \to W$$

zwischen den Vektorräumen $V = K^n$ und $W = K^m$.

Jeder solchen Abbildung ordnen wir eine $m \times n$ Matrix mit Einträgen zu in K vermöge

$$M(\varphi) := \begin{pmatrix} \vdots & \vdots & & \vdots \\ \varphi(e_1) & \varphi(e_2) & \cdots & \varphi(e_n) \\ \vdots & \vdots & & \vdots \end{pmatrix}$$

zu. Hierbei seien $e_1, \ldots, e_n$ die Standardbasisvektoren von $V = K^n$.

In obiger Schreibweise wollen wir nunmehr "unnötige Klammern" weggelassen. Wie das zu verstehen ist, soll das folgende Beispiel erläutern.

Beispiel: Sei $n = m$, also $V = W$ und sei $\varphi := id_V$. Dann ist

$$M(id_V) = \left(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \cdots \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right) \stackrel{!}{=} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Die so entstandene Matrix nennt man die <u>Einheitsmatrix</u> $E = E_n$ und es gilt

$$E = (\delta_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}} .$$

Hierbei ist $\delta_{ij} = 1$ oder 0, je nach dem ob $i = j$ oder $i \neq j$ ist. δ_{ij} ist das sogenannte <u>Kronecker-Symbol</u>.

Definition: Die Menge aller $m \times n$ Matrizen mit Einträgen in K bezeichnen wir mit $M_{m,n}(K)$. Man erklärt eine K–Vektorraumstruktur auf $M_{m,n}(K)$ via

$$\lambda \cdot (M_{ij}) = (\lambda M_{ij})$$

und

$$(M_{ij}) + (M'_{ij}) = (M_{ij} + M'_{ij})$$

für Matrizen $M = (M_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ und $M' = (M'_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ aus $M_{m,n}(K)$.

Wie wir in §11, Bemerkung (II) gesehen haben, bestimmt die Matrix $M(\varphi)$ (mit anderen Worten die Bilder $\varphi(e_i)$ der Basis $e_1, .., e_n$) die Abbildung φ eindeutig; und zu jeder Wahl von Spaltenvektoren existiert eine zugehörige K-lineare Abbildung $\varphi : K^n \to K^m$.

Satz: *Die Zuordnung $\varphi \mapsto M(\varphi)$ stiftet einen Isomorphismus zwischen den K–Vektorräumen $Hom_K(K^n, K^m)$ und $M_{m,n}(K)$.*

Dies ist eine offensichliche Konsequenz aus der Definition und §11, Bemerkung (II). Es ist nur noch die K-Linearität der Abbildung $\varphi \mapsto M(\varphi)$ zu zeigen.

Beweis: $M(\lambda\varphi + \mu\psi)$ hat in der i-ten Spalte den Vektor

$$(\lambda\varphi + \mu\psi)(e_i) = \lambda\varphi(e_i) + \mu\psi(e_i).$$

Dies ist aber auch der i-te Spaltenvektor von $\lambda M(\varphi) + \mu M(\psi)$ bei obiger Definition der Vektorraumstruktur auf $M_{m,n}(K)$. Also gilt

$$M(\lambda\varphi + \mu\psi) = \lambda M(\varphi) + \mu M(\psi).$$

Korollar:

$$\boxed{dim_K(Hom_K(K^n, K^m)) = mn}$$

Beweis: $M_{m,n}(K)$ ist offensichtlich isomorph zu K^{mn}.

Bemerkung: Wir identifizieren $M_{m,1}(K)$ mit K^m und ebenso $M_{1,n}(K)$ mit $(K^n)^*$.

20 Matrizenmultiplikation

Wir betrachten nun Abbildungen $\varphi \in Hom_K(K^k, K^n)$ und $\psi \in Hom_K(K^n, K^m)$

$$K^k \xrightarrow{\varphi} K^n \xrightarrow{\psi} K^m$$

und ihre Komposition

$$K^k \xrightarrow{\psi \circ \varphi} K^m$$

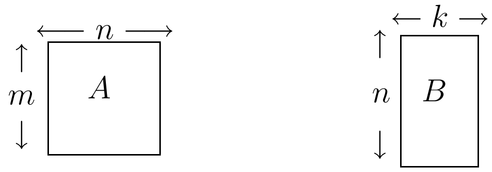
Frage: Wie bestimmt man $M(\psi \circ \varphi)$ aus $M(\varphi)$ und $M(\psi)$?

Dies führt uns zur Definition der Matrizenmultiplikation: Setzt man $A := M(\psi)$ und $B := M(\varphi)$, dann kann man $A \cdot B$ geeignet definieren, so daß

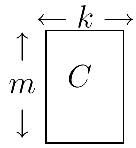
$$A \cdot B = M(\psi \circ \varphi)$$

gilt.

Definition: Seien $A^{m,n}$ und $B^{n,k}$ zwei Matrizen vom Typ:



mit Einträgen $a_{ij}, b_{uv} \in K$. Wir definieren als Matrizenprodukt $C = C^{m,k} = A \cdot B$ die Matrix



mit folgenden Einträgen c_{ij} :

$$\boxed{c_{ij} = \sum_{\gamma=1}^{n} a_{i\gamma} b_{\gamma j}}$$

Die Einträge von C berechnen sich also als Skalarprodukt von Zeilen- und Spaltenvektor von A und B.

Achtung: Im allgemeinen gilt $A \cdot B \neq B \cdot A$.

Zwei Beispiele:

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

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

oder

$$(\mu_1, ..., \mu_n) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \left(\sum_{i=1}^{n} \mu_i \lambda_i \right) = \sum_{i=1}^{n} \mu_i \lambda_i$$

<u>Schreibweise</u>: Bei 1×1 Matrizen lassen wir (oft) die Klammer fort.

Zeilen- bzw. Spaltenvektoren werden als Matrizen von Typ $1 \times n$ bzw. $n \times 1$ behandelt.

Behauptung:

$$\boxed{M(\psi \circ \varphi) = M(\psi) \cdot M(\varphi)}$$

Beachte: $M(\psi \circ \varphi)$ beschreibt die Komposition der Abbildungen durch das Matrizenprodukt $M(\psi) \cdot M(\varphi)$.

<u>Beweis</u>: Sei $M(\psi) = A$ und $M(\varphi) = B$. Dann gilt per Definition:

$$\begin{array}{ll} b_{\gamma j} = \varphi(e_j)_\gamma & \gamma\text{-te Komponente von } \varphi(e_j) \\ a_{i\gamma} = \psi(e_\gamma)_i & i\text{-te Komponente von } \psi(e_\gamma) \end{array}$$

also

$$\begin{aligned}
(\psi \circ \varphi)(e_j) &= \psi(\varphi(e_j)) \\
&= \psi\left(\sum_{\gamma=1}^{n} b_{\gamma j} e_\gamma\right) \\
&= \sum_{\gamma=1}^{n} b_{\gamma j} \psi(e_\gamma) \\
&= \sum_{\gamma=1}^{n} b_{\gamma j} \left(\sum_{i=1}^{m} a_{i\gamma} e_i\right) \\
&= \sum_{i=1}^{m} \left(\sum_{\gamma=1}^{n} a_{i\gamma} b_{\gamma j}\right) e_i
\end{aligned}$$

Somit gilt

$$\sum_{\gamma=1}^{n} a_{i\gamma} b_{\gamma j} = (\psi \circ \varphi)(e_j)_i.$$

Bemerkung: Aus §19 und der obigen Behauptung folgt, daß sich die Kompositionseigenschaften von K-linearen Abbildungen, wie z.B. das Assoziativgesetz und die Linearität, als Eigenschaften des Matrizenproduktes formulieren lassen. Es folgt also insbesondere:

$$\begin{aligned}
(A \cdot B) \cdot C &= A \cdot (B \cdot C) \\
(A + B) \cdot C &= A \cdot C + B \cdot C \\
A \cdot (B + C) &= A \cdot B + A \cdot C
\end{aligned}$$

<u>Rückgewinnung der Abbildung aus der Matrix</u>: Wir wollen uns dazu an die Identifikation von Spaltenvektoren und $n \times 1$-Matrizen erinnern:

$$M_{n,1}(K) = K^n.$$

Behauptung: *Es gilt*

$$\boxed{\varphi(x) = M(\varphi) \cdot x}$$

Man beachte, daß in dieser Formel der Funktionswert $\varphi(x)$ durch das Matrixprodukt $M(\varphi) \cdot x$ beschrieben wird.

Beweis: $x = \sum_{i=1}^{n} x_i e_i$. Es gilt

$$M(\varphi) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \vdots & & \vdots \\ \varphi(e_1) & \cdots & \varphi(e_n) \\ \vdots & & \vdots \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

nach der Definition von $M(\varphi)$. Andererseits gilt

$$\varphi(x) = x_1\varphi(e_1) + \cdots + x_n\varphi(e_n) = \begin{pmatrix} \vdots & & \vdots \\ \varphi(e_1) & \cdots & \varphi(e_n) \\ \vdots & & \vdots \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

nach der Definition des Matrizenprodukts. Daraus folgt die Behauptung.

Auf Grund der obigen Zusammenhänge zwischen K-linearen Abbildungen und Matrizen wollen wir im folgenden keinen Unterschied mehr machen zwischen linearen Abbildungen in $Hom_K(K^n, K^m)$ und Matrizen in $M_{m,n}(K)$.

21 *Duale Abbildungen in Termen der Matrix

Sei $\varphi : K^n \to K^m$ gegeben mit mit $\varphi(e_i) = \sum_{i=1}^m A_{ij} e_j$ und der zugehörigen Matrix $A = M(\varphi)$

$$A = (A_{ij}) \quad \in \; M_{m,n}(K) \; .$$

Die sogenannte transponierte Matrix $A' \in M_{n,m}(K)$ von A entsteht aus A durch Spiegeln (Vertauschen von Zeilen und Spalten)

$$(A')_{ij} = A_{ji} \quad , \quad \forall i, j \; .$$

Weiterhin gilt: Ist $A = M(\varphi)$ die Matrix von φ, dann ist $A' = M(\varphi^*)$ die Matrix der dualen Abbildung.

Genauer: Per Definition bildet die duale Abbildung

$$\varphi^* : (K^m)^* \to (K^n)^*$$

Elemente v^* des Dualraums $(K^n)^*$ ab auf die Abbildungskomposition $v^* \circ \varphi$. Somit gilt $\varphi^*(e_j^*) = e_j^* \circ \varphi$ für alle $i = 1,..m$. Es folgt für alle Vektoren e_i mit $i = 1,..,n$

$$\varphi^*(e_j^*)(e_i) = (e_j^* \circ \varphi)(e_i) = e_j^*(\varphi(e_i)) = e_j^*(\sum_{k=1}^m A_{ik} e_k) = A_{ij}\delta_{ij} = \sum_{k=1}^n A_{kj} e_k^*(e_i) \; .$$

Da dies für alle $i = 1,..n$ gilt, erhält man

$$\varphi^*(e_j^*) = \sum_{k=1}^n A_{kj} e_k^* = \sum_{k=1}^n (A')_{jk} e_k^* \; .$$

Mit anderen Worten: Der Übergang von der Abbildung $\varphi : K^n \to K^m$ zur dualen Abbildung $\varphi^* : (K^m)^* \to (K^n)^*$ entspricht auf dem Matrizenniveau dem Übergang von der Matrix $A = M(\varphi)$ zur transponierten Matrix $A' = M(\varphi^*)$.

Korollar: *Aus §17, Satz 2 folgt daher*

$$dim_K(Bild(A)) = dim_K(Bild(A'))$$

für jede lineare Abbildung $A : K^n \to K^m$.

Wir geben für die Aussage dieser Folgerung später einen unabhängigen Beweis in der Sprache der Matrizen!

22 Der Rang einer Matrix

Definition: Der Rang $r(A)$ einer Matrix $A \in M_{m,n}(K)$ ist definiert als die maximale Anzahl der Spaltenvektoren von A, welche linear unabhängig sind.

$$A = \begin{pmatrix} \vdots & \vdots & & \vdots \\ a_1 & a_2 & \cdots & a_n \\ \vdots & \vdots & & \vdots \end{pmatrix}$$

Die Spaltenvektoren a_i liegen in K^m.

Für den Rang gilt offensichtlich

$$0 \leq r(A) \leq n.$$

Beispiel: $r\begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} = 1$, denn $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ oder $\begin{pmatrix} 3 \\ 6 \end{pmatrix}$ ist jeweils ein maximales System linear unabhängiger Vektoren.

Für die obige Definition ist zu zeigen, daß sie unabhängig ist von der Wahl einer maximalen Teilmenge von linear unabhängigen Spaltenvektoren.

Wir zeigen dazu

Lemma: $r(A) = dim_K(Bild(A : K^n \to K^m))$

Beweis: Da $e_1, ..., e_n$ eine Basis von K^n ist, bilden die Bildvektoren $A(e_1), ..., A(e_n)$ (d.h. die Spalten von A) ein Erzeugendensystem von $Bild(A)$. Wie im Beweis des Basisergänzungssatzes gezeigt wurde, bildet jedes maximal linear unabhängige Teilsystem eines Erzeugendensystems eines Vektorraums eine Basis des Raums, in unserem Fall von $Bild(A)$. Es folgt $r(A) = dim_K(Bild(A))$.

Wegen $Bild(A) \subseteq K^m$ gilt insbesondere $r(A) \leq m$.

Hauptsatz (über Matrizen): *Es gilt Spaltenrang = Zeilenrang, also*

$$\boxed{r(A) = r(A')}$$

oder mit anderen Worten: Die maximale Anzahl linear unabhängiger Spaltenvektoren ist gleich der maximalen Anzahl linear unabhängiger Zeilenvektoren.

Beweis: $r(A) = dim_K(Bild(A)) = dim_K(Bild(A')) = r(A')$ wegen §21 und §17.

Der obige Satz ist eine der fundamentalsten Aussagen über Matrizen. Der obige Beweis benutzt Dualität. Der interessierte Leser möge dazu die entsprechenden (optionalen) Abschnitte des letzten Kapitels studieren. Auf Grund der Bedeutung der Aussage geben wir in den folgenden Abschnitten einen zweiten und davon unabhängigen Beweis des obigen Satzes in der Sprache der Matrizen. Dem Leser wird empfohlen zuerst den Beweis in der Sprache der Matrizen zu verstehen. Siehe Folgerung 3 in §25.

Beispiel: $A := \begin{pmatrix} 1 & 3 & 4 \\ 2 & 6 & 8 \end{pmatrix}$

$r(A) = 1$, betrachte z. B. $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und $r(A') = 1$, betrachte z. B. $(1 \quad 3 \quad 4)$

Wir ziehen nun einige Folgerungen aus dem Hauptsatz!

Folgerung 1: *Es gilt* $r(A) \leq \min\{n, m\}$

Folgerung 2: *Der Rang einer Matrix ändert sich nicht beim Vertauschen von Zeilen (Vertauschen von Spalten).*

Beweis: Für Spalten folgt dies aus der Definition, für Zeilenvertauschungen durch Zurückführen auf Spaltenvertauschungen mittels Transposition der Matrix und obigem Satz.

Wir schließen diesen Abschnitt mit einigen Definitionen, welche uns in der Sprache der Abbildungen bereits vertraut sind.

Definition: Eine Matrix $A \in M_{n,n}(K)$ heißt <u>regulär</u>, falls eine Matrix $B \in M_{n,n}(K)$ existiert mit $B \cdot A = E = E^{(n,n)}$.

<u>Bemerkung</u>: Offensichtlich ist A regulär genau dann, wenn die dazugehörige lineare Abbildung $x \mapsto A \cdot x$ ein Isomorphismus von K^n ist, da B der Umkehrfunktion dieser Abbildung entspricht.

<u>Bezeichnung</u>: $\mathcal{G}\ell(n, K)$ sei die Menge der regulären $n \times n$-Matrizen.

$\mathcal{G}\ell(n, K)$ bildet eine Gruppe bezüglich der Matrizenmultiplikation, $E^{(n,n)}$ ist das neutrale Element.

<u>Bemerkung</u>: Sei $A \in M_{n,n}(K)$. Dann sind äquivalent:

1. $A \in \mathcal{G}\ell(n, K)$
2. $r(A) = n$
3. $r(A') = n$
4. $\exists B \in M_{n,n}(K)$ mit $AB = E$
5. $\exists C \in M_{n,n}(K)$ mit $CA = E$

Wir geben noch einen Vergleichskatalog der wichtigsten Eigenschaften:

<u>Lineare Abbildungen</u>	<u>Matrizen</u>
$\varphi \circ \psi$ (Komposition)	$A \cdot B$ (Produkt)
$\lambda\varphi + \mu\psi$ (Summe)	$\lambda A + \mu B$ (Summe)
$(\varphi \circ \psi) \circ \phi = \varphi \circ (\psi \circ \phi)$	$(A \cdot B) \cdot C = A \cdot (B \cdot C)$
φ^* (Dual)	A' (transponierte Abbildung)
$(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$	$(A \cdot B)' = B' \cdot A'$
$(\lambda\varphi + \mu\psi)^* = \lambda\varphi^* + \mu\psi^*$	$(\lambda A + \mu B)' = \lambda A' + \mu B'$
$\varphi \circ (\lambda\psi + \mu\phi) = \lambda\varphi \circ \psi + \mu\varphi \circ \phi$	$A \cdot (\lambda B + \mu C) = \lambda A \cdot B + \mu A \cdot C$
$(\lambda\varphi + \mu\psi) \circ \phi = \lambda\varphi \circ \phi + \mu\psi \circ \phi$	$(\lambda A + \mu B) \cdot C = \lambda A \cdot C + \mu B \cdot C$
id_{K^n}	$E = E^{(n,n)}$
$\varphi : K^n \to K^n$ ist ein Isomorphismus	$A^{(n,n)}$ ist regulär
φ^{-1}	A^{-1}
$\varphi \circ \varphi^{-1} = \varphi^{-1} \circ \varphi = id_{K^n}$	$A \cdot A^{-1} = A^{-1} \cdot A = E^{(n,n)}$
$\mathcal{G}\ell(K^n)$	$\mathcal{G}\ell(n, K)$
$r(\varphi) = dim_K(Bild(\varphi))$	$r(A)$

23 Lineare Gleichungssysteme

Gegeben seien Koeffizienten $(a_{\nu,\mu})_{\substack{\nu=1,\dots,m\\ \mu=1,\dots,n}}$ und b_ν ($\nu = 1, ..., m$) aus einem Körper K. Gesucht werden Lösungsvektoren

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

mit den Koeffizienten $x_i \in K$, welche folgendes lineare Gleichungssystem erfüllen (hierbei durchläuft i die Zahlen $i = 1, ..., n$):

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

Für die Matrix $A = (a_{\nu,\mu}) \in M_{m,n}(K)$ und den Vektor

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^m$$

ist das obige lineare Gleichungssystem äquivalent zu der Matrixgleichung

$$A \cdot x = b.$$

Man beachte, daß hierbei $A \cdot x$ die Matrizenmultiplikation bezeichnet.

Aus dieser Umformulierung in die Sprache der Matrizen ergeben sich sofort die beiden fundamentalen Folgerungen:

Folgerung 1: *Das lineare Gleichungssystem besitzt eine Lösung genau dann, wenn $b \in Bild(A)$, wobei A als lineare Abbildung von K^n nach K^m aufgefaßt wird.*

$$\begin{array}{cccc} A: & K^n & \to & K^m \\ & \in & & \in \\ & x & \mapsto & b \end{array}$$

Folgerung 2: *Im Fall, daß überhaupt eine Lösung $x = x_0$ des linearen Gleichungssystems existiert, ist jede andere Lösung x des linearen Gleichungssystems von der Gestalt*

$$x = x_0 + u \qquad\qquad u \in Kern(A).$$

Umgekehrt ist jeder solcher Vektor eine Lösung. Die Lösungsmenge des linearen Gleichungssystems ist also

$$x_0 + Kern(A) = \{x_0 + u \mid u \in Kern(A)\}.$$

Beweis: Sei x_0 eine feste Lösung, das heißt $A \cdot x_0 = b$. Dann gilt für jede weitere Lösung:

$$\begin{aligned} A \cdot x = b \quad &\Longleftrightarrow \quad A \cdot x - A \cdot x_0 = 0 \\ &\Longleftrightarrow \quad A(x - x_0) = 0 \\ &\Longleftrightarrow \quad x - x_0 = u, \quad u \in Kern(A) \\ &\Longleftrightarrow \quad x = x_0 + u, \quad u \in Kern(A) \end{aligned}$$

Bezeichnung: Ein Gleichungssystem $A \cdot x = b$ wie oben heißt inhomogenes lineares Gleichungssystem. Das Gleichungssystem $A \cdot x = 0$ ist das zugeordnete homogene lineare Gleichungssystem. Ist $A \in M_{m,n}(K)$, dann ist n die Zahl der Unbekannten und m die Zahl der Gleichungen.

Bemerkung: $Kern(A)$ kann als Lösungsmenge des zugeordneten homogenen Gleichungssystems aufgefaßt werden.

Nun wollen wir uns dem Fall des homogenen Gleichungssystems zuwenden:

Jedes homogene lineare Gleichungssystem besitzt eine Lösung, nämlich die triviale Lösung:

$$x = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0.$$

Die Menge der Lösungen ist $Kern(A)$.

$$\begin{array}{cccc} A: & K^n & \to & K^m \\ & \in & & \in \\ & x & \mapsto & 0 \end{array}$$

Daß immer eine Lösung existiert, ist im allgemeinen nicht mehr richtig für das allgemeine inhomogene lineare Gleichungssystem.

Definition: Ein lineares Gleichungssystem $A \cdot x = b$ heiß universell lösbar, wenn das Gleichungssystem für jede Wahl von b eine Lösung x besitzt (wobei A fest gewählt sei). Offensichtlich ist $A \cdot x = b$ genau dann universell lösbar, wenn $Bild(A) = K^m$ gilt oder gleichbedeutend

$$A : K^m \to K^n$$

surjektiv ist.

Nicht jedes lineare Gleichungssystem ist universell lösbar. Wir formulieren dazu das folgende Kriterium:

Gegeben seien $A \in M_{m,n}(K)$ und $b \in K^m$. Gesucht ist ein $x \in K^n$ mit $A \cdot x = b$.

Bezeichnung: $(A, b) \in M_{m,n+1}(K)$ heißt erweiterte Matrix (gebildet zum obigen linearen Gleichungssystem definiert durch A und b).

Satz: *Das lineare Gleichungssystem gebildet zu A und b besitzt genau dann eine Lösung, wenn gilt:*

$$r(A) = r(A, b).$$

Beweis: Das lineare Gleichungssystem besitzt genau dann eine Lösung, wenn $b \in Bild(A)$ gilt, d.h. genau dann, wenn b eine Linearkombination der Spaltenvektoren von A ist.

Offensichtlich gilt $r((A, b)) \geq r(A)$ und somit gilt $r((A, b)) = r(A)$ genau dann, wenn b linear abhängig ist von den Spaltenvektoren von A ist (also wenn b eine Linearkombination der Spaltenvektoren ist).

Quadratische Matrizen: Wir betrachten nun den Fall $m = n$, das heißt die Zahl der Variablen ist gleich der Zahl der Gleichungen.

Aus dem Dimensionssatz (§13) folgt in diesem Fall, daß $A \cdot x = b$ eindeutig lösbar ist genau dann, wenn das Gleichungssystem universell lösbar ist. Dies ist offensichtlich genau dann der Fall, wenn

$$A : K^n \xrightarrow{\simeq} K^n$$

ein Isomorphismus ist, d.h. wenn A regulär ist. Dieser Fall ist besonders wichtig, da man die Lösung sofort durch Kenntnis der inversen Matrix A^{-1} erhält:

$$x = A^{-1} \cdot b.$$

Alle folgenden Abschnitte bis einschließlich §40 beschäftigen sich mit mehr oder weniger direkt mit dem Problem der Bestimmung von $r(A)$ und von A^{-1} und stehen somit in unmittelbarem Zusammenhang mit der Lösung von linearen Gleichungssystemen. In den nächsten beiden Abschnitten geben wir einen de facto Algorithmus zur Bestimmung des Rangs einer Matrix an (siehe Folgerung 2 im übernächsten Abschnitt §25).

24 Elementare Matrizenumformungen

Es sei $E_{\nu,\mu}(\lambda)$ die $n \times n$-Matrix, welche folgende Gestalt hat:

$$E_{\nu,\mu}(\lambda) = \begin{pmatrix} 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ 0 & \vdots & \ddots & & & & \vdots & 0 \\ \vdots & \vdots & & \ddots & & & \vdots & \vdots \\ \vdots & \vdots & & & \ddots & \lambda & \vdots & \vdots \\ 0 & \vdots & & & & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 \end{pmatrix},$$

$E_{\nu,\mu}(\lambda) = (a_{ij})$ mit $a_{ij} = \lambda\delta_{i\nu}\delta_{j\mu}$.

Als <u>elementare Matrizen</u> bezeichnen wir folgende Matrizen:

$$Diag(a_1, ..., a_n) = \begin{pmatrix} a_1 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & a_n \end{pmatrix}, \text{ wobei } a_1 \neq 0, ..., a_n \neq 0$$

und

$$D_{\nu,\mu}(\lambda) = E + E_{\nu,\mu}(\lambda) = \begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 \\ 0 & 1 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ 0 & \vdots & \ddots & & & & \vdots & 0 \\ \vdots & \vdots & & \ddots & & & \vdots & \vdots \\ \vdots & \vdots & & & \ddots & \lambda & \vdots & \vdots \\ 0 & \vdots & & & & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & 1 \end{pmatrix},$$

für alle $\nu \neq \mu$, wobei $\lambda \in K$ beliebig sein darf.

Es gelten die folgenden einfachen Rechenregeln:

Regel 1:

$$\boxed{Diag(a_1,...,a_n)\cdot Diag(b_1,...,b_n) = Diag(a_1b_1,...,a_nb_n)}$$

Der Beweis der Regel 1 erfolgt durch einfaches Nachrechnen.

Folgerung: *$Diag(a_1^{-1},...,a_n^{-1})$ ist die inverse Matrix zu $Diag(a_1,...,a_n)$.*

Regel 2:

$$\boxed{D_{\nu,\mu}(\lambda)\cdot D_{\nu,\mu}(\lambda') = D_{\nu,\mu}(\lambda+\lambda')}$$

Beweis:

$$\begin{aligned}(E+E_{\nu,\mu}(\lambda))\cdot(E+E_{\nu,\mu}(\lambda')) &= E+E_{\nu,\mu}(\lambda)+E_{\nu,\mu}(\lambda')+E_{\nu,\mu}(\lambda)E_{\nu,\mu}(\lambda')\\ &= E+E_{\nu,\mu}(\lambda+\lambda')+E_{\nu,\mu}(\lambda)E_{\nu,\mu}(\lambda')\\ &= D_{\nu,\mu}(\lambda+\lambda')+E_{\nu,\mu}(\lambda)E_{\nu,\mu}(\lambda')\end{aligned}$$

Daraus folgt die Regel 2, denn es gilt: $E_{\nu,\mu}(\lambda)E_{\nu,\mu}(\lambda') = 0$.

Beweis: $(E_{\nu,\mu}(\lambda)E_{\nu,\mu}(\lambda'))_{ij} = \sum_{k=1}^{n} \underbrace{(\lambda\delta_{i\nu}\delta_{k\mu})}_{a_{ik}}\underbrace{(\lambda'\delta_{k\nu}\delta_{j\mu})}_{a_{kj}} = \lambda\lambda'\sum_{k=1}^{n}\delta_{i\nu}\delta_{k\mu}\delta_{k\nu}\delta_{j\mu}$

Dieser Ausdruck ist 0, außer wenn gilt $i=\nu, k=\mu, k=\nu$ und $j=\mu$. Insbesondere muß gelten: $\mu=k=\nu$, wenn der zugehörige Summand $\neq 0$ sein soll. Da wir $\nu\neq\mu$ angenommen hatten, ist die Summe $=0$ für alle i,j.

Folgerung: *$D_{\nu,\mu}(-\lambda)$ ist die inverse Matrix zu $D_{\nu,\mu}(\lambda)$.*

Es folgt unmittelbar

Lemma: *Die elementaren Matrizen $Diag(a_1,...,A_n)$ und $D_{\nu,\mu}(\lambda)$ sind invertierbar, d.h. sie haben Rang $=n$.*

Hilfssatz 1: *Sei* $A \in M_{m,n}(K)$ *und* $U \in \mathcal{G}\ell(n,K)$*, dann gilt* $r(A \cdot U) = r(A)$.

Der Beweis folgt unmittelbar aus der analogen Aussage Hilfssatz'.

Hilfssatz 1': *Sei* $\varphi : K^n \to K^m$ *eine* K*–lineare Abbildung und* $\psi : K^n \to K^n$ *ein Isomorphismus, dann gilt*

$$dim_K(Bild(\varphi)) = dim_K(Bild(\varphi \circ \psi)).$$

Beweis: $Bild(\varphi \circ \psi) = (\varphi \circ \psi)(K^n) = \varphi(\psi(K^n)) \overset{\psi\ surj}{=} \varphi(K^n) = Bild(\varphi)$

Hilfssatz 2: *Sei* $A \in M_{m,n}(K)$ *und* $V \in \mathcal{G}\ell(m,K)$*, dann gilt* $r(V \cdot A) = r(A)$.

Beweis: Sei $\tilde{\psi}$ die bijektive K-lineare Abbildung $\tilde{\psi} : K^m \to K^m$, welche durch die invertierbare Matrix V definiert wird. Zur Abkürzung sei $W = Bild(\varphi)$. Dann ist $r(A) = dim_K(W)$. Wegen $r(V \cdot A) = dim_K(Bild(\tilde{\psi} \circ \varphi))$ und $Bild(\tilde{\psi} \circ \varphi)) = \tilde{\psi}(Bild(\varphi)) = \tilde{\psi}(W)$ genügt daher zum Beweis

$$dim_K(W) = dim_K(\tilde{\psi}(W)) \ .$$

Die gewünschte Dimensionsgleichheit folgt somit aus dem 2.Struktursatz §12, denn die Einschränkung des Isomorphismus $\tilde{\psi}$ auf den Untervektorraum $W \subset K^m$ definiert eine bijektive K-lineare Abbildung (d.h. einen Vektorraumisomorphismus)

$$\tilde{\psi}|W : \quad W \longrightarrow \tilde{\psi}(W) \ .$$

Nun wollen wir folgende Strategie verwenden:

Wir wenden die beiden Hilfssätze an für elementare Matrizen U oder V. Wir versuchen so die Gestalt der ursprünglichen Matrix A zu vereinfachen. Hilfreich ist dabei, daß die Rechts- und Linksmultiplikation mit elementaren Matrizen auf einfache Weise gedeutet werden kann.

Wir beschränken uns auf die Erläuterung im Fall der Linksmultiplikation. Die grundlegende Beobachtung ist dabei folgende:

$$Diag(a_1,...,a_m) \cdot \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} a_1 a_{11} & \cdots & a_1 a_{1n} \\ \vdots & & \vdots \\ a_m a_{m1} & \cdots & a_m a_{mn} \end{pmatrix}$$

sowie

$$E_{\nu,\mu}(\lambda)\cdot\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn}\end{pmatrix} = \begin{pmatrix} & 0 & \\ \lambda a_{\mu 1} & \cdots & \lambda a_{\mu n} \\ & 0 & \end{pmatrix} \quad \begin{array}{c}\textit{das } \lambda\textit{-fache der } \mu\textit{-ten Zeile} \\ \textit{steht in der } \nu\textit{-ten Zeile}\end{array}$$

$$\begin{aligned} D_{\nu,\mu}(\lambda)\cdot\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn}\end{pmatrix} &= (E + E_{\nu,\mu}(\lambda))\cdot\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn}\end{pmatrix} \\ &= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{\nu 1} + \lambda a_{\mu 1} & \cdots & a_{\nu n} + \lambda a_{\mu n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn}\end{pmatrix} \end{aligned}$$

Zusammengefaßt läßt sich wegen Hilfsatz 2 sagen:

Der Rang $r(A)$ einer Matrix $A \in M_{m,n}(K)$ ändert sich nicht bei folgenden Abänderungen von A :

(I) multiplizieren der Zeilen mit nichtverschwindenden Skalaren $a_1, ..., a_n$.

(II) addieren des λ-fachen der μ-ten Zeile zur ν-ten Zeile für beliebige $\lambda \in K$ und $\nu \neq \mu$.

(III) beliebige Vertauschungen von Zeilen (benutze Permutationsmatrizen aus Aufgabe 1 vom Übungsblatt 8)

Bemerkung: Eine analoge Aussage gilt für die entsprechenden Spaltenumformungen, da diese der Rechtsmultiplikation mit elementaren Matrizen entsprechen. Analog ändert sich der Rang nicht bei beliebigen Vertauschungen von Spalten.

Bemerkung: Jede Permutationsmatrix ist ein Produkt von Elementarmatrizen. Da jede Permutation ein Produkt von Transpositionen ist, führt man dies auf folgende Formel zurück

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

25 Eliminationsalgorithmus

Ziel ist, eine gegebene Matrix $A \in M_{m,n}(K)$ schrittweise auf folgende Form zu bringen:

$$(*) \quad \left(\begin{array}{cccccccccccc}
1 & * & \cdots & \cdots & * & * & * & * & * & \cdots & * & * \\
0 & 1 & * & & & * & * & & & & & * \\
0 & 0 & 1 & * & & \vdots & \vdots & & & & & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \vdots & \vdots & & & & & \vdots \\
0 & & & 0 & 1 & * & * & & & & & * \\
0 & 0 & \cdots & 0 & 0 & 1 & * & * & * & \cdots & * & * \\
 & & & & & & & & & & & \\
0 & 0 & \cdots & 0 & 0 & 0 & & & & & & \\
0 & & & & & 0 & & & & & & \\
\vdots & & & & & \vdots & & & & B^{(k)} & & \\
0 & & & & & 0 & & & & & & \\
0 & 0 & \cdots & 0 & 0 & 0 & & & & & &
\end{array}\right)$$

Die Matrix A sei mittels der Verfahren $(I) - (III)$ angewendet auf Zeilen und Spalten (!) auf die Form $(*)$ gebracht worden. Ist $B^{(k)} \neq 0$, dann kann man A weiter vereinfachen bis das neu entstandene $B^{(k')}$ verschwindet. k ist dabei die Zahl der Zeilen und Spalten des Blocks links oben, welcher vereinfacht wurde.

Methode: Sei ein Eintrag $b_{ij}^{(k)}$ von $B^{(k)}$ nicht 0, dann vertausche von A die Zeilen $k+1$ mit i und die Spalten $k+1$ und j. Der Eintrag $b_{ij}^{(k)}$ kommt dadurch an die Stelle $k+1, k+1$.

Nun multipliziere die Zeile $k+1$, so daß $b_{k+1,k+1} = 1$ wird und ziehe geeignete Vielfache der $k+1$-sten Zeile von den Zeilen $k+2, ..., n$ ab.

Als Resultat erhält man:

$$
\begin{pmatrix}
1 & * & \cdots & \cdots & * & * & * & * & * & \cdots & * & * \\
0 & 1 & * & & & * & * & & & & & * \\
0 & 0 & 1 & * & & \vdots & \vdots & & & & & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \vdots & \vdots & & & & & \vdots \\
0 & & & 0 & 1 & * & * & & & & & * \\
0 & 0 & \cdots & 0 & 0 & 1 & * & * & * & \cdots & * & * \\
 & & & & & & & & & & & \\
0 & 0 & \cdots & 0 & 0 & 0 & 1 & * & * & \cdots & * & * \\
0 & & & & & 0 & 0 & & & & & \\
\vdots & & & & & \vdots & \vdots & & & & & \\
0 & & & & & 0 & 0 & & & B^{(k')} & & \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & & & & &
\end{pmatrix}
$$

Zusammenfassung: Durch obiges Verfahren kann jede Matrix $A \in M_{m,n}(K)$ nach endlich vielen Schritten auf die Gestalt

$$
A^{(k)} = \begin{pmatrix}
1 & * & \cdots & \cdots & * & * & * & * & * & \cdots & * & * \\
0 & 1 & * & & & * & * & & & & & * \\
0 & 0 & 1 & * & & \vdots & \vdots & & & & & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \vdots & \vdots & & & & & \vdots \\
0 & & & 0 & 1 & * & * & & & & & * \\
0 & 0 & \cdots & 0 & 0 & 1 & * & * & * & \cdots & * & * \\
 & & & & & & & & & & & \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & & & & & 0 & 0 & & & & & 0 \\
\vdots & & & & & \vdots & \vdots & & & & & \vdots \\
0 & & & & & 0 & 0 & & & & & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0
\end{pmatrix}
$$

mit $0 \leq k \leq \min(m, n)$ gebracht werden. k steht dabei für die Zahl der Zeilen und Spalten des Blockes links oben.

Es gilt dabei $r(A) = r(A^{(k)})$, da die vorgenommenen Umformungen den Rang einer Matrix nicht verändern, wie in §22 und §24 gezeigt wurde.

Fortsetzung der Umformung: Durch elementare Spaltenumformungen kann nun die obige Matrix weiter umgeformt werden, so daß sie die Normalform

$$N = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & & & 0 & 0 & & & & & 0 \\ 0 & 0 & 1 & 0 & & \vdots & \vdots & & & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots & \vdots & & & & & \vdots \\ 0 & & & 0 & 1 & 0 & 0 & & & & & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & & & & & 0 & 0 & & & & & 0 \\ \vdots & & & & & \vdots & \vdots & & & & & \vdots \\ 0 & & & & & 0 & 0 & & & & & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} = \begin{pmatrix} E^{(k,k)} & 0^{(k,n-k)} \\ 0^{(m-k,k)} & 0^{(m-k,n-k)} \end{pmatrix}$$

erhält. Sei nämlich

$$B = \begin{pmatrix} 1 & b_{12} & b_{13} & \cdots & \cdots & b_{1n} \\ 0 & 1 & b_{23} & \cdots & \cdots & b_{2n} \\ 0 & 0 & 1 & b_{34} & \cdots & b_{3n} \\ \vdots & & \ddots & \ddots & & \vdots \end{pmatrix},$$

dann subtrahiere nacheinander

die $b_{12} \cdot (1.Spalte)$ von der 2. Spalte

die $b_{13} \cdot (1.Spalte)$ von der 3. Spalte

$\vdots$

die $b_{1n} \cdot (1.Spalte)$ von der n. Spalte.

Man erhält die neue Matrix, welche sich nur in der ersten Zeile von B unterscheidet. Sie hat die Gestalt:

$$B' = \begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & b_{23} & \cdots & \cdots & b_{2n} \\ 0 & 0 & 1 & b_{34} & \cdots & b_{3n} \\ \vdots & & \ddots & \ddots & & \vdots \end{pmatrix},$$

Nun subtrahiere

die $b_{23} \cdot (2.Spalte)$ von der 3. Spalte

die $b_{24} \cdot (2.Spalte)$ von der 4. Spalte

$\vdots$

die $b_{2n} \cdot (2.Spalte)$ von der n. Spalte.

Iteriert man dieses Verfahren, so erhält man die Matrix N wie oben angegeben.

Bemerkung: Der Rang der entstehenden Matrix $N = \begin{pmatrix} E^{(k,k)} & 0^{(k,n-k)} \\ 0^{(m-k,k)} & 0^{(m-k,n-k)} \end{pmatrix}$ ist offensichtlich gleich k.

Bemerkung: Die Methode, welche eine beliebige Matrix A auf Normalform bringt, erlaubt gleichzeitig eine einfache Bestimmung der Lösungsmenge der Gleichung $Ax = b$.

Die oben beschriebene Eliminationsmethode läßt sich wie folgt zusammenfassen:

Satz: *Sei $A \in M_{m,n}(K)$. Dann existieren Matrizen $U \in \mathcal{G}\ell(m,k)$ und $V \in \mathcal{G}\ell(n,K)$, so daß gilt*

$$U \cdot A \cdot V = \begin{pmatrix} E^{(k,k)} & 0^{(k,n-k)} \\ 0^{(m-k,k)} & 0^{(m-k,n-k)} \end{pmatrix}$$

mit $k = Rang(A) = r(A)$.

Zusatz: *Hierbei können U und V als Produkt von elementaren Matrizen $Diag(a_1, ..., a_n)$ und $D_{\nu,\mu}(\lambda)$ gewählt werden.*

Folgerung 1: *Jede invertierbare Matrix $A \in \mathcal{G}\ell(n,K)$ ist ein Produkt von Elementarmatrizen.*

<u>Beweis</u>: Unser Algorithmus zeigt, daß Produkte U, V von Elementarmatrizen existieren mit
$$U \cdot A \cdot V = \begin{pmatrix} E^{(k,k)} & 0 \\ 0 & 0 \end{pmatrix}.$$
Aus $k = Rang(A) = n$ folgt
$$U \cdot A \cdot V = E$$
oder
$$A = U^{-1}V^{-1} = (D_1 \cdots D_r)^{-1}(D_1' \cdots D_s')^{-1} = D_r^{-1} \cdots D_1^{-1} D_s'^{-1} \cdots D_1'^{-1}.$$
Da inverse Elementarmatrizen wieder Elementarmatrizen sind, folgt die Behauptung.

Zur Erinnerung (siehe §24): Sei $A \in M_{m,n}(K)$ und $U \in \mathcal{G}\ell(m, K)$ und $V \in \mathcal{G}\ell(n, K)$. Dann gilt $r(A) = r(UA) = r(AV) = r(UAV)$.

Definition: Zwei Matrizen A_1 und A_2 und $M_{m,n}$ heißen <u>äquivalent</u>, falls invertierbare Matrizen $U \in \mathcal{G}\ell(m, k)$ und $V \in \mathcal{G}\ell(n, K)$ existieren, so daß gilt
$$U \cdot A_1 \cdot V = A_2.$$

Folgerung 2: *Jedes $A \in M_{m,n}(K)$ ist äquivalent zu*
$$\begin{pmatrix} E^{(k,k)} & 0 \\ 0 & 0 \end{pmatrix}$$
mit $k = r(A)$.

Dies ist nur eine Umformulierung der obigen Sätze mit Hilfe der Definition.

Folgerung 3: *Für jede Matrix A in $M_{m,n}(K)$ gilt Spaltenrang gleich Zeilenrang*
$$r(A) = r(A')\ .$$

<u>Beweis</u>: Ist A äquivalent zu

$$\begin{pmatrix} E^{(k,k)} & 0^{(k,n-k)} \\ 0^{(m-k,k)} & 0^{(m-k,n-k)} \end{pmatrix} ,$$

dann ist A' äquivalent zu

$$\begin{pmatrix} E^{(k,k)} & 0^{(k,m-k)} \\ 0^{(n-k,k)} & 0^{(n-k,m-k)} \end{pmatrix} ,$$

denn $A = UDV$ impliziert $A' = V'D'U'$. Berücksichtige: Die transponierte Matrix einer invertierbaren Matrix sind wieder invertierbar. Es folgt $r(A) = k = r(A')$.

Folgerung 4: *Zwei Matrizen A_1 und A_2 in $M_{m,n}(K)$ sind genau dann äquivalent, wenn gilt $r(A_1) = r(A_2)$.*

<u>Beweis</u>: Sei A_1 äquivalent zu A_2. Dann gilt $r(A_1) = r(A_2)$ nach §24.
Ist umgekehrt $r(A_1) = r(A_2) = k$. Dann existieren invertierbare Matrizen U_1, V_1 bzw $U_2, V_2,$ so daß

$$U_1 \cdot A_1 \cdot V_1 = \begin{pmatrix} E^{(k,k)} & 0 \\ 0 & 0 \end{pmatrix} = U_2 \cdot A_2 \cdot V_2$$

nach obigem Satz. Also ist

$$A_1 = (U_1^{-1}U_2)A_2(V_2V_1^{-1}).$$

Es folgt $A_2 = UA_1V$ mit $U \in \mathcal{G}\ell(m,K)$ und $V \in \mathcal{G}\ell(n,K)$ und A_1 und A_2 sind äquivalent.

26 'Gauß-Bruhat'-Zerlegung

In diesem Abschnitt studieren wir eine Variante der Methode von §25. Wir machen dazu folgende

Annahme: Sei $A \in M_{n,n}(K)$ mit $r(A) = n$.

Behauptung: Dann kann A durch folgende Operationen

(i) $D_{\nu,\mu}(\lambda)$ von links an A multiplizieren, wobei $\nu > \mu$

(ii) Vertauschen von Spalten

(iii) Multiplizieren von Matrizen $Diag(a_1, ..., a_n)$ von links an A, wobei $a_i \neq 0$ für alle $i = 1, ..., n$ gilt

auf die Normalform

$$= \begin{pmatrix} 1 & * & \cdots & \cdots & * & * \\ 0 & 1 & * & & & * \\ 0 & 0 & 1 & * & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & & & 0 & 1 & * \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix}$$

gebracht werden.

<u>Beweis</u>: Wir führen dieselbe Konstruktion durch wie bisher. Wir müssen nur überlegen, daß an der 'kritischen Stelle' des Arguments alles gut geht:

$$\begin{pmatrix} 1 & * & \cdots & \cdots & * & * & * & * & * & \cdots & * & * \\ 0 & 1 & * & & & * & * & & & & & * \\ 0 & 0 & 1 & * & & \vdots & \vdots & & & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots & \vdots & & & & & \vdots \\ 0 & & & 0 & 1 & * & * & & & & & * \\ 0 & 0 & \cdots & 0 & 0 & 1 & * & * & * & \cdots & * & * \\ & & & & & & & & & & & \\ 0 & 0 & \cdots & 0 & 0 & 0 & & & & & & \\ 0 & & & & & 0 & & & & & & \\ \vdots & & & & & \vdots & & & & B & & \\ 0 & & & & & 0 & & & & & & \\ 0 & 0 & \cdots & 0 & 0 & 0 & & & & & & \end{pmatrix}$$

<u>Behauptung</u>: Es gibt einen nichttrivialen Koeffizienten in der ersten Zeile von B.

<u>Beweis</u>: Wäre dies nicht der Fall, dann wäre die $k+1$-te Zeile der Matrix $= 0$ und der Zeilenrang $< n$ (Widerspruch).

Ein Vertauschen von Spalten liefert $b_{1,1} \neq 0$. Dann schließt man weiter wie bisher.

<u>Beachte</u>: Alle $D_{\nu,\mu}(\lambda)$ mit $\nu > \mu$ in (i) und alle Matrizen in (iii) haben die Gestalt

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & \ddots & 0 & & \\ & & & \ddots & & \\ & * & & & \ddots & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} \quad bzw \quad \begin{pmatrix} a_1 & & & & & \\ & \ddots & & & & \\ & & \ddots & 0 & & \\ & & & \ddots & & \\ & * & & & \ddots & \\ & & & & & \ddots \\ & & & & & & a_n \end{pmatrix}$$

mit $a_i \neq 0 \qquad i = 1, ..., n,$ sind also untere Dreiecksmatrizen.

Lemma: *Die Menge der unteren (bzw. oberen) Dreiecksmatrizen mit nichtverschwindenden Diagonaleneinträgen bilden eine Untergruppe der Gruppe $\mathcal{G}\ell(n, K)$.*

Insbesondere sind inverse von unteren Dreiecksmatrizen wieder untere Dreiecksmatrizen.

(ohne Beweis)

Satz (Gauß-Bruhat-Zerlegung): *Jede Matrix $A \in \mathcal{G}\ell(n, K)$ läßt sich nach geeigneter Vertauschung von Spalten als Produkt $U \cdot V$ einer unteren und oberen Dreiecksmatrix*

$$\begin{pmatrix} * & & & & & \\ & \ddots & & & & \\ & & \ddots & 0 & & \\ & & & \ddots & & \\ & * & & & \ddots & \\ & & & & & \ddots \\ & & & & & & * \end{pmatrix} \cdot \begin{pmatrix} * & & & & & \\ & \ddots & & & & \\ & & \ddots & * & & \\ & & & \ddots & & \\ & 0 & & & \ddots & \\ & & & & & \ddots \\ & & & & & & * \end{pmatrix}$$

schreiben.

Beweis: Der Beweis folgt aus dem oben beschriebenen Eliminationsverfahren und aus folgender Beobachtung: Seien $U, A \in M_{n,n}(K)$ und $A = (a_1 \cdots a_n)$ habe die Spaltenvektoren $a_1, ..., a_n \in K^n$. Dann gilt

$$U \cdot A = U \cdot (a_1 \cdots a_n) = (U \cdot a_1 \cdots U \cdot a_n)$$

Die Spaltenvektoren von $U \cdot A$ sind also die Vektoren $U \cdot a_n, ..., U \cdot a_n$. Dies zeigt, daß Spaltenvertauschung mit Linksmultiplikation von Matrizen kommuliert.

Folgerung : Im obigen Eliminationsverfahren kann ich also alle Spaltenvertauschungen sofort auf A anwenden (dies gibt eine neue Matrix, sagen wir $\tilde{A}$) und anschließend alle Operationen vom Typ (i) und (iii) durchführen.

Da die Operationen (i) und (iii) Linksmultiplikation mit unteren Dreiecksmatrizen sind, kann man diese zu einer einzigen Multiplikation mit einer geeigneten unteren Dreiecksmatrix $\tilde{U}$ zusammenfassen (Lemma!). Es folgt

$$\tilde{U} \cdot \tilde{A} = V \qquad \textit{oder} \qquad \tilde{A} = \tilde{U}^{-1} \cdot V.$$

Setzt man $U = \tilde{U}^{-1}$, dann ist dies wieder eine untere Dreiecksmatrix (Lemma). Es folgt

$$\tilde{A} = U \cdot V.$$

q.e.d.

27 Matrizen und Abbildungen (II)

Sei nun V ein K–Vektorraum versehen mit einer festen Basis $B = b_1, \ldots, b_n$. Die Abbildung

$$\begin{array}{cccc} M_B : & V & \xrightarrow{\simeq} & K^n \\ & v = \sum\limits_{i=1}^{n} \lambda_i b_i & \longmapsto & \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \end{array}$$

welche jedem Vektor $v \in V$ seine B-Koordinaten zuordnet, ist offensichtlich ein Isomorphismus. Die Umkehrabbildung bildet die Standardbasis $e_1, .., e_n$ von K^n auf B ab.

$$M_B^{-1}(e_i) = b_i \qquad \forall i = 1, ..., n$$

Sei nun W ein zweiter K–Vektorraum mit Basis $B' = b'_1, \ldots, b'_n$ und

$$\varphi : V \to W$$

eine K–lineare Abbildung. Wir wollen dann φ wieder eine Matrix zuordnen. Dies geschieht wie folgt. Wir betrachten das Diagram

$$\begin{array}{ccc} V & \xrightarrow{\qquad\varphi\qquad} & W \\ {\scriptstyle M_B}\Big\downarrow \cong & & \cong \Big\downarrow{\scriptstyle M_{B'}} \\ K^n & \xrightarrow{M_{B'} \circ \varphi \circ M_B^{-1}} & K^m \end{array}$$

Die Idee ist es – nicht φ sondern – die K–lineare Abbildung

$$M_{B'} \circ \varphi \circ M_B^{-1} : \quad K^n \longrightarrow K^m$$

zu beschreiben. Letzterer entspricht nach §19 einer Matrix, welche wir

$$M_{BB'}(\varphi) \in M_{m,n}(K)$$

nennen wollen. Identifiziert man diese Matrix wieder mit der dazugehörenden Abbildung in $Hom_K(K^n, K^m)$, dann gilt:

$$M_{BB'}(\varphi) = M_{B'} \circ \varphi \circ M_B^{-1}.$$

Offensichtlich gilt für Vektorräume U, V, W mit Basen $B, B', B"$ der folgende

Satz: : *Seien* $U \xrightarrow{\varphi} V$ *und* $V \xrightarrow{\psi} W$ K*–lineare Abbildungen. Dann gilt*

$$M_{B'B"}(\psi) \cdot M_{BB'}(\varphi) = M_{BB"}(\psi \circ \varphi)$$

Beweis:

$$\begin{aligned} M_{B'B"}(\psi) \cdot M_{BB'}(\varphi) &= (M_{B"} \circ \psi \circ M_{B'}^{-1}) \circ (M_{B'} \circ \varphi \circ M_B^{-1}) \\ &= M_{B"} \circ (\psi \circ \varphi) \circ M_B^{-1} = M_{BB"}(\psi \circ \varphi) \ . \end{aligned}$$

Direkte Beschreibung der Matrix: Schließlich findet man auf Grund der Definition die folgende "direkte" Beschreibung von $M_{BB'}(\varphi)$:

$$M_{BB'}(\varphi) = (\, M_{B'}(\varphi(b_1)) \cdots M_{B'}(\varphi(b_n)) \,)$$

Beweis: Die Abbildung $M_{BB'}(\varphi) = M_{B'} \circ \varphi \circ M_B^{-1}$ bildet den Standardbasisvektor e_i ab auf die i-te Spalte der Matrix. Dies ist

$$(M_{B'} \circ \varphi \circ M_B^{-1})(e_i) = (M_{B'} \circ \varphi)(b_i) = M_{B'}(\varphi(b_i))$$

28 Basiswechsel

In diesem Abschnitt soll untersucht werden, wie sich bei festem $\varphi : V \to W$ die zugehörige Matrix $M_{BB'}(\varphi) \in M_{m,n}(K)$ ändert, wenn man die Basen B bzw. B' abändert.

Zur Erinnerung: B ist Basis von V, B' ist Basis von W.

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d, "M_B"', "\cong"] & W \arrow[d, "\cong"', "M_{B'}"] \\ K^n \arrow[r, "M_{B'}\circ\varphi\circ M_B^{-1}"] & K^m \end{tikzcd}$$

Wir wollen nun $M_{BB'}(\varphi) = M_{B'} \circ \varphi \circ M_B^{-1}$ schreiben.

Wählt man nun eine andere Basis C resp. C' von V resp. W, dann gilt:

$$\begin{aligned} M_{CC'}(\varphi) &= M_{C'} \circ \varphi \circ M_C^{-1} = M_{C'} \circ M_{B'}^{-1} \circ (M_{B'} \circ \varphi \circ M_B^{-1}) \circ M_B \circ M_C^{-1} \\ &= M_{C'} \circ M_{B'}^{-1} \circ M_{BB'}(\varphi) \circ M_B \circ M_C^{-1} \ . \end{aligned}$$

Im Spezialfall $V = W$, $B = B'$ und $C = C'$ setze $T := M_C \circ M_B^{-1}$

$$\begin{tikzcd} K^n \arrow[rr, "T"] \arrow[dr, "\simeq", "M_B^{-1}"'] & & K^n \\ & V \arrow[ur, "\simeq", "M_C"'] & \end{tikzcd} \ .$$

Dann gilt $T \in \mathcal{G}\ell(n, K)$ sowie

$$M_{CC}(\varphi) = T \circ M_{BB}(\varphi) \circ T^{-1} \ .$$

Dies beschreibt die Änderung der Matrix des Endomorphismus $\varphi : V \to V$ beim Übergang von der Basis B zur Basis C. Die Matrix T heißt Basiswechselmatrix.

<u>Bemerkung</u>: Um von Matrizen N und M aus $M_{n,n}(K)$ zu zeigen, daß eine Matrix $T \in \mathcal{G}\ell(n, K)$ existiert mit

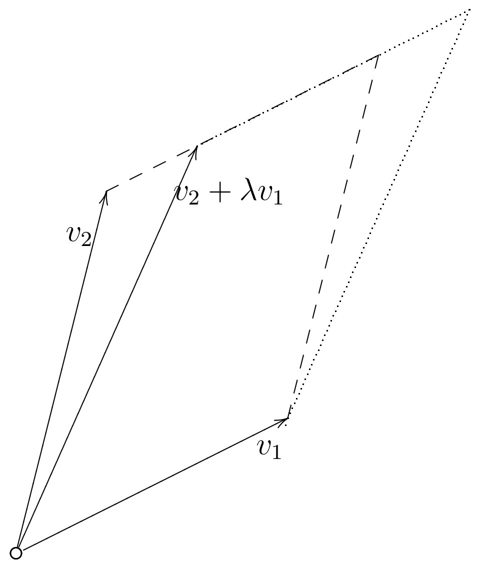
$$N = TMT^{-1}$$

genügt es also nach obigen Ausführungen zu wissen, daß N und M als Matrixdarstellung ein und desselben Endomorphismus $\varphi \in End(V)$ zu verschiedenen Basen C und B von V aufgefaßt werden können. Dies werden wir uns in §39 und §40 zu Nutzen machen!

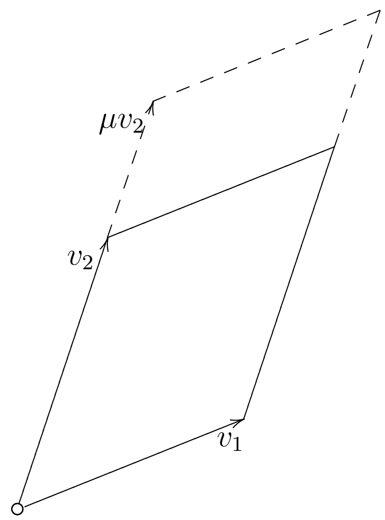
Determinantentheorie

29 Determinantenfunktionen

Die Fläche der Raute $[v_1, v_2]$ in der Ebene $\mathbb{R}^2$ ist scherungsinvariant $D(v_1, v_2) = D(v_1, v_2 + \lambda v_1)$

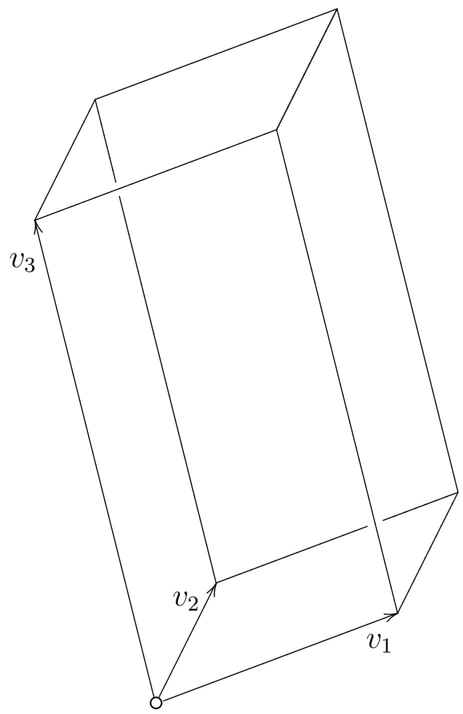


hat die Entartungseigenschaft $D(v_1, v_1) = 0$ und erfüllt die Skalierungseigenschaft $D(\lambda v_1, v_2) = \lambda D(v_1, v_2)$ und $D(v_1, \mu v_2) = \mu D(v_1, v_2)$



Betrachtet man einen '**orientierten**' Flächenbegriff, dann gilt dies für alle $\lambda \in \mathbb{R}$.

Ähnliche Eigenschaften gelten für das orientierte Volumen eines dreidimensionalen Spates $[v_1, v_2, v_3]$ im $\mathbb{R}^3$.



Sei V ein n-dimensionaler K–Vektorraum und K ein beliebiger (!) Körper. Obige Eigenschaften motivieren folgenden Begriff eines verallgemeinerten orientierten Volumens.

Definition: Eine Determinantenfunktion auf V ist eine Funktion $D : V^n \to K$

$$V^n \ni v_1, ..., v_n \mapsto D(v_1, ..., v_n) \in K$$

mit der Eigenschaft

(D1) $D(..., v_i, ...)$ ist linear in v_i bei fester Wahl von $v_1, ..., v_{i-1}, v_{i+1}, ..., v_n$ für alle $i = 1, ..., n$.

$$D(..., \lambda v_i + \mu v_i', ...) = \lambda D(..., v_i, ...) + \mu D(..., v_i', ...)$$

(D2) (Entartungseigenschaft) Für $v_i = v_j$ und $i \neq j$ gilt $D(..., v_i, ..., v_j, ...) = 0$.

Intuitiv: Für $V = \mathbb{R}^3$ sollte $|D(v_1, v_2, v_3)|$ etwas mit dem Volumen des Spates $[v_1, v_2, v_3]$ zu tun haben."

Lemma (Orientierungseigenschaft): Determinantenfunktionen sind alternierend, d.h. es gilt

$$D(v_{\sigma(1)}, ..., v_{\sigma(n)}) = sign(\sigma) \cdot D(v_1, ..., v_n)$$

für alle $\sigma \in S_n$ und alle $v_1, .., v_n \in V$. Eine andere Bezeichnung für alternierend ist schiefsymmetrisch.

Beweis: Da S_n von Transpositionen erzeugt wird, genügt es zu zeigen

$$D(..., v_i, ..., v_j, ...) = -D(..., v_j, ..., v_i, ...).$$

Dazu setze wir zum einfacheren Schreiben $A(v, w) := D(..., v, ..., w, ...)$. $A(v, w)$ ist K–linear in v bei festem w und K–linear in w bei festem v. Es gilt außerdem $A(v, v) = A(w, w) = 0$. Es folgt

$$\begin{aligned} 0 &= A(v + w, v + w) \\ &= A(v, v + w) + A(w, v + w) \\ &= A(v, v) + A(v, w) + A(w, v) + A(w, w) \\ &= A(v, w) + A(w, v) \end{aligned}$$

Daraus folgt sofort $A(v, w) = -A(w, v)$ und damit die Behauptung.

Nun wollen wir die Eigenschaften von Determinantenfunktionen diskutieren.

Zuerst zur Eindeutigkeit:

Sei $v_i = \sum\limits_{j=1}^{n} \lambda_j^i \cdot b_j$, mit den Koordinaten $\lambda_j^i \in K$ (bezüglich einer festen Basis $b_1, ..., b_n$ von V). Dann ist

$$\begin{aligned} D(v_1, ..., v_n) &= D(\sum_{j_1=1}^{n} \lambda_{j_1}^1 \cdot b_{j_1}, ..., \sum_{j_n=1}^{n} \lambda_{j_n}^n \cdot b_{j_n}) \\ &= \sum_{j_1=1}^{n} \lambda_{j_1}^1 \cdot D(b_1, \sum_{j_2=1}^{n} \lambda_{j_2}^2 \cdot b_{j_2}, ..., \sum_{j_n=1}^{n} \lambda_{j_n}^n \cdot b_{j_n}) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{j_1=1}^{n} \lambda_{j_1}^1 \sum_{j_2=1}^{n} \lambda_{j_2}^2 \cdot D(b_{j_1}, b_{j_2}, ..., \sum_{j_n=1}^{n} \lambda_{j_n}^n \cdot b_{j_n}) \\
&= \sum_{1 \le j_1,...,j_n \le n} \lambda_{j_1}^1 \cdots \lambda_{j_n}^n \cdot D(b_{j_1}, ..., b_{j_n}) \\
&= \sum_{\substack{1 \le j_1,...,j_n \le n \\ alle\ paarw.\ versch.}} \lambda_{j_1}^1 \cdots \lambda_{j_n}^n \cdot D(b_{j_1}, ..., b_{j_n}) \\
&= \sum_{\sigma \in S_n} \lambda_{\sigma(1)}^1 \cdots \lambda_{\sigma(n)}^n \cdot D(b_{\sigma(1)}, ..., b_{\sigma(n)}) \\
&\overset{alternierend}{=} \sum_{\sigma \in S_n} \lambda_{\sigma(1)}^1 \cdots \lambda_{\sigma(n)}^n \cdot sign(\sigma) \cdot D(b_1, ..., b_n) \\
&= c(D) \cdot \left(\sum_{\sigma \in S_n} \lambda_{\sigma(1)}^1 \cdots \lambda_{\sigma(n)}^n sign(\sigma) \right)
\end{aligned}
$$

mit $c(D) = D(b_1, ..., b_n) \in K$.

Eine Bemerkung zu dieser Rechnung: Wegen (D2) tragen nur Multiindices $(j_1, .., j_n)$ mit paarweise verschiedenen Einträgen bei. Solche sind gerade von der Gestalt $(j_1, .., j_n) = (\sigma(1), .., \sigma(n))$ für Permutationen $\sigma \in S_n$.

Folgerung: *Je zwei Determinantenfunktionen D_1 und D_2 sind proportional.*

Folgerung: *Sind $b_1, ..., b_n$ linear unabhängig in V und ist D eine nicht identisch verschwindende Determinantenfunktion auf V, dann gilt*

$$D(b_1, ..., b_n) \neq 0.$$

Beweis: Wäre $c(D) = D(b_1, ..., b_n) = 0$, folgt aus obiger Formel

$$D(v_1, ...v_n) = 0$$

für alle Vektoren $v_1, .., v_n$ und somit $D = 0$ im Widerspruch zur Annahme.

Intuitiv besagt letzteres: Sind $v_1, v_2, .., v_n$ linear unabhängig in $\mathbb{R}^n$, dann ist das verallgemeinerte Volumen des n-Spats $[v_1, v_2, ..., v_n]$ nicht null.

<u>Nun zur Existenz von Determinantenfunktionen</u>:

Sei $b_1, ..., b_n$ eine festgewählte Basis von V mit $v_i = \sum_{j=1}^{n} \lambda_j^i \cdot b_j$. Dann definiert

$$D(v_1, ..., v_n) := \sum_{\sigma \in S_n} sign(\sigma) \lambda^1_{\sigma(1)} \cdots \lambda^n_{\sigma(n)} \ .$$

eine Funktion mit den Eigenschaften (D1) und (D2):

Offensichtlich ist D linear als Funktion der Koordinaten $\lambda_1^i, ..., \lambda_n^i$ des Vektors v_i – bei fester Wahl der anderen λ_j^k für $k \neq i$ – denn in jedem Summand kommt nur jeweils ein einziger Term vom Typ λ_*^i (mit Hochindex i) vor. Also ist D linear in den Koordinaten des Vektors v_i und D erfüllt die Eigenschaft (D1).

Andererseits erfüllt D auch Eigenschaft (D2): Sei nämlich $v_i = v_j$ für $i \neq j$. Es bezeichne dann τ die Transposition der Ziffern i, j. Durchläuft σ die geraden Permutationen in $A_n \subset S_n$, dann durchläuft $\sigma \circ \tau$ die ungeraden Permutationen mit negativem Signum. Beachte dazu $sign(\sigma\tau) = sign(\sigma) sign(\tau) = -sign(\sigma)$ wegen $sign(\tau) = -1$. Daraus folgt das Verschwinden des Ausdrucks

$$\begin{aligned} D(v_1, ..., v_n) &= \sum_{\sigma \in S_n} sign(\sigma) \cdot \prod_{k=1}^{n} \lambda^k_{\sigma(k)} \\ &= \sum_{\sigma \in A_n} \left(sign(\sigma) \cdot \prod_{k=1}^{n} \lambda^k_{\sigma(k)} + sign(\sigma\tau) \cdot \prod_{k=1}^{n} \lambda^k_{\sigma\tau(k)} \right) \\ &= \sum_{\sigma \in A_n} sign(\sigma) \cdot \prod_{k \neq i,j} \lambda^k_{\sigma(k)} \cdot \left(\lambda^i_{\sigma(i)} \cdot \lambda^j_{\sigma(j)} - \lambda^i_{\sigma(j)} \cdot \lambda^j_{\sigma(i)} \right) \end{aligned}$$

Benutze die Gleichungen $\lambda^i_{\sigma(i)} = \lambda^j_{\sigma(i)}$ und $\lambda^i_{\sigma(j)} = \lambda^j_{\sigma(j)}$ und das Kommutativgesetz. Beachte, die Annahme $v_i = v_j$ impliziert Koordinatengleichheit $\lambda^i_k = \lambda^j_k$ für alle $k = 1, .., n$.

30 Das Volumen

Definition: Sei $[v_1, ..., v_n]$ der von Vektoren $v_1, ..., v_n \in \mathbb{R}^n$ aufgespannte Spat. Dann definieren wir das Volumen des n-Spats als den Absolutbetrag

$$\boxed{vol[v_1, ..., v_n] = \left| \sum_{\sigma \in S_n} sign(\sigma) \prod_{i=1}^{n} \lambda^i_{\sigma(i)} \right|}$$

Diese Definition erfüllt die Skalierungseigenschaft, die Scherungseigenschaft (siehe Bilder) und die Entartungseigenschaft (D2), wie man dies von der Anschauung für den Volumenbegriff erwarten würde. Dies läßt sich leicht aus (D1) und (D2) ableiten! Zum Beispiel gilt $D(v_1, v_2 + \lambda v_1) = D(v_1, v_2) + \lambda D(v_1, v_1) = D(v_1, v_2)$. Umgekehrt kann man auch zeigen, daß die drei genannten Typen von geometrischen Eigenschaften – allerdings in Form einer orientierten Skalierungseigenschaft wie sie nach Weglassen der Betragsstriche gilt – die Eigenschaften (D1) und (D2) implizieren und somit sogar äquivalent zu den Eigenschaften (D1) und (D2) sind.

Beispiel: Sei $V := \mathbb{R}^2$. Wähle Vektoren $v_1 := \begin{pmatrix} x' \\ y' \end{pmatrix}$ und $v_2 := \begin{pmatrix} x \\ y \end{pmatrix}$ und sei b_1, b_2 die Standardbasis. Dann gilt

$$D(v_1, v_2) = \sum_{\sigma \in S_n} sign(\sigma) \lambda^1_{\sigma(1)} \lambda^2_{\sigma(2)}$$

$$= sign(id) \cdot x' \cdot y + sign((12)) \cdot y' \cdot x = x' \cdot y - y' \cdot x \ .$$

Die Fläche der Raute $[v_1, v_2]$ ist damit gleich

$$|x'y - y'x| \ .$$

31 Der Determinantenhomomorphismus

In diesem Abschnitt wollen wir in vollkommener Analogie zum Vorgehen in §4 den Determinantenhomomorphismus

$$det : \mathcal{G}\ell(V) \to K^*$$

einführen.

Vorbemerkung: Sei V ein endlich dimensionaler K–Vektorraum mit Basis $b_1, ..., b_n$ und einer festen nichttrivialer Determinantenfunktion $D(v_1, ..., v_n)$, zum Beispiel

$$D(v_1, ..., v_n) = \sum_{\sigma \in S_n} sign(\sigma) \lambda^1_{\sigma(1)} \cdots \lambda^n_{\sigma(n)}.$$

Für jeden Endomorphismus $\varphi : V \to V$ ist dann

$$D_\varphi(v_1, ..., v_n) := D\left(\varphi(v_1), ..., \varphi(v_n)\right)$$

wieder eine Determinantenfunktion: Offensichtlich ist D_φ wieder multilinear (D1) und erfüllt (D2). Also ist nach §29 D_φ proportional zu D. Es gibt also eine Konstante $c = c(\varphi)$, so daß für alle $v_1, ..., v_2 \in V$ gilt

$$D_\varphi(v_1, ..., v_n) = c(\varphi) \cdot D(v_1, ..., v_n) \ .$$

Die Konstante $c(\varphi)$ ist unabhängig von der Wahl von D. Ist nämlich $D' = \lambda \cdot D$ und $\lambda \neq 0$, dann ist $D'_\varphi = \lambda \cdot D_\varphi$. Also gilt $c'(\varphi) = c(\varphi)$, da $D \neq 0$.

Die Determinante: Die so definierte Zahl $c(\varphi)$ beschreibt die Veränderung des orientierten Volumens unter der Abbildung φ. Diese Zahl aus K nennt man die Determinante $det(\varphi)$ des Endomorphismus φ. Wählt man $v_1, ..., v_n$ linear unabhängig (beliebig), dann ist nach $D(v_1, .., v_n) \neq 0$ und man erhält

$$\boxed{det(\varphi) = \frac{D(\varphi(v_1), ..., \varphi(v_n))}{D(v_1, ..., v_n)}}$$

für eine beliebige, nichttriviale Determinantenfunktion D.
Beachte: Dies ist insbesondere auch unabhängig von der spezifischen Wahl der v_i, und hängt somit nur von φ ab.

Übungsaufgabe: Vergleiche mit der Definition von $sign(\sigma)$ in §4.

Lemma: *Für $\varphi \in \mathcal{G}\ell(n,K)$ gilt $det(\varphi) \neq 0$.*

Beweis: Sind $v_1,..,v_n$ linear unabhängig, dann auch $\varphi(v_1),..,\varphi(v_n)$. Nach §29 gilt daher $D\left(\varphi(v_1),...,\varphi(v_n)\right) \neq 0$ und somit auch $det(\varphi) \neq 0$.

Lemma: *Es gilt*

$$det(\varphi \circ \psi) = det(\varphi) \cdot \det(\psi)$$

für alle Endomorphismen φ, ψ von V.

Beweis:

$$\begin{aligned} det(\varphi \circ \psi) \cdot D(v_1,...,v_n) &= D\Big((\varphi \circ \psi)(v_1),...,(\varphi \circ \psi)(v_n)\Big) \\ &= D\Big(\varphi(\psi(v_1)),...,\varphi(\psi(v_n))\Big) \\ &= det(\varphi) \cdot D\Big(\psi(v_1),...,\psi(v_n)\Big) \\ &= det(\varphi) \cdot det(\psi) \cdot D(v_1,...,v_n) \end{aligned}$$

Für linear unabhängige $v_1,...,v_n$ gilt $D(v_1,...,v_n) \neq 0$ und daher folgt die Behauptung.

Folgerung: *Die zugehörige Abbildung det : $\mathcal{G}\ell(V) \to K^*$, welcher einem Automorphismus φ seine Determinante zuordnet, ist ein Gruppenhomomorphismus. Der Kern ist die Untergruppe $Sl(V)$ aller Automorphismen φ mit $det(\varphi) = 1$.*

Folgerung: *Es gilt $det(id_V) = 1$ sowie $det(\varphi^{-1}) = det(\varphi)^{-1}$ für alle $\varphi \in Gl(V)$.*

32 Matrixdeterminanten

Im Spezialfall $V = K^n$ ist eine K-lineare Abbildung $\varphi : V \to V$ durch eine Matrix M gegeben. Wir schreiben dann auch $det(M)$ anstelle von $det(\varphi)$. Aus der Definition ergibt sich unmittelbar die Formel

$$\boxed{det(M) = \sum_{\sigma \in S_n} sign(\sigma) \prod_{i=1}^{n} M_{\sigma(i),i}}$$

Dies erhält man, wenn man für $b_1, ..., b_n$ die Standardbasis wählt und die Determinantenfunktion D so, daß gilt $D(b_1, ..., b_n) = 1$. Dann ist nämlich (wenn m_i die Spalten von M bezeichne):

$$det(M) = D(m_1, ..., m_n).$$

Folgerung 1: *Die Determinante einer Matrix ändert beim Vertauschen von 2 Spalten ihr Vorzeichen.*

Beweis: Siehe das Lemma von §29.

Folgerung 2: *Multipliziert man die i-te Spalte einer Matrix M mit einer Zahl $a \in K$, dann ist die Determinante der so erhaltenen Matrix das a-fache der Determinante der ursprünglichen Matrix.*

Folgerung 3: *Addiert man in einer Matrix ein Vielfaches der i-ten Spalte zur j-ten Spalte, so ändert sich die Determinante nicht im Fall $i \neq j$.*

Beweis: Die letzten beiden Folgerungen ergeben sich aus

$$\begin{aligned} D(..., m_i, ..., m_j + \lambda m_i, ...) &= D(..., m_i, ..., m_j, ...) + D(..., m_i, ..., \lambda m_i, ...) \\ &= D(..., m_i, ..., m_j, ...) + \lambda D(..., m_i, ..., m_i, ...) \\ &= D(..., m_i, ..., m_j, ...) + 0 \end{aligned}$$

Satz: *Für die transponierte Matrix gilt $det(M') = det(M)$.*

<u>Beweis</u>:

$$\begin{aligned}
det(M') &= \sum_{\sigma \in S_n} sign(\sigma) \prod_{i=1}^{n} M_{i,\sigma(i)} \\
&\overset{i=\sigma^{-1}(j)}{=} \sum_{\sigma \in S_n} sign(\sigma) \prod_{j=1}^{n} M_{\sigma^{-1}(j),j} \\
&\overset{\tau=\sigma^{-1}}{=} \sum_{\tau \in S_n} sign(\tau^{-1}) \prod_{j=1}^{n} M_{\tau(j),j} \\
&= \sum_{\tau \in S_n} sign(\tau) \prod_{j=1}^{n} M_{\tau(j),j} \\
&= det(M),
\end{aligned}$$

unter Benutzung von $sign(\tau^{-1}) = sign(\tau)^{-1} = sign(\tau)$.

Folgerung 4: *Die Aussagen der obigen Folgerungen gelten auch für Zeilen anstelle von Spalten.*

Dreiecksmatrizen

33 Dreiecksmatrizen

Für eine Reihe von Anwendungen braucht man folgende Identität:

Satz:

$$det \begin{pmatrix} a & * & \cdots & * \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix} = a \cdot det(B)$$

Beweis: Da alle Einträge mit $\sigma(1) \neq 1$ wegen $M_{\sigma(1),1} = 0$ verschwinden gilt

$$det(M) = \sum_{\sigma \in S_n} sign(\sigma) \prod_{i=1}^{n} M_{\sigma(i),i} = \sum_{\substack{\sigma \in S_n \\ \sigma(1)=1}} sign(\sigma) \prod_{i=1}^{n} M_{\sigma(i),i} \; .$$

Die Menge aller Permutationen σ mit $\sigma(1) = 1$ kann mit den Permutationen der $n-1$-elementigen Menge $\{2, ..., n\}$ identifiziert werden. Es folgt

$$det(M) = \sum_{\sigma \in S_{n-1}} sign(\sigma) M_{11} \prod_{i=2}^{n} M_{\sigma(i),i} = M_{11} \cdot det(B) \; .$$

Folgerung: *Für eine obere Dreiecksmatrix M, d.h. $M_{\mu,\nu} = 0$ für $\mu > \nu$, gilt*

$$det(M) = M_{11} \cdot M_{22} \cdot ... \cdot M_{nn} = \prod_{i=1}^{n} M_{ii}.$$

Folgerung:

(i) $det\left(Diag(a_1, ..., a_n)\right) = \prod_{i=1}^{n} a_i$ für beliebige $a_i \in K$.

(ii) $det\left(D_{\nu\mu}(\lambda)\right) = 1$ für alle $\nu \neq \mu$

<u>Beweis</u>: $Diag(a_1, ..., a_n)$ und $D_{\nu\mu}(\lambda)$ für $\mu > \nu$ sind obere Dreiecksmatrizen. Für $D_{\nu\mu}(\lambda)$ mit $\mu < \nu$ reduziert man die Aussage durch Transponieren auf den bekannten Fall.

<u>Folgerung</u>: Für die Determinante einer unteren Dreiecksmatrix M, d.h. $M_{\mu\nu} = 0$ für $\mu < \nu$ gilt

$$det(M) = M_{11} \cdot M_{22} \cdot ... \cdot M_{nn} = \prod_{i=1}^{n} M_{ii}.$$

<u>Beweis</u>: $det(M) = det(M') = \prod_{i=1}^{n} M'_{ii} = \prod_{i=1}^{n} M_{ii}$

34 Vandermonde-Determinante

Die Determinante der Vandermonde-Matrix

$$V(x_1, ..., x_n) \; = \; \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & & & & \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix}$$

ist gerade die Diskriminante $\Delta(x_1, .., x_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$.

Satz: $det(V(x_1, .., x_n)) = \Delta(x_1, .., x_n)$.

Beweis: Im ersten Schritt wollen wir die 1. Zeile von allen Zeilen subtrahieren, dabei ändert sich die Determinante nicht. Man erhält

$$det(V(x_1, .., x_n)) \; = \; det \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 0 & x_2 - x_1 & x_2^2 - x_1^2 & \cdots & x_2^{n-1} - x_1^{n-1} \\ \vdots & & & & \\ 0 & x_n - x_1 & x_n^2 - x_1^2 & \cdots & x_n^{n-1} - x_1^{n-1} \end{pmatrix} .$$

Dies ist wegen der Dreiecksgestalt gleich

$$det \begin{pmatrix} x_2 - x_1 & x_2^2 - x_1^2 & \cdots & x_2^{n-1} - x_1^{n-1} \\ \vdots & & & \\ x_n - x_1 & x_n^2 - x_1^2 & \cdots & x_n^{n-1} - x_1^{n-1} \end{pmatrix} .$$

Nun wendet man Folgerung 4 aus §32 auf jede Zeile an und erhält

$$\prod_{j>1} (x_j - x_1) \cdot det \begin{pmatrix} 1 & x_2 + x_1 & x_1^2 + x_1 x_2 + x_2^2 & \cdots & \sum_{i=0}^{n-2} x_1^{n-2-i} x_2^i \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_k + x_1 & x_1^2 + x_1 x_k + x_k^2 & \cdots & \sum_{i=0}^{n-2} x_1^{n-2-i} x_k^i \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n + x_1 & x_1^2 + x_1 x_n + x_n^2 & \cdots & \sum_{i=0}^{n-2} x_1^{n-2-i} x_n^i \end{pmatrix} .$$

In der verbleibenden Determinante subtrahiert man nun nacheinander des x_1-fache der i-ten Spalte von der $i+1$-ten Spalte beginnend von rechts und man erhält

$$\prod_{j>1}(x_j - x_1) \cdot det \begin{pmatrix} 1 & x_2 & x_2^2 & \cdots & x_2^{n-2} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-2} \\ \vdots & & & & \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-2} \end{pmatrix} .$$

Dies zeigt $det\,(V(x_1, ..., x_n)) = \prod_{j>1}(x_j - x_1) \cdot det\,(V(x_2, ..., x_n))$. Somit folgt per Induktion die Behauptung.

35 Determinantenkriterium für Regularität

Satz: *Für $M \in M_{n,n}(K)$ sind äquivalent:*

(i) $rang(M) = n$

(ii) $det(M) \neq 0$

<u>Beweis</u>: Zu jedem $M \in M_{n,n}(K)$ existieren Matrizen $U, V \in \mathcal{G}\ell(n, K)$ mit

$$UMV = \begin{pmatrix} E^{r,r} & 0 \\ 0 & 0 \end{pmatrix}$$

mit $r = rang(M)$ nach §25. Rechts steht eine obere Dreiecksmatrix. Aus §33 folgt daher

$$det \begin{pmatrix} E^{r,r} & 0 \\ 0 & 0 \end{pmatrix} = \begin{cases} 1 & r = n \\ 0 & r < n \end{cases}.$$

Da U, V invertierbar sind, folgt aus §31 sofort $det(U) \neq 0$ und $det(V) \neq 0$. Wegen

$$det(U) \cdot det(M) \cdot det(V) = det \begin{pmatrix} E^{r,r} & 0 \\ 0 & 0 \end{pmatrix}$$

folgt also

$$det(M) = 0 \iff det \begin{pmatrix} E^{r,r} & 0 \\ 0 & 0 \end{pmatrix} \iff r = rang(M) < n.$$

Zusammenfassung: *Für $M \in M_{n,n}(K)$ sind äquivalent:*

(i) $rang(M) = n$

(ii) $det(M) \neq 0$

(iii) $M \in \mathcal{G}\ell(n, K)$, *d.h. M ist invertierbar*

(iv) *M ist bijektiv*

(v) *M ist injektiv*

(vi) *M ist surjektiv*

(vii) *Das lineare Gleichungssystem $M \cdot x = y$ mit $x, y \in K^n$ besitzt für jede Wahl von y genau eine Lösung x.*

36 Laplacescher Entwicklungssatz

Seien $m_1, ..., m_n$ die n Spaltenvektoren einer Matrix $M \in M_{n,n}(K)$. Wir betrachten die Determinante einer Matrix als Funktion der i-ten Spalte ($1 \leq i \leq n$ fest). Sei x ein Spaltenvektor im K^n. Die Funktion

$$L(x) := det(m_1 \ ... \ m_{i-1} \ x \ m_{i+1} \ ... \ m_n)$$

ist linear in der Vektorvariable x. Daher gilt

$$L(x) = \sum_{j=1}^{n} L(e_j) x_j,$$

wobei

$$x = \sum_{j=1}^{n} x_j e_j \qquad\qquad (e_1, ..., e_n \ Standardbasis)$$

und

$$\begin{aligned}
L(e_j) &= det\begin{pmatrix} M_{1,1} & \cdots & M_{1,i-1} & 0 & M_{1,i+1} & \cdots & M_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ M_{j,1} & \cdots & M_{j,i-1} & 1 & M_{j,i+1} & \cdots & M_{j,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ M_{n,1} & \cdots & M_{n,i-1} & 0 & M_{n,i+1} & \cdots & M_{n,n} \end{pmatrix} \\
&= (-1)^{i-1} det\begin{pmatrix} 0 & M_{1,1} & \cdots & M_{1,i-1} & M_{1,i+1} & \cdots & M_{1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & M_{j,1} & \cdots & M_{j,i-1} & M_{j,i+1} & \cdots & M_{j,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & M_{n,1} & \cdots & M_{n,i-1} & M_{n,i+1} & \cdots & M_{n,n} \end{pmatrix} \\
&= (-1)^{i-1}(-1)^{j-1} det\begin{pmatrix} 1 & * & \cdots & * & * & \cdots & * \\ 0 & M_{1,1} & \cdots & M_{1,i-1} & M_{1,i+1} & \cdots & M_{1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & M_{j-1,1} & \cdots & M_{j-1,i-1} & M_{j-1,i+1} & \cdots & M_{j-1,n} \\ 0 & M_{j+1,1} & \cdots & M_{j+1,i-1} & M_{j+1,i+1} & \cdots & M_{j+1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & M_{n,1} & \cdots & M_{n,i-1} & M_{n,i+1} & \cdots & M_{n,n} \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
&= (-1)^{i+j} det \begin{pmatrix} M_{1,1} & \cdots & M_{1,i-1} & M_{1,i+1} & \cdots & M_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ M_{j-1,1} & \cdots & M_{j-1,i-1} & M_{j-1,i+1} & \cdots & M_{j-1,n} \\ M_{j+1,1} & \cdots & M_{j+1,i-1} & M_{j+1,i+1} & \cdots & M_{j+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ M_{n,1} & \cdots & M_{n,i-1} & M_{n,i+1} & \cdots & M_{n,n} \end{pmatrix} \\
&= (-1)^{i+j} \cdot det \begin{pmatrix} \text{Matrix, welche aus } M \text{ durch} \\ \text{Streichen der } i\text{-ten Spalte und} \\ \text{der } j\text{-ten Zeile entsteht} \end{pmatrix} \\
&:= \tilde{M}_{ij}
\end{aligned}$$

Setzt man für x jetzt $x = m_i$ beziehungsweise $x = m_k$ $(i \neq k)$ ein, erhält man

$$L(m_i) = det(m_1, ..., m_i, ..., m_n) = det(M)$$

beziehungsweise

$$L(m_k) = det(m_1, ..., m_k, ...m_{i-1}, m_k, m_{i+1}, ..., m_n) = 0,$$

da der Rang der Matrix $(m_1, ..., m_k, ..., m_k, ..., m_n)$ echt kleiner als n ist. Beachtet man noch $x_j = M_{ji}$ bzw. $x_j = M_{jk}$ für $x = m_i$ bzw. $x = m_k$, folgt

$$\left.\begin{aligned}
L(x) &= \sum_{j=1}^{n} L(e_j)x_j = \sum_{j=1}^{n} \tilde{M}_{ij}x_j \\
L(m_i) &= \sum_{j=1}^{n} \tilde{M}_{ij}M_{ji} = det(M) \\
L(m_k) &= \sum_{j=1}^{n} \tilde{M}_{ij}M_{jk} = 0
\end{aligned}\right\} (*)$$

Definition: Die <u>Komplementärmatrix</u> $\tilde{M}$ von M ist definiert als Matrix mit den Einträgen

$$\tilde{M}_{ij} = (-1)^{i+j} \cdot det \begin{pmatrix} \text{Matrix, welche aus } M \text{ durch} \\ \text{Streichen der } i\text{-ten Spalte und} \\ \text{der } j\text{-ten Zeile entsteht} \end{pmatrix}$$

Die Gleichungen $(*)$ schreiben sich dann in der Form

Satz: *Für die Komplementärmatrix $\tilde{M}$ von M gilt:*

$$\boxed{\tilde{M} \cdot M = det(M) \cdot E}$$

Korollar: *Im Fall $det(M) \neq 0$ gilt*

$$\boxed{M^{-1} = det(M)^{-1} \cdot \tilde{M}.}$$

Einen Spezialfall der obigen Formel $det(M) \cdot E = \tilde{M} \cdot M$ nennt man

<u>Laplacescher Entwicklungssatz</u>:

$$det(M) = \sum_{j=1}^{n} (-1)^{i+j} det \begin{pmatrix} \text{Matrix, welche aus } M \text{ durch} \\ \text{Streichen der } i\text{-ten Spalte und} \\ \text{der } j\text{-ten Zeile entsteht} \end{pmatrix} \cdot M_{ji}$$

Man nennt dies auch die Entwicklung nach der i-ten Spalte der Matrix. Es handelt sich dabei um den Spezialfall $det(M) = \sum_{j=1}^{n} \tilde{M}_{ij} M_{ji}$ der Formeln $(*)$.

37 Cramersche Regel

Gegeben sei ein $n \times n$ lineares Gleichungssystem

$$M \cdot x = b$$

mit einer invertierbaren Matrix M. Das heißt es gilt $det(M) \neq 0$. Gesucht ist eine Formel für die Komponenten x_i des Lösungsvektors x.

Satz: *Es gilt*

$$x_i = \frac{det \begin{pmatrix} m_{1,1} & \cdots & m_{1,i-1} & b_1 & m_{1,i+1} & \ldots & m_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ m_{n,1} & \cdots & m_{n,i-1} & b_n & m_{n,i+1} & \ldots & m_{n,n} \end{pmatrix}}{det(M)}.$$

In der oberen Matrix wurde der i-te Spaltenvektor durch den Vektor b ersetzt.

Beweis: Da M nach Annahme invertierbar ist, gilt für $x = M^{-1} \cdot b$. Da $det(M) \neq 0$ ist, gilt nach §36

$$x_i = (M^{-1} \cdot b)_i \ = \ det(M)^{-1} \sum_{j=1}^{n} \tilde{M}_{ij} b_j \ .$$

Durch Entwicklung nach der i-ten Spalte (Laplacescher Entwicklungssatz) stimmt dies überein mit

$$det(M)^{-1} \cdot det \begin{pmatrix} m_{1,1} & \cdots & m_{1,i-1} & b_1 & m_{1,i+1} & \ldots & m_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ m_{n,1} & \cdots & m_{n,i-1} & b_n & m_{n,i+1} & \ldots & m_{n,n} \end{pmatrix} \ .$$

38 Das charakteristische Polynom

Sei V ein endlich dimensionaler K-Vektorraum und es sei $\varphi : V \to V$ ein Endomorphismus von V. Dann definiert

$$\chi_\varphi(x) \;=\; det(x \cdot id - \varphi)$$

eine Funktion der Variable $x \in K$, das charakteristische Polynom.

In der Tat ist diese Funktion **polynomial** vom Grad $n = dim_K(V)$.

Beweis: Durch Wahl einer Basis gilt obdA $V = K^n$ und der Endomorphismus φ wird durch eine Matrix $M \in M_{n,n}(K)$ gegeben. Somit gilt

$$\chi_\varphi(x) \;=\; \chi_M(x) \;=\; det(x \cdot E - M)\;.$$

Die rechte Seite ist eine Summe mit $\#S_n$ Summanden, welche jeweils Produkte von n Koeffizienten der Matrix $x \cdot E - M$ sind. Jeder Faktor ist vom Grad 0 oder 1 (konstant außerhalb der Diagonale). Somit ist $\chi_M(x)$ polynomial in x vom Grad $\leq n$.

Bemerkung : Der Term vom höchsten Grad n des charakteristischen Polynoms kommt vom Summand $(x - M_{11}) \cdot \ldots \cdot (x - M_{nn})$, welcher zur trivialen Permutation $\sigma \in S_n$ gehört. Es folgt

$$\chi_M(x) \;=\; x^n \;+\; \; niedere\ x - Potenzen\;.$$

Bemerkung: Da das charakteristische Polynom von φ und nicht von der Wahl einer Basis abhängt, folgt sofort $\chi_M(x) = \chi_{TMT^{-1}}(x)$. Hierbei ist T die Basiswechselmatrix. Wir geben für diese fundamentale Eigenschaft des charakteristischen Polynoms im nächsten Lemma einen zweiten Beweis.

Definition: Zwei Matrizen $M, N \in M_{n,n}(K)$ heißen konjugiert, wenn eine invertierbare Matrix $T \in \mathcal{G}\ell(n,K)$ existiert mit

$$N \;=\; T \cdot M \cdot T^{-1}.$$

In diesem Fall schreiben wir $N \overset{k}{\sim} M$. Analog heißen zwei Endomorphismen $\varphi, \psi \in End(V)$ konjugiert, wenn es einen Automorphismus $\tau \in Gl(V)$ gibt mit

$$\psi \;=\; \tau \circ \varphi \circ \tau^{-1}\;.$$

Bemerkung: Offensichtlich definiert dies eine Äquivalenzrelation auf den Matrizen in $M_{n.n}(K)$ (resp. analog den Endomorphismen in $End(V)$):

(i) $N \overset{k}{\sim} N$

(ii) $N \overset{k}{\sim} M \quad \Rightarrow \quad M \overset{k}{\sim} N$

(iii) $N \overset{k}{\sim} M, \quad M \overset{k}{\sim} L \quad \Rightarrow \quad N \overset{k}{\sim} L$

Lemma: *Konjugierte Endomorphismen haben das gleiche charakteristische Polynom.*

Beweis:

$$\begin{aligned}
\chi_{\tau\circ\varphi\circ\tau^{-1}}(x) &= det(x \cdot id - \tau \circ \varphi \circ \tau^{-1}) \\
&= det(\tau \circ x \cdot id \circ \tau^{-1} - \tau \circ \varphi \circ \tau^{-1}) \\
&= det(\tau \circ (x \cdot id - \varphi) \circ \tau^{-1}) \\
&= det(\tau) \cdot det(x \cdot id - \varphi) \cdot det(\tau^{-1}) \\
&= det(x \cdot id - \varphi) \\
&= \chi_\varphi(x)
\end{aligned}$$

Beispiel: Wir berechnen nun das charakteristische Polynom einer oberen Dreiecksmatrix.

$$\chi_{\begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}}(x) = det(\begin{pmatrix} x - a_{11} & & * \\ & \ddots & \\ 0 & & x - a_{nn} \end{pmatrix}) \overset{§33}{=} \prod_{i=1}^{n}(x - a_{ii})$$

Die Nullstellen des charakteristischen Polynoms sind also in diesem Fall genau die Diagonaleneinträge der Dreiecksmatrix. Insbesondere zerfällt das charakteristische Polynom bereitsüber dem Körper K vollständig in Linearfaktoren.

Definition: Eine Matrix $M \in M_{n,n}(K)$ heißt trigonalisierbar bzw. diagonalisierbar, falls sie konjugiert ist zu einer oberen Dreiecksmatrix bzw. Diagonalmatrix.

Folgerung: *Ist M trigonalisierbar (oder diagonalisierbar), dann zerfällt $\chi_M(X)$ über dem Körper K in ein Produkt von Linearfaktoren.*

Beweis: Aus

$$M \overset{k}{\sim} \begin{pmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{pmatrix}$$

folgt wegen der Konjugationsinvarianz des charakteristischen Polynoms

$$\chi_M(X) \;=\; \chi_{\begin{pmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{pmatrix}}(X) \;=\; \prod_{i=1}^{n}(X - a_{ii}) \; .$$

Die Umkehrung dieser Aussage beweisen wir im nächsten Kapitel mit Hilfe der Theorie der Eigenvektoren.

Appendix

Hinweis: Um nicht erklären zu müssen, wie man mit Matrizen rechnet, deren Einträge Unbestimmte sind, haben wir das charakteristische Polynom $\chi_\varphi(x)$ als eine polynomiale Funktion der Variable $x \in K$ mit Werten in K aufgefasst

$$K \ni x \;\mapsto\; \chi_\varphi(x) \in K \; .$$

Im allgemeinen ist ein formales Polynom aber nicht durch seine Funktionswerte eindeutig bestimmt. Dieses Problem tritt aber nur für endliche Körper K auf. Beispiel: $f(X) = X^3 + X$ ist als Funktion im Körper $K = \{0, 1\}$ mit zwei Elementen identisch Null, obwohl es als formales Polynom nicht verschwindende Koeffizienten in K besitzt.

Lemma: *Sei $\#K = \infty$, dann gilt: Stimmen zwei formale Polynome $P(X)$ und $Q(X)$ für alle Einsetzungen $x \in K$ überein, dann sind die Koeffizienten der beiden Polynome $P(X)$ und $Q(X)$ gleich.*

Beweis : Sei die Differenz $f(X) = P(X) - Q(X)$ ein Polynom vom Grad $< n$

$$f(X) = \sum_{i=1}^{n} a_i \cdot X^{i-1} = a_1 + a_2 \cdot X + ...a_n \cdot X^{n-1}$$

(für gewisse $a_i \in K$.) Nach Annahme gilt dann $f(x) = 0$ für alle $x \in K$. Wähle n paarweise verschiedene Elemente $x_1, ..., x_n$ in K. Wir behaupten

$$f(x_1) = ... = f(x_n) = 0 \quad \Rightarrow \quad f(X) = 0 \ .$$

Die linke Seite definiert nämlich folgendes lineares Gleichungssystem für die Koeffizienten a_i des Polynoms $f(X)$

$$V(x_1, ..., x_n) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} f(x_1 \\ \vdots \\ f(x_n) \end{pmatrix} = 0.$$

Die Vandermonde-Matrix $V(x_1, ..., x_n)$ ist nach §33) invertierbar. Daraus folgt $a_1 = \cdots = a_n = 0$. Man erhält daher $f(X) = 0$ beziehungsweise wie behauptet $P(X) = Q(X)$ (koeffizientenweise Gleichheit).

<u>Folgerung</u>: Falls der Körper K unendliche viele Elemente enthält, gibt es keinen Unterschied zwischen formalen Polynomen und polynomialen Funktionen. Somit sind die Koeffizienten des charakteristischen Polynoms in diesem Fall durch die Funktionswerte $\chi_\varphi(x), x \in K$ eindeutig bestimmt.

39 Eigenwerte und Eigenvektoren

Definition: Ein Vektor $v \neq 0$ heißt Eigenvektor eines Endomorphismus $\varphi : V \to V$, falls eine Zahl $\lambda \in K$ existiert, so daß gilt:

$$\varphi(v) = \lambda \cdot v \qquad (Eigenvektorgleichung)\ .$$

Die Zahl $\lambda \in K$ ist dann eindeutig bestimmt und heißt Eigenwert von φ zum Eigenvektor v.

Beweis der Eindeutigkeit: Aus $\varphi(v) = \lambda_1 \cdot v = \lambda_2 \cdot v$ folgt $(\lambda_1 - \lambda_2) \cdot v = 0$. Daraus folgt $\lambda_1 = \lambda_2$ wegen $v \neq 0$.

Bezeichnung: $V(\lambda) := \{v \in V \mid \varphi(v) = \lambda \cdot v\}$ nennt man den Eigenraum des Endomorphismus φ zum Eigenwert λ.

Aus der Definition folgt, daß $V(\lambda)$ der Kern der linearen Abbildung $\varphi - \lambda \cdot id$ ist. Um die Eigenvektoren zu bestimmen, muß man also zuerst alle möglichen Eigenwerte bestimmen und dann die Kerne der Abbildungen $\varphi - \lambda \cdot id$ berechnen. Die Eigenwerte sind aber gereade die Nullstellen des charakteristischen Polynoms $\chi_\varphi(x)$, wie das nächste Lemma zeigt.

Lemma:

1. *$V(\lambda) = Kern(\lambda \cdot id - \varphi)$ Insbesondere ist $V(\lambda)$ ein Unterverktorraum von V.*

2. *$V(\lambda) \neq \{0\}$ genau dann, wenn λ eine Nullstelle des charakteristischen Polynoms $\chi_\varphi(x)$ ist, d.h. $\chi_\varphi(\lambda) = 0$.*

3. *Ein Vektor $v \neq 0$ in V ist ein Eigenvektor von φ genau dann, wenn v in einem der Räume $V(\lambda)$ liegt.*

Beweis: 3. ist offensichtlich eine Umschreibung der Definition.

1. $v \in V(\lambda) \iff \varphi(v) = \lambda v \iff (\lambda id - \varphi)(v) = 0 \iff v \in Kern(\lambda id - \varphi)$

2. $V(\lambda) \neq \{0\} \iff Kern(\lambda id - \varphi) \neq \{0\} \iff \lambda id - \varphi$ nicht bijektiv $\iff \chi_\varphi(\lambda) = det(\lambda id - \varphi) = 0$. Siehe §35.

Folgerung: *Es gibt höchstens $dim_K(V)$ verschiedene Eigenwerte λ von $\varphi \in End(V)$.*

Beweis: Der Vandermondeschluß im Appendix des vorigen Kapitels §38 zeigt: Jedes Polynom vom Grad $\leq n$ hat höchstens n paarweise verschiedene Nullstellen in einem Körper K. Insbesondere hat also $\chi_\varphi(X)$ höchstens $dim_K(V)$ verschiedene Nullstellen (= Eigenwerte).

Gegeben sie eine $n \times n$ Matrix mit Einträgen in K. Weiterhin sei $\#K = \infty$. Gesucht ist eine Basis B, bezüglich der die Matrix M Dreiecksgestalt oder sogar Diagonalengestalt besitzt. Zur Erinnerung: In diesem Fall hieß M trigonalisierbar bzw. diagonalisierbar.

Satz: *Eine Matrix $M \in M_{n,n}(K)$ ist genau dann trigonalisierbar, wenn ihr charakteristisches Polynom $\chi_M(x)$ sich als Produkt von Linearfaktoren*

$$\chi_M(x) = \prod_{i=1}^{n}(x - \lambda_i) \qquad , \qquad \lambda_i \in K$$

schreiben läßt.

Beweis: Gegeben sei eine Matrix $M \in M_{n,n}(K)$ und $\lambda_1, ..., \lambda_n \in K$, so daß $\chi_M(K) = \prod_i (X - \lambda_i)$. Wähle ein λ_i und einen dazugehörigen Eigenvektor v_i. Wähle eine Basis $b_1, ..., b_n$ so, daß $b_1 = v_i$. Es folgt für diese Basis

$$M_{B,B}(M) = \begin{pmatrix} \lambda_i & * & \cdots & * \\ 0 & & & \\ \vdots & & M_{n-1} & \\ 0 & & & \end{pmatrix},$$

da $M_{b_1} = \lambda_i v_i + 0b_2 + \cdots + 0b_n$. Die Dreiecksungleichung impliziert

$$\chi_M(X) = (X - \lambda_i) \cdot \chi_{M_{n-1}}(X).$$

Aus der Annahme an $\chi_M(X)$ folgt somit $\chi_{M_{n-1}}(X) = \prod\limits_{j \neq i}(X - \lambda_j)$, da beide für unendlich viele Einsetzungen $X \neq \lambda_i$ aus K denselben Wert haben. Somit ist per Induktion M_{n-1} bereits trigonalisierbar. D.h. ändert man $b_2, ..., b_n$ zu einer geeigneten Basis $b'_2, ..., b'_n$ ab, dann erhält M_{n-1} Dreiecksgestalt. Bezüglich der neuen Basis $v_i, b'_2, ..., b'_n$ hat M also Dreiecksgestalt:

$$M_{B,B}(M) = \begin{pmatrix} \lambda_i & & & * \\ & \lambda_? & & \\ & & \ddots & \\ 0 & & & \lambda_? \end{pmatrix}.$$

Für eine geeignete Basiswechselmatrix T (vgl. §21) folgt somit

$$TMT^{-1} = \begin{pmatrix} \lambda_i & & & * \\ & \lambda_? & & \\ & & \ddots & \\ 0 & & & \lambda_? \end{pmatrix}.$$

q.e.d.

Wie man sieht, kann man die Eigenwerte in beliebig vorgegebener Reihenfolge in die Diagonale bringen.

Satz: *Für $K = \mathbb{C}$ ist jede Matrix $M \in M_{n,n}(\mathbb{C})$ trigonalisierbar.*

<u>Beweis</u>: Da nach dem Fundamentalsatz der Algebra jedes Polynom mit Koeffizienten in $\mathbb{C}$ ein Produkt von Linearfaktoren (Mit Koeff. in $\mathbb{C}$) ist, folgt dies aus dem letzten Satz.

<u>Achtung</u>: Dies ist für andere Körper nicht immer möglich: Die Matrix

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

ist für $0 < \theta < 180°$ und $K = \mathbb{R}$ nicht trigonalisierbar. Ihre Eigenwerte sind $\lambda_{1/2} = cos(\theta) \pm i \sin(\theta)$ und liegen daher nicht in $\mathbb{R}$:

Über $K = \mathbb{C}$ ist die obige Matrix dagegen sogar diagonalisierbar! Sei e_1, e_2 die Standardbasis von $\mathbb{C}^2$. Definiere $b_1 := e_1 + ie_2$ und $b_2 := e_1 - ie_2$. Dann gilt

$$\begin{aligned}
M(b_1) &= (\cos\theta e_1 - \sin\theta e_2) + i(\sin\theta e_1 + \cos\theta e_2) \\
&= (e_1 + ie_2)(\cos\theta + i\sin\theta) \\
&= \lambda_1 \cdot b_1 \\
M(b_2) &= (\cos\theta e_1 - \sin\theta e_2) - i(\sin\theta e_1 + \cos\theta e_2) \\
&= (e_1 - ie_2)(\cos\theta + i\sin\theta) \\
&= \lambda_2 \cdot b_2
\end{aligned}$$

Bezüglich der Basis b_1, b_2 erhält also $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ die Gestalt

$$\begin{pmatrix} \cos\theta + i\sin\theta & 0 \\ 0 & \cos\theta - i\sin\theta \end{pmatrix}.$$

40 Diagonalisierbarkeit

Satz: *Sei $M \in M_{n,n}(K)$ und seien $\lambda_1, ..., \lambda_n \in K$ paarweise verschiedene Eigenwerte von M. Dann gilt:*

(1) $\chi_M(X) = \prod_{i=1}^{n}(X - \lambda_i)$

(2) M ist diagonalisierbar.

Hilfssatz: *Seien $\lambda_1, ..., \lambda_n \in K$ paarweise verschiedene Eigenwerte eines Endomorphismus M und $v_1, ..., v_n$ zugehörige Eigenvektoren, dann sind $v_1, ..., v_n$ linear unabhängig.*

Beweis: Gegeben sei eine K-lineare Relation

$$(*) \qquad c_1 v_1 + ... + c_n v_n \ = \ 0$$

der Länge N (= Anzahl der Koeffizienten $c_i \neq 0$). Wir müssen zeigen $N = 0$ und geben einen Widerspruchsbeweis.

Sei $N \geq 1$. Dann kann die gegebene Relation der Länge N zu einer Relation der Länge $N - 1$ verkürzt werden. Iteriert liefert dies eine Relation der Länge 1 (von der Form $0 + c_j v_j + 0 = 0$ mit $c_j \neq 0$). Dies impliziert $v_j = 0$ im Widerspruch zur Annahme $v_j \neq 0$ (Eigenvektorbedingung).

Die Längenreduktion: M angewendet auf die Gleichung $(*)$ liefert wegen $M(v_i) = \lambda_i v_i$ die Gleichung

$$(A) \qquad \lambda_1 c_1 v_1 + ... + \lambda_n c_n v_n \ = \ 0 \ .$$

Sei $N \geq 1$ und obdA $c_1 \neq 0$. Multiplizieren wir $(*)$ mit λ_1

$$(B) \qquad \lambda_1 c_1 v_1 + ... + \lambda_1 c_n v_n \ = \ 0 \ ,$$

dann liefert die Differenz der Gleichungen $(A) - (B)$ die neue Relation

$$(**) \qquad 0 + (\lambda_2 - \lambda_1) c_2 v_2 + ... + (\lambda_n - \lambda_1) c_n v_n \ = \ 0 \ .$$

Beachte: Da $\lambda_1, ..., \lambda_n$ paarweise voneinander verschieden sind, sind alle Koeffizienten $(\lambda_i - \lambda_1) \neq 0$. Somit ist $(**)$ eine Relation der Länge $N-1$.

Beweis des Satzes: Seien $v_1, ..., v_n$ zugehörige Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_n$. Die Vektoren $v_1, ..., v_n$ sind n linear unabhängige Vektoren in K^n wegen des Hilfssatzes. Also ist $v_1, ..., v_n$ eine Basis B des K^n. Bezüglich dieser Basis gilt

$$M_{B,B}(M) = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

Die Matrix ist somit diagonalisierbar. Außerdem gilt trivialerweise

$$\chi_M(X) = \prod_{i=1}^{n} (X - \lambda_i).$$

Folgerung: *Seien $M, N \in M_{n,n}(K)$ und seien $\lambda_1, ..., \lambda_n$ paarweise verschiedene Eigenwerte sowohl von M als auch von N. Dann sind M und N konjugiert.*

Beweis: Aus obigem Satz folgt die Existenz einer invertierbaren Matrix $T \in \mathcal{G}\ell(n, K)$ mit $T \cdot M \cdot T^{-1} = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$ und analog die Existenz einer invertierbaren Matrix $S \in \mathcal{G}\ell(n, K)$ mit $S \cdot M \cdot S^{-1} = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Es folgt $M = RNR^{-1}$ mit $R = T^{-1}S$.

Bilinearformen und quadratische Formen

41 Bilinearformen

Seien V, W zwei K–Vektorräume (möglicherweise auch unendlich dimensional). Eine Funktionen

$$\begin{array}{rccc} B: & V \times W & \to & K \\ & (v,w) & \mapsto & B(v,w) \end{array}$$

heißt **Bilinearform**, falls B in der ersten Variable K–linear ist bei fester zweiter Variable und umgekehrt. Das heißt wir fordern

$$\begin{array}{lll} (1) & B(\lambda v + \mu v', w) & = \lambda B(v,w) + \mu B(v',w) \\ (2) & B(v, \lambda w + \mu w') & = \lambda B(v,w) + \mu B(v,w') \end{array} .$$

für alle $\lambda, \mu \in K$ und alle $v, v' \in V, w, w' \in W$. Sind $V' \subset V$ und $W' \subset W$ K-Untervektorräume, dann ist die Einschränkung einer Bilinearform auf $V \times W$ eine Bilinearform auf $V' \times W'$.

Beispiele:

(1) Die Determinante $\quad B(v,w) = det\left(\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\right) = v_1w_2 - v_2w_1$ ist bilinear auf $V = W = K^2$.

(2) $V = W = K^n$, das Skalarprodukt $B(x,y) = \sum\limits_{i=1}^{n} x_i y_i$

(3) $W = V^\star, \quad B(v, v^\star) = v^\star(v) =< v^\star, v >$ (die kanonische Paarung)

(4) Ist $B(v,w)$ eine Bilinearform $B : V \times W$, dann ist auch $\tilde{B}(w,v) := B(v,w)$ eine Bilinearform $\tilde{B} : W \times V \to K$.

(5) $W = V = \mathbb{R}$-Vektorraum aller stetigen Funktionen auf $[a,b]$. $B(f,g) := \int\limits_a^b f(t)g(t)dt \in \mathbb{R}$ definiert eine Bilinearform.

(6) $V = W = M_{n,n}(K)$. Dann definiert $B(X,Y) = Spur(X \cdot Y)$ eine Bilinearform.

Definition: Eine Bilinearform heißt nichtausgeartet, wenn gilt:

$$\begin{array}{lll} (I) & \forall v \in V & : B(v,w) = 0 \quad \Rightarrow w = 0 \\ (II) & \forall w \in W & : B(v,w) = 0 \quad \Rightarrow v = 0 \end{array}$$

Bemerkung: $B(0,w) = 0 = B(v,0)$ ist wegen der Linearität immer erfüllt.

Beispiel: B ist nichtausgeartet in den Beispielen (1),(2),(3),(5),(6).

Definition: Im Fall $v = W$ heißt eine Bilinearform symmetrisch, wenn gilt

$$B(v,w) = B(w,v)$$

für alle $v \in V$ und $w \in W$. Gilt $B(v,w) = -B(w,v)$ heißt die Form alternierend.

Beispiel: B ist symmetrisch in den Beispielen (2),(5) und (6). Im Fall (1) genau dann wenn $2 = 0$ gilt in K.

42 Matrixbeschreibung

Seien nun V und W endlich dimensionale K-Vektorräume. Seien $b_1, ..., b_n$ und $b'_1, ..., b'_m$ Basen von V respektive W. Für eine Bilinearform

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

auf den K–Vektorräumen V und W gilt

$$B(x, y) = B(\sum_{i=1}^{n} x_i b_i, \sum_{j=1}^{m} y_j b'_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i B(b_i, b'_j) y_j \ .$$

Die dadurch festgelegte Matrix

$$S = \big(B(b_i, b'_j)\big)_{\substack{i=1,...,n \\ j=1,...,m}}$$

ist durch B eindeutig bestimmt. Wir erhalten $B(v, w) =< v, w >_S$ mit

$$\boxed{< x, y >_S = \ x' \cdot S \cdot y \ = \ (x_1, ..., x_n) \cdot S \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}} \ .$$

Umgekehrt liefert diese Formel für jedes $S \in M_{n,m}(K)$ eine Bilinearform $<,>_S$ auf $V \times W$. Es folgt

Lemma: *Die obige Zuordnung stiftet eine Bijektion zwischen den Bilinearformen auf $V \times W$ und den Matrizen S in $M_{n,m}(K)$. Zwei Bilinearformen sind gleich dann und nur dann, wenn die zugeordneten Matrizen S übereinstimmen.*

Zusatz: *Eine Form ist nichtausgeartet genau dann wenn $n = m$ gilt und die zugehörige Matrix S invertierbar ist.*

Beweis des Zusatzes: Nichtausgeartet bedeutet $x'S = 0 \Rightarrow x = 0$ und $Sy = 0 \Rightarrow y = 0$, oder wegen $S'x = (x'S)'$ dazu äquivalent: S' und S sind injektiv. Die impliziert einerseits $n \leq m$ und $m \leq n$, also $n = m$ und dann wegen der Dimensionsformel auch die Surjektivität – d.h dann Bijektivität – von S. Ist umgekehrt S bijektiv, dann gilt $det(S) = det(S') \neq 0$, also sind S und S' bijektiv und insbesondere injektiv.

Satz: *Sei $< \, , \, > : V \times V \to K$ eine Bilinearform und $b_1, ..., b_n$ eine Basis von V. Dann ist*

$$< y, x >_S = < x, y >_{S'}$$

und $< \, , \, >_S$ genau dann symmetrisch, wenn die Matrix $S = (< b_1, b_j >)$ eine symmetrische Matrix $S = S'$ ist.

Beweis: $< y, x >_S = y' \cdot S \cdot x = (y' \cdot S \cdot x)' = x' \cdot S' \cdot y = < x, y >_{S'}$.

Analog entsprechen alternierende Bilinearformen antisymmetrischen Matrizen.

Basiswechsel: Bei fester Wahl der Basen B und B' kann man also auf diese Weise die Bilinearformen auf $V \times W$ mit den $n \times m$ Matrizen aus $M_{n,m}(K)$ identifizieren. Im Fall $V = K^n$ und $W = K^m$ hat man eine kanonische Basiswahl, die Standardbasen. Im allgemeinen muß man willkürliche Basen wählen. Es ist daher von Interesse, wie die Matrix S von der Wahl der Basen abhängt. Der Einfachheit halber beschränken wir uns auf den Fall $V = W$ und $B = B'$. Sei also $V = W$ und $b_1, ..., b_n$ eine Basis von V. Die $< \, , \, >$ zugehörige Matrix sei

$$S = (< b_1, b_j >) \, .$$

Sei $\widetilde{b_1}, ..., \widetilde{b_n}$ eine andere Basis von V mit zugehöriger Matrix

$$\tilde{S} = \left(< \widetilde{b_1}, \widetilde{b_j} >\right) .$$

Der (inverse) Basiswechsel sei gegeben durch $\tilde{b_i} = \sum\limits_{j=1}^{n} T_{ij} b_j$.

Lemma: *Für einen Basiswechsel $T = (T_{ij}) \in \mathcal{G}\ell(n, \mathbb{R})$ wie oben gilt*

$$\boxed{\tilde{S} = T \cdot S \cdot T'}$$

Beweis: Aus $< \widetilde{b_i}, \widetilde{b_j} > = < \sum_{k=1}^n T_{ik} b_k, \sum_{l=1}^n T_{jl} b_l > = \sum_{k,l} T_{ik} < b_k, b_l > T'_{lj}$ folgt wegen $T_{jl} = T'_{lj}$ unmittelbar die gesuchte Matrixgleichung $\tilde{S} = T \cdot S \cdot T'$.

Achtung: *'Bilinearformen' transformieren sich also anders als 'Homomorphismen' bei Basiswechsel.*

43 Quadratische Formen

In diesen Paragraph sei K ein Körper mit $2 \neq 0$ und V ein K–Vektorraum.

Jeder symmetrischen Bilinearform $< , > : V \times V \to K$ wird durch Einschränken auf die Diagonale die folgende quadratische Form $q(v) =< v, v >$

$$q : \quad V \ \to \ K$$

zugeordnet. Für diese quadratische Form gilt offensichtlich

$$q(\lambda v) = \lambda^2 q(v)$$

für $\lambda \in K$ und $v \in V$.

Polynomdarstellung: Sei $V = K^n$ und $< v, w >=< v, w >_S$. Die zugeordnete quadratische Form $q(x) = q_S(x)$ schreibt sich als ein homogenes quadratisches Polynom mit Koeffizienten $\lambda_{ij}, i \leq j$ aus K

$$q(x) \ =< x, x >_S = \ x' \cdot S \cdot x \ = \ \sum_{i,j} x_i S_{ij} x_j$$

$$= \ \sum_{i=j} x_i S_{ij} x_j + 2 \sum_{i<j} x_i S_{ij} x_j \ = \ \sum_{i \leq j} \lambda_{ij} x_i x_j \ = \ q(x_1, ..., x_n) \ .$$

Ist umgekehrt ein solches homogen quadratisches Polynom in den Unbestimmten $x_1, ..., x_n$ und den Koeffizienten λ_{ij} Koeffizienten aus K gegeben

$$q(x_1, ..., x_n) = \sum_{i \leq j} \lambda_{ij} x_i x_j \qquad \lambda_{ij} \in K \ ,$$

dann definiert man eindeutig dazu die symmetrische Matrix S mit den Diagonaleneinträgen λ_{ii} und Nebendiagonaleinträgen $\frac{1}{2}\lambda_{ij}$ $(i < j)$. (Hier haben wir die Annahme $2 \neq 0$ benutzt. Die zugehörige symmetrische Bilinearform

$$< x, y >_S = x' \cdot S \cdot y$$

auf $V = K^n$ liefert dann gerade $q(v) =< v, v >_S$.

<u>Koordinatenfreier Zugang</u>: Dieselbe Rechnung in koordinatenfreier Form ist das binomische Gesetz - oder die sogenannte Polarisierungsformel - und zeigt erneut, daß die symmetrische Bilinearform $< \, , \, >$ sich vollständig aus der quadratischen Form q rekonstruieren läßt.

Die Binomische Formel: (Polarisierung)

$$\boxed{< v, w > = \tfrac{1}{2} \left[q(v+w) - q(v) - q(w) \right]}$$

<u>Beweis</u>: Der Beweis benutzt die Symmetrie von $< \, , \, >$ sowie $\frac{1}{2} \in K$. Aus $< v+w, v+w > = < v, v > + < w, v > + < v, w > + < w, w >$ folgt

$$\begin{aligned} q(v+w) - q(v) - q(w) &= \; < v+w, v+w > - < v, v > - < w, w > \\ &= \; < w, v > + < v, w > \\ &= \; 2 < v, w > \end{aligned}$$

<u>Zusammenfassung</u>: Durch Übergang von der nichtlinearen quadratischen Funktion $q(x_1, ..., x_n)$ zur zugehörigen symmetrischen Bilinearform definiert durch S wie oben oder definiert durch die Polarisierungsformel (siehe letzte Seite), wird also quadratische Formen $q(x)$ zu einem Objekt der linearen Algebra oder besser gesagt der bilinearen Algebra!

44 Orthogonalbasen

In diesem Paragraph sei wieder $\frac{1}{2} \in K$.

Ziel dieses Abschnittes ist es, zu einer gegebenen symmetrischen Bilinearform

$$< \, , \, > : V \times V \to K$$

auf einem endlichdimensionalen K–Vektorraum V eine Basis $b_1, ..., b_n$ zu finden für die die zugehörige Matrix S Diagonalengestalt besitzt. Eine solche Basis heißt Orthogonalbasis. Es gilt dann

$$< b_i, b_j > = 0 \qquad , \qquad \forall \, i \neq j.$$

Satz: *Unter der Annahme* $\frac{1}{2} \in K$ *besitzt jede symmetrische Bilinearform auf einem endlichdimensionalen* K*–Vektorraum* V *eine Orthogonalbasis.*

Wir benutzen folgende Sprechweise: Der K-Untervektorraum

$$W^{\perp} := \{ v \in V \mid \, < w, v > = 0 \quad \forall w \in W \}$$

von V nennt man das Orthokomplement von W. Ist W ein K-Untervektorraum von V, dann auch $W^{\perp}$. Die Bildung des Orthokomplements hängt natürlich von der Wahl der symmetrischen K–Bilinearform $< \, , \, >$ auf V ab.

Beweis des Satzes: Ist $< \, , \, >$ identisch null, dann ist nichts zu zeigen. Sei daher die Bilinearform nicht identisch Null. Wegen der Polarisierungsformel (§28) gibt es dann mindestens einen Vektor $v \in V$ mit

$$q(v) = < v, v > \neq 0.$$

Sei Kv der eindimensionale Raum, der von v aufgespannt wird und $(Kv)^{\perp}$ sein Orthokomplement bezüglich der Form $< \, , \, >$.

Lemma: *Für* $< v, v > \neq 0$ *ist* $V = Kv \oplus (Kv)^{\perp}$ *eine direkte Summe.*

Wir zeigen zuerst, daß jeder Vektor $w \in V$ sich als Summe

$$w = \frac{< v, w >}{< v, v >} \cdot v + \left(w - \frac{< v, w >}{< v, v >} \cdot v \right) = w_1 + w_2$$

schreibt mit $w_1 \in Kv$ und $w_2 \in (Kv)^\perp$.
$w_1 \in Kv$ ist klar und $w_2 \in (Kv)^\perp$ ist auch klar wegen

$$< v, w_2 > = < v, w > - \frac{< v, w >}{< v, v >} \cdot < v, v > = 0.$$

Man erhält also

$$V = Kv + (Kv)^\perp.$$

Andererseits ist $Kv \cap (Kv)^\perp = \{0\}$ wegen $< v, v > \neq 0$. Also ist die Summe direkt und die Behauptung bewiesen.

<u>Den Beweis des Satzes vollendet man nun wie folgt</u>:

Da die Dimension von $(Kv)^\perp$ echt kleiner als die Dimension von V ist, besitzt $(Kv)^\perp$ per Induktion bereits eine Orthogonalbasis $b_1, ..., b_{n-1}$ (bzgl. der Einschränkung von $< \ , \ >$). Die gewünschte Orthogonalbasis von V ist dann $b_1, ..., b_{n-1}, b_n = v$, da v nach Konstruktion orthogonal zu allen $b_1, ..., b_{n-1}$ ist.
q.e.d.

In Matrixsprechweise folgt wegen §42 aus dem letzten Satz das

Korollar: *Sei K ein Körper mit $\frac{1}{2} \in K$ und sei $S = S'$ eine symmetrische Matrix in $M_{n,n}(K)$. Dann existiert eine Matrix $T \in \mathcal{G}\ell(n, K)$ so daß TST' diagonal wird*

$$T \cdot S \cdot T' = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix} \quad , \quad a_1, .., a_n \in K \ .$$

T ist hierbei die Basiswechselmatrix beim Übergang von der Standardbasis des K^n zu einer Orthogonalbasis von S.

45 Orthogonale Abbildungen

Sei K ein Körper mit $\frac{1}{2} \in K$.

Lemma 1: *Für symmetrische K-Bilinearformen $< \, , \, >_V$ und $< \, , \, >_W$ auf K–Vektorräumen V resp. W sei $q_V(v) =< v, v >_V$ und $q_W(w) =< w, w >_W$. Für eine K-lineare Abbildung $\varphi : V \to W$ sind dann folgende Eigenschaften äquivalent*

(i) *Längenerhaltung:* $q_W(\varphi(v)) = q_V(v) \qquad \forall v \in V$.

(ii) *Winkelerhaltung:* $< \varphi(v), \varphi(w) >_W \, = \, < v, w >_V \qquad \forall v, w \in V$

(iii) $< \varphi(b_i), \varphi(b_j) >_W \, = \, < b_i, b_j >_V \qquad \forall b_i, b_j$ *einer V-Basis $b_1, ..., b_n$.*

<u>Beweis</u>: (iii) $\Rightarrow$ (ii): $< \varphi(v), \varphi(v') >_W$ ist eine symmetrische Bilinearform auf V, welche nach (iii) dieselbe Matrix S bezüglich der Basis $b_1, .., b_n$ besitzt wie $< v, v' >_V$. Daraus folgt die Gleichheit der Bilinearformen, also Aussage (ii). Die Implikation (ii) $\Rightarrow$ (i) folgt indem man $v = w$ setzt. (i) $\Rightarrow$ (iii) folgt durch Polarisierung: $2\cdot < \varphi(b_i), \varphi(b_j) >_W$ ist

$$< \varphi(b_i + b_j), \varphi(b_i + b_j) >_W - < \varphi(b_i), \varphi(b_i) >_W - < \varphi(b_j), \varphi(b_j) >_W$$

und $2\cdot < b_i, b_j >_V$ ist

$$< b_i + b_j, b_i + b_j >_V - < b_i, b_i >_V - < b_j, b_j >_V \quad .$$

Nach der Annahme (i) stimmen die rechten Seiten beider Formeln überein. Somit sind die linken Seiten gleich. Daraus folgt (iii).

Definition: Eine Abbildung $\varphi : V \to W$, welche eine der drei äquivalenten Eigenschaften (i)-(iii) erfüllt, nennt man eine <u>Isometrie</u> (bezgl. $< \, , \, >_V$ und $< \, . \, >_W$).

Definition: Ist $V = W$ und $< ., . >_V =< ., . >_W$ eine symmetrische Bilinearform auf V, dann nennt man eine <u>bijektive</u> <u>Isometrie</u> $\varphi \in \mathcal{G}\ell(V)$ auch <u>orthogonale</u>

Abbildung bezüglich der Bilinearform $< \, , \, >_V$. Die Menge der orthogonalen Abbildungen bezeichnen wir mit $O_B(V)$.

Lemma 2: *$O_B(V)$ ist eine Untergruppe von $\mathcal{G}\ell(V)$.*

Beweis: Offensichtlich ist $id \in O_B(V)$. Aus $\varphi \in O_B(V)$ folgt $\varphi^{-1} \in O_B(V)$ wegen $< \varphi(\tilde{v}), \varphi(\tilde{w}) > = < \tilde{v}, \tilde{w} >$ für $\tilde{v} = \varphi^{-1}(v)$ und $\tilde{w} = \varphi^{-1}(w)$. Dies impliziert $< v, w > = < \varphi^{-1}(v), \varphi^{-1}(w) >$. Außerdem sind Produkte von $\varphi, \psi \in O_B(V)$ wieder orthogonal $\varphi \circ \psi \in O_B(V)$, denn

$$< (\varphi \circ \psi)(v), (\varphi \circ \psi)(w) > = < \varphi(\psi(v)), \varphi(\psi(w)) >$$

$$\stackrel{\varphi \in O_B(V)}{=} \; < \psi(v), \psi(w) > \; \stackrel{\psi \in O_B(V)}{=} \; < v, w > \ .$$

Lemma 3:(Matrixkriterium) *Ist $V = K^n$ und $< v, w >_V = v'Sw$ für eine symmetrische Matrix $S = S'$, dann ist eine Matrix M eine Isometrie genau dann wenn gilt*

$$M' \cdot S \cdot M = S \ .$$

Beweis: Die beiden Bilinearformen $< M(v), M(w) >_V = (Mv)'S(Mw) = v'M'SMw = < v, w >_{M'SM}$ und $< v, w >_V = v'Sw = < v, w >_S$ stimmen genau dann überein, wenn gilt $M'SM = S$.

Lemma 4: *Sei $S = S'$ eine symmetrische invertierbare Matrix in $M_{n,n}(K)$. Dann ist jede Isometrie $M : K^n \to K^n$ der Bilinearform $< v, w >_S$ ist invertierbar und erfüllt $det(M) \in \{\pm 1\}$.*

Beweis: Aus dem Matrixkriterium folgt durch Determinantenbildung

$$det(M') \cdot det(S) \cdot det(M) \ = \ det(S) \ .$$

Wegen $det(M') = det(M)$ und $det(S) \neq 0$ (letzteres nach Annahme) folgt $det(M)^2 = 1$ beziehungsweise $det(M) = \pm 1$. Insbesondere ist $det(M) \neq 0$ und somit M invertierbar. Beachte, daß aus der Annahme $\frac{1}{2} \in K$ folgt $1 \neq -1$, und daß das Polynom $x^2 - 1$ höchstens zwei verschiedene Nullstellen in K besitzen kann, nämlich dann ± 1.

Bemerkung: Im allgemeinen ist eine Isometrie nicht notwendigerweise invertierbar. Zum Beispiel ist im Fall $S = 0$ jede lineare Abbildung M eine Isometrie!

Definition: Sei $S = S'$ eine symmetrische invertierbare Matrix in $M_{n,n}(K)$. Der Kern des Gruppenhomomorphismus

$$det : O_S(n, K) \to \{-1, 1\}$$

heißt spezielle orthogonale Gruppe. Dies definiert eine Untergruppe $SO_S(n, K)$ der orthogonalen Gruppe $O_S(n, K)$.

Beispiel: Im Spezialfall $V = \mathbb{R}^2$ und $S = E$ ist ein Endomorphismus M in der speziellen orthogonalen Gruppe genau dann wenn gilt

$$M \ = \ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \ .$$

Die Spiegelung $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ ist zwar orthogonal, hat dagegen Determinante -1.

Beweis: Dies folgt aus dem Matrixkriterium $M'SM = S$ oder hier $M'M = E$ wegen $S = E$. Hat M die reellen Einträge a, b, c, d, dann liefert das Matrixkriterium die Bedingungen: $a^2 + c^2 = 1$, $b^2 + d^2 = 1$ und $ab + cd = 0$. Mit anderen Worten (a, b) und (c, d) müssen zwei orthogonale Vektoren der Länge 1 in der Ebene $\mathbb{R}^2$ bilden. Es gibt dann (Analysis!) ein θ mit $a = \cos\theta$ und $c = \sin\theta$. Dann ist notwenig (c, d) proportional zum Normalenvektor $(-\sin\theta, \cos\theta)$ von $(\cos\theta, \sin\theta)$. Die Matrix M ist also in $O(2, \mathbb{R})$ genau dann wenn

$$M \ = \ \begin{pmatrix} \cos\theta & \lambda\cdot(-\sin\theta) \\ \sin\theta & \lambda\cdot\cos\theta \end{pmatrix} \quad , \quad \lambda^2 = 1 \ ,$$

und es gilt $det(M) = \lambda$.

46 Der Satz von Sylvester

Wir betrachten in den beiden nächsten Paragraphen ausschließlich den Fall des Körpers $K = \mathbb{R}$.

Für eine Orthogonalbasis $b_1, ..., b_n$ einer symmetrischen K–Bilinearform $< \, , \, >$ auf einem Vektorraum V gilt

$$< b_i, b_j > = a_i \cdot \delta_{ij} \quad , \quad a_i \in K$$

für alle i, j. Die Form $< \, , \, >$ ist nichtausgeartet genau dann wenn gilt $a_i \in K^*$. Eine Orthogonalbasis $b_1, ..., b_n$ heißt <u>Orthonormalbasis</u>, falls $a_i = 1$ für alle i.

Ersetzt mit die Vektoren b_i einer <u>Orthogonalbasis</u> von $< \, , \, >$ durch skalare Vielfache $\lambda_i \cdot b_i$ erhält wieder eine Orthogonalbasis von $< \, , \, >$. Die 'neuen Diagonaleinträge ' sind dann $\lambda_i^2 \cdot a_i$. Waren alle a_i Quadrate in K^*, so kann man auf diese Weise alle $\tilde{a}_i$ zu 1 machen. Mit anderen Worten: $\lambda_1 \tilde{b}_1, .., \lambda_n b_n$ bildet dann eine Orthonormalbasis. Etwas allgemeiner zeigt dieser Schluß

Folgerung: *Jede nichtausgeartete symmetrische Bilinearform $< \, , \, >$ auf $V = K^n$ besitzt im Fall $K = \mathbb{R}$ eine Orthogonalbasis $b_1, .., b_n$ mit der Eigenschaft*

$$< b_i, b_j > = a_i \cdot \delta_{ij} \quad , \quad a_i \in \{\pm 1\} \ .$$

Zusatz: *Im Fall $K = \mathbb{C}$ besitzt $< \, , \, >$ sogar eine Orthonormalbasis.*

Weiterhin gilt

Sylvester'sche Trägheitssatz: *Sei $K = \mathbb{R}$. Dann ist die Anzahl der Vektoren b_i mit $< b_i, b_i > = 1$ resp $< b_i, b_i > = -1$ einer Orthogonalbasis einer nichtausgearteten symmetrischen Bilinearform $< \, , \, >$ auf $\mathbb{R}^n$ unabhängig von der Wahl der Orthogonalbasis.*

<u>Beweis</u>: Seien $b_1, .., b_n$ und $\tilde{b}_1, .., \tilde{b}_n$ Orthogonalbasen mit $< b_i, b_i >= a_i$ und $< \tilde{b}_i, \tilde{b}_i >= \tilde{a}_i$. Sei

$$a_i = 1 \text{ für } i = 1, .., r \text{ und } a_i = -1 \text{ für } i = r+1, .., n$$

sowie

$\tilde{a}_i = 1$ für $i = 1, .., s$ und $\tilde{a}_i = -1$ für $i = s+1, .., n$.

Für $x \in V$ gilt $x = \sum_{i=1}^{n} x_i b_i = \sum_{i=1}^{n} y_i \tilde{b}_i$. Die Koordinaten $y_i = L_i(x)$ sind Linearform in x und es gilt

$$(*) \begin{cases} < x, x > = x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_n^2 \\ < x, x > = L_1(x)^2 + \cdots + L_s(x)^2 - L_{s+1}(x)^2 - \cdots - L_n(x)^2 \end{cases}$$

Im Fall $r < s$ besitzen die $r + (n - s) = n + (r - s) < n$ Gleichungen

$$x_1 = \cdots x_r = L_{s+1}(x) = \cdots = L_n(x) = 0$$

mindestens eine nichttriviale Lösung $x \neq 0$ (da n Unbekannte x_i!). Ein Positivitätsargument zeigt dann wegen der Identität $(*)$ sofort

$$x_{r+1} = \cdots = x_n = L_1(x) = \cdots = L_s(x) = 0 \; .$$

Dies erzwingt $x = 0$ im Widerspruch zu $x \neq 0$. Es folgt $r \leq s$ und aus Symmetriegründen dann $r = s$.

Bemerkung: Wie der Sylvestersche Trägheitssatz zeigt, besitzt eine nichtausgeartete symmetrische Bilinearform über einem Körper K (mit $\frac{1}{2} \in K$) zwar immer eine Orthogonalbasis aber nicht immer eine Orthonormalbasis.

47 Der Sylvestertyp

Wir wollen ein Kriterium angeben, wann eine gegebene symmetrische nichtausgeartete Bilinearform

$$< \, , \, > : V \times V \to \mathbb{R}$$

auf einem endlich dimensionalen $\mathbb{R}$-Vektorraum eine Orthonormalbasis besitzt.

Sei genauer $V = \mathbb{R}^n$ mit Standardbasis $e_1, .., e_n$. Dann bestimmt die symmetrische Matrix S mit den Einträgen $S_{ij} = < e_i, e_j >$ die Form $< \, , \, >$ eindeutig. Sei $1 \leq r \leq n$. Die Determinante der $r \times r$ Teilmatrix

$$m_r(S) = det \begin{pmatrix} s_{11} & \cdots & s_{1r} \\ \vdots & & \vdots \\ s_{r1} & \cdots & s_{rr} \end{pmatrix}$$

heißt r-ter Hauptminor von S.

Notation: Für reelle λ, μ schreiben wir

$$\lambda \sim \mu ,$$

falls $\lambda = c \cdot \mu$ für positives $c > 0$ in $\mathbb{R}$ gilt.

Arbeitshypothese: Wir nehmen an für die symmetrische Matrix $S = S'$ gilt

$$m_1(S) \neq 0, m_2(S) \neq 0, ..., m_s(S) \neq 0 \ .$$

Wir nehmen weiterhin an, es sei $r < s \leq n$ und wir hätten bereits r Vektoren $b_1, ..., b_r \in V$ konstruiert mit den folgenden Eigenschaften:

$$\mathbb{R}b_1 + ... + \mathbb{R}b_\nu = \mathbb{R}e_1 + ... + \mathbb{R}e_\nu$$

erzeugen denselben Teilraum $V_\nu \subset V$ für alle $\nu = 1, .., r$. Weiterhin sei $b_1, .., b_r$ eine Orthogonalbasis von V_r

$$(< b_i, b_j >) = diag(< b_1, b_1 >, ..., < b_n, b_n >) \ .$$

Schließlich soll gelten

$$\prod_{i=1}^{\nu} < b_i, b_i > \sim \ m_\nu(S)$$

für alle $\nu = 1, ..., r$.

Wir wollen versuchen einen weiteren Vektor b_{r+1} hinzuzufügen, derart daß die entsprechende Aussage für die Vektoren $b_1, .., b_{r+1}$ und $r+1$ anstelle von r gilt: Aus der Arbeitshypothese und der Annahmen an $b_1, ..b_r$ folgt $< b_i, b_i > \neq 0$ für alle $i = 1, .., r$ da ansonsten ein $m_j(S)$ Null sein müsste.

<u>Ansatz</u>: Der Vektor

$$b_{r+1} := e_{r+1} - \sum_{i=1}^{r} \frac{< b_i, e_{r+1} >}{< b_i, b_i >} \cdot b_i$$

ist daher wohldefiniert. Nach Konstruktion gilt dann

$$\mathbb{R}b_1 + ... + \mathbb{R}b_{r+1} = \mathbb{R}e_1 + ... + \mathbb{R}e_{r+1} \ \ =: V_{r+1}$$

und $b_1, ..., b_r, b_{r+1}$ ist eine Orthogonalbasis der Einschränkung der Form $< ., . >$ auf V_{r+1}. Wegen $< b_i, b_j >= \delta_{ij}$ gilt nämlich

$$< b_{r+1}, b_j >=< e_{r+1}, b_j > - \sum_{i=1}^{r} \frac{< b_i, e_{r+1} >}{< b_i, b_i >} \cdot < b_i, b_j >= 0$$

für alle $j = 1, .., r$. Die Einschränkung von $< \, , \, >$ auf diesen $r+1$-dimensionalen Raum läßt sich bezüglich beider Basen beschreiben. Bis auf einen Basiswechsel $T \in \mathcal{G}\ell(r+1, \mathbb{R})$ gilt daher

$$(< b_i, b_j >) = T \cdot (< e_i, e_j >) \cdot T'$$

für $1 \leq i, j \leq r+1$. Wegen $det(T) \cdot det(T') = (det(T))^2 > 0$ folgt daraus durch Determinantenbildung

$$\prod_{i=1}^{r+1} < b_i, b_i > \sim \ m_{r+1}(S) \ ,$$

da $b_1, ..., b_{r+1}$ eine Orthogonalbasis von V_{r+1} ist. Dies zeigt den Induktionsschritt.

Da der Induktionsanfang für $b_1 = e_1$ trivialerweise erfüllt ist folgt

Satz: *Sei $S = S' \in M_{n,n}(\mathbb{R})$ eine reelle symmetrische Matrix, deren Hauptminoren alle nicht verschwinden. Dann gibt es eine invertierbare reelle Matrix T mit der Eigenschaft*

$$TST' \;=\; diag\Big(m_1(S), m_2(S)/m_1(S), ..., m_n(S)/m_{n-1}(S)\Big) \;.$$

Insbesondere kann man den Sylvestertyp der Matrix S von den Vorzeichen der Minorenquotienten $m_{i+1}(S)/m_i(S)$ ablesen!

<u>Beispiel</u>: Für $S = \begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix}$ ist $m_1(S) = 2$ und $m_2(S) = 2$.

Normierte Räume

48 Komplexe Konjugation

Für alle nun folgenden Paragraphen dieses Kapitels ist der Körper K entweder $K = \mathbb{R}$ oder $K = \mathbb{C}$. Wir betrachten zusätzlich einen **Automorphismus des Körpers** K

$$\sigma(z+w) = \sigma(z) + \sigma(w) \quad ,$$

$$\sigma(z \cdot w) = \sigma(z) \cdot \sigma(w) \quad .$$

Hierbei wählen wir für σ im folgenden die Identität – im Fall $K = \mathbb{R}$ – beziehungsweise $\sigma(z) = \overline{z}$ (die komplexe Konjugation) – im Fall $K = \mathbb{C}$. Die **komplexe Konjugation** ist definiert durch

$$\sigma(x+iy) = x - iy$$

für $x, y \in \mathbb{R}$. Obige Eigenschaften sind leicht zu verifizieren (siehe Appendix). In beiden Fällen gilt $\sigma^2 = id$. Für $z \in \mathbb{C}$ gilt $\sigma(z) = z$ genau dann wenn z reell ist.

<u>Fortsetzung auf Matrizen</u>: Ist M eine Matrix mit Koeffizienten in K, dann schreiben wir $\sigma(M)$ oder $\overline{M}$ für die Matrix, welche aus M entsteht wenn man auf alle Koeffizienten die komplexe Konjugation σ anwendet

$$\sigma(M)_{ij} = \sigma(M_{ij}) \ .$$

<u>Definition</u>: Für eine Matrix $M \in M_{m,n}(K)$ sei $M^t \in M_{n,m}(K)$ die Matrix mit den Einträgen $(M^t)_{ij} = \sigma(M_{ji})$. Mit anderen Worten

$$M^t = \sigma(M') = (\sigma(M))'$$

$$(M \cdot N)^t = N^t \cdot M^t \quad , \quad (M^t)^t = M \ .$$

Appendix

Für komplexe Zahlen $z = a + bi$ definieren wir

$$N(z) = z\overline{z} = a^2 + b^2.$$

Hierbei ist $\overline{z} = a - bi$ die konjugiert komplexe Zahl. Es gilt

$$N(z) \geq 0 \textit{ und } N(z) = 0 \iff z = 0.$$

Außerdem gilt

$$N(z \cdot w) = N(z) \cdot N(w) \qquad \forall z, w \in \mathbb{C}.$$

<u>Beweis</u>: Wegen $Norm(z \cdot w) = zw \cdot \overline{zw} \stackrel{!}{=} z \cdot w \cdot \overline{z} \cdot \overline{w} = z\overline{z} \cdot w\overline{w} = N(z) \cdot N(w)$ genügt zu zeigen

$$\overline{z \cdot w} = \overline{z} \cdot \overline{w} \ .$$

Beweis von $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$:

$$\begin{aligned}
\overline{(a+bi) \cdot (a'+b'i)} &= (aa' - bb') - (ab' + a'b)i \\
\overline{(a+bi)} \cdot \overline{(a'+b'i)} &= (a - bi) \cdot (a' - b'i) \\
&= (aa' - (-b)(-b')) - (a(-b') + a'(-b))i \\
&= (aa' - bb') - (ab' + a'b)i
\end{aligned}$$

Analog zeigt man $\overline{z + w} = \overline{z} + \overline{w}$.

Folgerung: *Die Abbildung $\sigma(z) = \overline{z}$ ist ein Körper Automorphismus von $\mathbb{C}$.*

<u>Beweis</u>: $z \to \overline{z}$ ist eine bijektive Abbildung von $\mathbb{C}$ auf $\mathbb{C}$ wegen $\overline{(\overline{z})} = z$ und sie respektiert Addition sowie Multiplikation.

49 Sesquilinearformen

Definition: Ist V eine K-Vektorraum, dann heißt eine Abbildung

$$< .,. >: \ V \times V \to K$$

sesquilinear, falls sie $\mathbb{R}$-linear ist und falls gilt

$$< \lambda v, \mu w >= \sigma(\lambda)\mu < v, w > \quad , \quad \forall v, w \in V \ \forall \lambda, \mu \in K \ .$$

Eine solche Sesquilinearform heißt hermitesch, falls gilt

$$< v, w > = \ \sigma(< w, v >) \ .$$

Insbesondere folgt $< v, v > = \ \sigma(< v, v >)$ und somit

Lemma 1: *Für hermitesche Sesquilinearformen gilt $< v, v > \in \ \mathbb{R}$ für alle $v \in V$.*

Beispiel: Im Fall $K = \mathbb{R}$ ist sesquilinear gleichbedeutend mit K-linear und hermitesch gleichbedeutend mit symmetrisch.

Beispiel: Im Fall $V = \mathbb{C}^n$ und $K = \mathbb{C}$ sei $< .,. >$ die Standardform $< v, w >= \sum_{i=1}^n \overline{v}_i w_i$ für $v = (v_1, .., v_n)$ und $w = (w_1, .., w_n)$. Diese ist sesquilinear und hermitesch.

Definition: Eine hermitesche, Sesquilinearform auf V heißt positiv definit, falls

$$< v, v > > \ 0 \quad , \quad \forall \ 0 \neq v \in V$$

beziehungsweise heißt semidefinit, falls gilt

$$< v, v > \geq \ 0 \quad , \quad \forall \ 0 \neq v \in V \ .$$

Beispiel: Die sesquilineare Standardform auf dem $\mathbb{C}$-Vektorraum $V = \mathbb{C}^n$ ist definit; ebenso die bilineare Standardform auf dem $\mathbb{R}$-Vektorraum $\mathbb{R}^n$.

Übungsaufgabe: Sei V der K-Vektorraum der stetigen Funktionen $f : [a, b] \to K$ auf einem reellen Intervall (für $a \leq b$ aus $\mathbb{R}$). Dann definiert

$$< f, g > = \int_a^b \overline{f}(t) \cdot g(t) dt$$

eine positiv definite, hermitesche Sesquilinearform auf V. (Das Integral wird hierbei definiert als das Integral über den Realteil $\varphi(t)$ plus i mal das Integral über den Imaginärteil $\psi(t)$ des Produktes $\overline{f}(t) \cdot g(t) = \varphi(t) + i\psi(t)$. Sowohl $\varphi(t)$ als auch $\psi(t)$ sind stetige Funktionen und somit Riemann integrierbar auf dem Intervall).

1.Bemerkung: Die Einschränkung einer (positiv definiten) hermiteschen Sesquilinearform von V auf einen K-Untervektorraum $W \subset V$ ist wieder eine (positiv definite) hermitesche Sesquilinearform.

2.Bemerkung: Jede positiv definite hermitesche Sesquilinearform ist nichtausgeartet.

Beweis: Zu zeigen ist: $< v, w > = 0$ für alle $w \in V$ impliziert $v = 0$. Zum Beweis genügt es $w = v$ zu setzen wegen $< v, v > > 0$ für $v \neq 0$!

3.Bemerkung: Jede Sesquilinearform auf $V = K^n$ wird wie folgt beschrieben durch eine quadratische Matrix $H = (H_{ij}) \in M_{n,n}(K)$

$$H_{ij} = < e_i, e_j >$$

$$< v, w > = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma(v_i) H_{ij} w_j \ .$$

In Matrixsprechweise

$$< v, w > = v^t \cdot H \cdot w \quad , \quad v, w \in K^n \ .$$

Lemma 2: *Die der Matrix H zugeordnete Sesquilinearform $< ., . >$ auf $V = K^n$ ist genau dann hermitesch, wenn gilt*

$$H^t = H \ .$$

Beweis: Übungsaufgabe!

4.Bemerkung: Ist $< v, w > = v^t \cdot H \cdot w$ eine positiv definite hermitesche Sesquilinear auf $V = K^n$, dann schreiben wir kurz $H > 0$. Wir behaupten (eine Variante von Bemerkung 2)

$$H > 0 \quad \Rightarrow \quad det(H) \neq 0 \ .$$

Beweis: Aus $H \cdot v = 0$ folgt $v^t H v = 0$, also $v = 0$. Somit $Kern(H) = 0$. $H : K^n \to K^n$ injektiv, also bijektiv und somit invertierbar.

50 Normen

Sei wieder $K = \mathbb{R}$ oder $K = \mathbb{C}$ und $\sigma : K \to K$ sei wie im letzten Abschnitt und sei $< .,. >$ eine positiv definite hermitesche Sesquilinearform auf V.

Unter den obigen Annahmen definiert

$$\begin{array}{l} V \to \mathbb{R}_{\geq 0} \\ v \;\mapsto {}_{+}\sqrt{< v, v >} = \|v\| \end{array}$$

eine eine reellwertige Funktion auf V. Diese ist wohldefiniert, da nach Annahme $< v, v > \geq 0$ gilt. Insbesondere ist $\|v\| = 0$ genau dann, wenn $v = 0$. Der Wert $\|.\|$ hängt wohlgemerkt von der Wahl der Sesquilinearform ab.

Geometrische Deutung: Bei fester Wahl der Form $< .,. >$ hat $\|.\|$ die Eigenschaft einer Längenfunktion. Genauer: $\|v\|$ wird als **Norm** bezeichnet und als Länge des Vektors v gedeutet.

Eigenschaften der Norm: Setze $|\lambda| = {}_{+}\sqrt{\overline{\lambda}\lambda}$ für $\lambda \in \mathbb{C}$. Beachte $|x + iy| = {}_{+}\sqrt{x^2 + y^2}$ ist nach Pythagoras die euklidsche Länge des Vektors $\lambda = x + iy = (x, y) \in \mathbb{R}^2$ im Fall $\lambda \in \mathbb{C}$. Analog ist $|\lambda| = {}_{+}\sqrt{\lambda^2}$ der Absolutbetrag der reellen Zahl λ im Falle $\lambda \in \mathbb{R}$.

Offensichtlich gilt $\|\lambda \cdot v\|^2 = \overline{\lambda}\lambda \cdot \|v\|^2$. Es folgt

$$\|\lambda \cdot v\| \;=\; |\lambda| \cdot \|v\| \quad , \quad \lambda \in K, v \in V \; .$$

Schwarzsche Ungleichung: *Sei $< .,. >$ eine positiv definite, hermitesche Sesquilinearform auf einem K-Vektorraum V. Für alle $v, w \in V$ gilt dann*

$$\boxed{\;|< v, w >| \;\leq\; \|v\| \cdot \|w\|\;}$$

Gleichheit gilt genau dann, wenn v, w linear abhängig sind.

Beweis: Einerseits ist wegen der Definitheit

$$0 \overset{pos.}{\leq} < \lambda v + w, \lambda v + w > \ .$$

Andererseits ist die rechte Seite wegen der Sesquilinearität und der hermiteschen Symmetrie $\overline{\lambda} < v, \lambda v + w > + < w, \lambda v + w > \overset{sesquilinear.}{=} |\lambda|^2 < v, v > + \overline{\lambda} < v, w > + \lambda < w, v > + < w, w >$, also gleich

$$|\lambda|^2 \|v\|^2 + 2 \cdot Re(\lambda < w, v >) + \|w\|^2 \ .$$

Sei $v \neq 0$ (im Fall $v = 0$ ist die Schwarzsche Ungleichung trivial). Wir spezialisieren dann λ und setzen $\lambda = \frac{-<v,w>}{\|v\|^2}$. Man erhält

$$\begin{aligned} 0 &\leq \frac{|< v, w >|^2}{\|v\|^2} - 2\frac{< v, w >}{\|v\|^2} < w, v > + \|w\|^2 \\ &= \frac{-|< v, w >|^2 + \|w\|^2 \|v\|^2}{\|v\|^2} \end{aligned}$$

Wegen $\|v\| > 0$ folgt daraus $| < v, w > |^2 \leq \|v\|^2 \|w\|^2$. Dies beweist die Schwarzsche Ungleichung durch Wurzelziehen.

Im Gleichheitsfall $| < v, w > | = \|v\| \cdot \|w\|$ ist entweder $v = 0$ oder

$$< \lambda v + w, \lambda v + w >= |\lambda|^2 \|v\|^2 + 2 \cdot Re(\lambda < w, v >) + \|w\|^2$$

hat $\lambda_0 = \frac{-<v,w>}{\|v\|^2} \in \mathbb{R}$ als Nullstelle. Wegen Definitheit folgt daraus $\lambda_0 v + w = 0$. Somit sind v und w linear abhängig. Dies beweist den Zusatz.

Folgerungen:

1. $\left| \sum\limits_{i=1}^{n} \overline{x}_i y_i \right|^2 \leq \sum\limits_{i=1}^{n} |x_i|^2 \cdot \sum\limits_{i=1}^{n} |y_i|^2$

2. $\left| \int\limits_a^b \overline{f}(t) g(t) dt \right|^2 \leq \int\limits_a^b |f(t)|^2 dt \cdot \int\limits_a^b |g(t)|^2 dt$

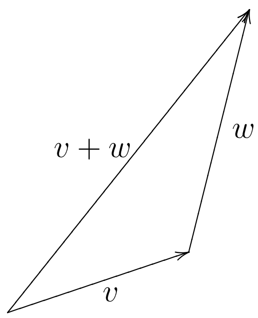
Dreiecksungleichung: *Sei $< .,. >$ eine positiv definite hermitesche Sesquilinearform auf dem K-Vektorraum V. Für die zugehörige Längenfunktion gilt*

$$\|v+w\| \ \leq \ \|v\| + \|w\| \ .$$

<u>Beweis</u>: Aus $|Re(z)| = |x| \leq \sqrt{x^2+y^2} = |z|$ für $z = x+iy$ aus $\mathbb{C}$ folgt

$$\begin{array}{rcl} \|v+w\|^2 & = & < v+w, v+w > \\ & = & \|v\|^2 + 2Re(< w,v >) + \|w\|^2 \\ & \leq & \|v\|^2 + |2 < w,v > | + \|w\|^2 \\ & \underset{Schw.Ungl.}{\leq} & \|v\|^2 + 2\|w\|\|v\| + \|w\|^2 \\ & = & (\|v\| + \|w\|)^2 \end{array}$$

Die Dreiecksungleichung folgt durch Wurzelziehen, da die beiden äußeren Terme $\|v+w\|^2$ und $(\|v\|+\|w\|)^2$ nicht negativ sind.



<u>Geometrische Deutung</u>: In der Ebenengeometie besitzt die Schwarzsche Ungleichung eine <u>explizite</u> Deutung! Dies ist der Spezialfall $K = \mathbb{R}$ und $V = \mathbb{R}^2$ für das standard Skalarprodukt $< .,. >$ des $\mathbb{R}^2$. Es gilt dann nämlich

$$cos(\alpha) \ = \ \frac{< v,w >}{\|v\|\|w\|} \ ,$$

wobei α der Winkel zwischen den Vektoren v, w der Ebene $\mathbb{R}^2$ ist. Die Schwarzsche Ungleichung ergibt sich in diesem Spezialfall aus der Tatsache, daß der Cosinus nur Werte zwischen -1 und 1 annimmt.

* Spinornorm

Für die symmetrische Matrix

$$S = diag(1, -1, ., -1)$$

vom Sylvestertyp $(1, n-1)$ liegen $g = -E$ und $g = S$ in der orthogonalen Gruppe $O_S(\mathbb{R}, n)$, denn in beiden Fällen gilt $g'Sg = S$. Für beliebiges $g \in SO_S(n, \mathbb{R})$ ist damit auch $g' = Sg^{-1}S$ in $SO_S(n, \mathbb{R})$.

$g \in O_S(n, \mathbb{R})$ bedeutet $g'Sg = S$. Dies impliziert $(g'Sg)_{11} = S_{11}$ oder mit anderen Worten $(*)$

$$(g_{11})^2 - \sum_{i=2}^{n} (g_{1i})^2 = 1 \ .$$

Für g' anstelle von g folgt analog $(**)$

$$(g_{11})^2 - \sum_{i=2}^{n} (g_{i1})^2 = 1 \ .$$

Insbesondere ist $|g_{11}| \geq 1$ nicht Null und besitzt ein eindeutig bestimmtes Vorzeichen.

Definition: Für $g \in O_S(n, \mathbb{R})$ sei $N(g) \in \{\pm 1\}$ das Vorzeichen von g_{11}.

Lemma: *Die Abbildung* $N : O_S(n, \mathbb{R}) \to \{\pm 1\}$ *definiert einen surjektiven Gruppenhomomorphismus.*

Wegen $N(-E) = -1$ führt man den Beweis zurück auf den Nachweis der folgenden Aussage: Für $g, h \in O_S(n, \mathbb{R})$ mit $g_{11} > 0, h_{11} > 0$ gilt

$$(gh)_{11} = g_{11} \cdot h_{11} + \sum_{i=2}^{n} g_{1i} \cdot h_{i1} > 0 \ .$$

Dies folgert man aber leicht aus $(*), (**)$ mittels der Schwarzschen Ungleichung

$$|\sum_{i=2}^{n} g_{1i} \cdot h_{i1}| \leq \sqrt{\sum_{i=2}^{n} (g_{1i})^2} \cdot \sqrt{\sum_{i=2}^{n} (h_{i1})^2} < g_{11} \cdot h_{11} \ .$$

51 Selbstadjungierte Abbildungen

Sei wieder $K = \mathbb{R}, \mathbb{C}$ und $< .,. >$ eine hermitesche Sesquilinearform.

Definition: Seien N und M Endomorphismen eines K-Vektorraums V. Wir nennen N die zu M **Adjungierte** bezüglich der Bilinearform $< .,. >$, falls

$$< Mv, w > = < v, Nw > \qquad , \quad \forall v, w \in V$$

gilt. Wir benutzen in diesem Fall die Notation

$$N = M^* \ .$$

Bemerkung: Bei dieser Bezeichnungweise wird angenommen, daß die Form $< .,. >$ fest gewählt ist. M^* ist abhängig von der Wahl der Form $< .,. >$.

Lemma 1: *Sei $V = K^n$ ein endlich dimensionaler K-Vektorraum und $< .,. >$ eine nicht ausgeartete Sesquilinearform auf V. Dann besitzt jeder Endomorphismus $M : V \to V$ eine eindeutig bestimmte Adjungierte M^*. Es gilt*

$$(M^*)^* = M \ .$$

Beweis: Sei $< v, w > = v^t H w$. Da $< .,. >$ nicht ausgeartet ist, ist H invertierbar. Wegen $< Mv, w > = (Mv)^t H w = v^t (M^t) H w = v^t H (H^{-1} M^t H) w = < v, (H^{-1} M^t H) w >$ ist das gesuchte M^* gleich

$$M^* = H^{-1} \cdot M^t \cdot H \ .$$

Offensichtlich ist M^* eindeutig bestimmt und man zeigt leicht $(M^*)^* = M$.

Beispiel: Im Fall der Standardform $< v, w >= v^t \cdot w$ gilt $M^* = M^t$.

Definition: Eine K-lineare Abbildung $M : V \to V$ heißt **selbstadjungiert** bezüglich der Sesquilinearform $< .,. >$, falls gilt

$$M^* = M$$

beziehungsweise **normal**, falls gilt

$$M^* \cdot M \;=\; M \cdot M^* \;.$$

Beispiele: 1) Jede selbstadjungierte Abbildung ist normal.

2) Jede antiselbstadjungierte Abbildung M, d.h. $M^* = -M$, ist normal.

3) Jeder Endomorphismus M (eines endlich dimensionalen K-Vektorraums) ist die Summe zweier normaler Endomorphisen

$$M = M_1 + M_2$$

mit $M_1^* = M_1$ und $M_2^* = -M_2$. Setze $M_1 = \frac{1}{2}(M+M^*)$ und $M_2 = \frac{1}{2}(M-M^*)$.

4) M genau dann normal, wenn die beiden in Punkt 3) definierten Zerlegungskomponenten M_1 und M_2 miteinander kommutieren

$$M_1 \cdot M_2 = M_2 \cdot M_1 \;.$$

5) Für $K = \mathbb{C}$ ist M genau dann selbstadjungiert, wenn $i \cdot M$ antiselbstadjungiert ist.

Lemma 2: *Sei $< .,. >$ eine hermitesche Sesquilinearform auf einem K-Vektorraum V und sei $M : V \to V$ ein selbstadjungierter Endomorphismus. Dann definiert*

$$< M(v), w >$$

wieder eine hermitesche Sesquilinearform auf V. Insbesondere gilt für alle $v \in V$

$$< M(v), v > \in \quad \mathbb{R} \;.$$

Beweis: Die Sesquilinearität ist unmittelbar klar. Weiterhin gilt $< M(w), v > = \; < w, M^*(v) > = \; < w, M(v) >$, denn M ist selbstadjungiert. Andererseits gilt $< w, M(v) > = \; \sigma(< M(v), w >)$. Also ist $< M(.), . >$ hermitesch.

Annahme: Für die folgenden Betrachtungen machen wir der Einfachheit halber etwas schärfere Annahmen. Sei dazu V ein endlich dimensionaler K-Vektorraum, und $< .,. >$ sei eine positiv definite hermitesche Sesquilinearform. Wie wir bereits in § 49 Bemerkung 4) gezeigt haben, ist dann die Form $< .,. >$ nicht ausgeartet. Konkret gilt $< v, w > = \ v^t H w$ für eine hermitesche Matrix $H = H^t$ mit $det(H) \neq 0$.

Formenvergleich: Für selbstadjungierte Endomorphismen $M : V \to V$ gilt nach Definition $M^* = M$. Zur Erinnerung: $M^* = H^{-1} \cdot M^t \cdot H$ (Lemma 1). Wegen $H = H^t$ folgt daher aus $M^* = M$ die Hermitizität der Matrix $H_1 = H \cdot M$

$$H_1^t = H_1 \ .$$

Die hermitesche Sesquilinearform $< M(v), w >$ ist also gegeben durch

$$< M(v), w >= v^t \cdot H_1 \cdot w \quad , \quad H_1 = H \cdot M = M^t \cdot H \ .$$

Sei nun allgemeiner $< v, w >_1 = v^t H_1 w$ eine beliebige hermitesche Sesquilinearform auf V. Dann gilt $(H_1)^t = H_1$. Setzt man $M = H^{-1} \cdot H_1$, so ist M selbstadjungiert bezüglich der Form $< .,. >$ wegen

$$M^* = H^{-1}(H^{-1}H_1)^t H = H^{-1} H_1^t (H^{-1})^t H = H^{-1} H_1 = M \ .$$

Korollar: *Sei V ein endlich dimensionaler K-Vektorraum mit einer fixierten positiv definiten hermiteschen Sesquilinearform $< .,. >$. Dann entsprechen die hermiteschen Sesquilinearformen $< .,. >_1$ auf V eineindeutig den bezüglich $< .,. >$ selbstadjungierten Endomorphismen $M : V \to V$ vermöge der Zuordnung*

$$< v, w >_1 = < M(v), w > \ .$$

Bemerkung: Sind X, Y antiselbstadjungierte Endomorphismen $X^* = -X, Y^* = -Y$. Dann ist offensichtlich (!) auch der Kommutator

$$[X, Y] = X \cdot Y - Y \cdot X$$

wieder antiselbstadjungiert. In der Physik spielen Orts- und Impulsoperatoren eine Rolle. Für diese gilt die Kommutatorrelation

$$[X, Y] = 2\pi i \cdot \hbar \cdot id_V \ .$$

Für Zustände $v \in V$ mit $\|v\| = 1$ beobachtet man die (quadratischen) Erwartungswerte $\|Xv\|$ und $\|Yv\|$ von Ort und Impuls. In obiger Situation gilt dann als formale Anwendung der Schwarzschen Ungleichung die sogenannte Heisenbergsche Unschärferelation.

$$\pi \cdot \hbar \quad \leq \quad \|Xv\| \cdot \|Yv\| \ .$$

Beweis: $2\pi\hbar = | < v, 2\pi i\hbar \cdot v > | = | < v, (XY - YX)v > | = | < -Xv, Yv > + < Yv, Xv > | \ \leq \ 2 \cdot \|Xv\| \cdot \|Yv\|$ wegen der Schwarzschen Ungleichung und der Dreiecksungleichung.

52 Der Spektralsatz

Sei K entweder $\mathbb{R}$ oder $\mathbb{C}$.

Satz: *Sei V ein endlicher K-Vektorraum, $< \, , \, >$ eine positiv definite hermitesche Sesquilinearform auf V, sowie $\varphi \in End_K(V)$ ein selbstadjungierter Endomorphismus bezüglich der Form $< \, , \, >$. Unter dieser Vorausetzung existiert eine Orthonormalbasis $b_1, ..., b_n$ des Vektorraums V*

$$< b_i, b_j > = \delta_{ij} \qquad 1 \leq i, j \leq n = dim_K(V)$$

welche aus Eigenvektoren von φ besteht

$$\varphi(b_i) = \lambda_i \cdot b_i \quad , \quad i = 1, ..., n$$

mit reellen Eigenwerten

$$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n \ .$$

Korollar: *Jede reelle symmetrische Matrix $M \in M_{n,n}(\mathbb{R})$ (jede hermitesche komplexe Matrix in $M_{n,n}(\mathbb{C})$) hat n reelle Eigenwerte. Das charakteristische Polynom zerfällt in ein Produkt reeller Linearfaktoren*

$$\chi_M(t) = \prod_{i=1}^{n} (t - \lambda_i) \qquad \lambda_i \in \mathbb{R}.$$

Das letzte Korollar folgt, indem man $V = \mathbb{R}^n$ beziehungsweise $\mathbb{C}^n$ setzt und für $< \, , \, >$ die Standardpaarung wählt

$$< x, y > = x^t \cdot y = \sum_{i=1}^{n} \overline{x}_i y_i \ .$$

Die M zugeordnete Abbildung ist dann bezüglich $< \, , \, >$ selbstadjungiert, und das Korollar folgt unmittelbar aus dem obigen Satz.

53 Hauptachsentransformation

Satz 1: *Sei $S = S' \in M_{n,n}(\mathbb{R})$ eine symmetrische Matrix. Dann existiert eine orthogonale Matrix $T \in O(n, \mathbb{R})$, so daß gilt*

$$TST^{-1} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad , \quad T'T = TT' = E$$

für $\lambda_1 \geq ... \geq \lambda_n$ aus $\mathbb{R}$. Die Zahlen λ_i sind die Eigenwerte der Matrix S.

Zusatz: *Wegen $T' = T^{-1}$ für $T \in O(n, \mathbb{R})$ kann man auch schreiben*

$$TST' = TST^{-1} = Diag(\lambda_1, ..., \lambda_n) \ .$$

Insbesondere ist S definit dann und nur dann, wenn gilt $\lambda_n > 0$.

Beweis: Eine symmetrische Matrix S ist selbstadjungiert bezüglich der Standardform $< \, , \, >$ des $\mathbb{R}^n$. Der Spektralsatz für die Standardform $< .,. >$ zeigt, daß somit eine Orthogonalbasis $b_1, ..., b_n$ des $\mathbb{R}^n$ existiert, welche aus Eigenvektoren von S besteht

$$S \cdot b_i = \lambda_i \cdot b_i.$$

Sei $e_1, ..., e_n$ die Standardbasis des $\mathbb{R}^n$ und T die eindeutig bestimmte lineare Basiswechsel Abbildung mit

$$T^{-1}(e_i) = b_i \qquad i = 1, ..., n \ .$$

Dann gilt $TST^{-1}(e_i) = \lambda_i \cdot e_i$ für alle $i = 1, ..., n$. In Matrixsprechweise

$$TST^{-1} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} .$$

Zu zeigen ist, daß T^{-1} (und damit auch T) orthogonal sind:

$$< T^{-1}v, T^{-1}w > = < v, w > \qquad \forall v, w.$$

Dazu genügt wegen der Äquivalenz (iii) $\Leftrightarrow$ (ii) aus §45 etwas schwächer

$$< T^{-1}(e_i), T^{-1}(e_j) > = < e_i, e_j > = \delta_{ij} \qquad \forall i, j = 1, .., n.$$

Dies ist aber klar nach Definition von T wegen

$$< T^{-1}(e_i), T^{-1}(e_j) > = < b_i, b_j >= \delta_{ij}$$

da nach dem Spektralsatz die b_i eine Orthonormalbasis bilden $< b_i, b_j > = \delta_{ij}$. Dies beweist den Satz.

Eine geometrische Deutung: Die Matrix T^{-1} führt die orthonormale Standardbasis $e_1, ..., e_n$ in die orthonormale Basis $b_1, ..., b_n$ über. Daher sind T und T^{-1} orthogonale Matrizen. (Siehe §45 die Äquivalenz von (i)-(iii).)

Der Name Hauptachsentransformation erklärt sich wie folgt:

Sei $\mathbf{E}$ der **Kegelschnitt**

$$\{x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x' \cdot S \cdot x = 1\}$$

und $\mathbf{E}^{\#}$ der Kegelschnitt

$$\{x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x' \cdot Diag(\lambda_1, ..., \lambda_n) \cdot x = 1\}.$$

Dann gilt $T(\mathbf{E}) = \mathbf{E}^{\#}$ bezw. $\mathbf{E} = M(\mathbf{E}^{\#})$, denn

$$\begin{array}{rcc} x \in \mathbf{E}^{\#} & \Leftrightarrow & x' \cdot Diag(\lambda_1, ..., \lambda_n) \cdot x = 1 \\ & \Leftrightarrow & x' \cdot M' \cdot S \cdot M \cdot x = 1 \\ & \Leftrightarrow & (Mx)' \cdot S \cdot (Mx) = 1 \\ & \Leftrightarrow & Mx \in \mathbf{E} \end{array}$$

wegen $M^{-1}SM = Diag(\lambda_1, ..., \lambda_n)$ ist für $M = T^{-1} \in O(n, \mathbb{R})$.

Die Winkel und Längen erhaltende orthogonale Abbildung $T : \mathbb{R}^n \to \mathbb{R}^n$ bildet also den Kegelschnitt $\mathbf{E}$ bijektiv auf den Kegelschnitt $\mathbf{E}^{\#}$ ab und umgekehrt. Während sich $\mathbf{E}$ in schiefer Lage befindet

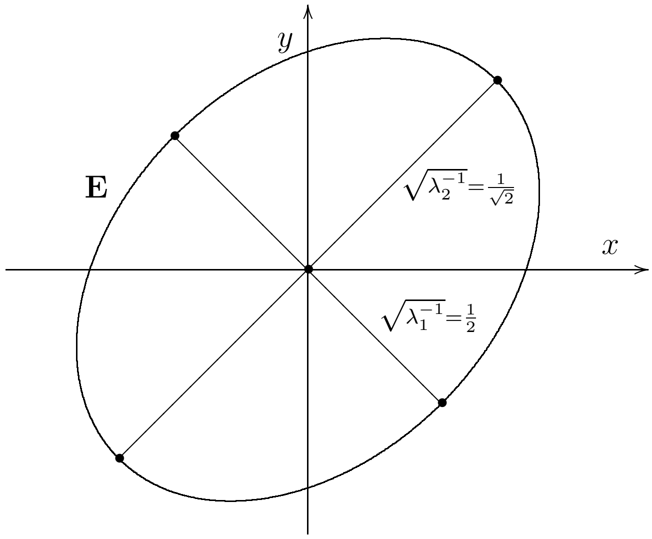
$$S_{11}x_1^2 + 2S_{12}x_1x_2 + \cdots + 2S_{nn-1}x_{n-1}x_n + S_{nn}x_n^2 = 1 \ ,$$

liegt der Kegelschnitt $\mathbf{E}^{\#}$ achsenparallel und ist gegeben durch die Gleichung

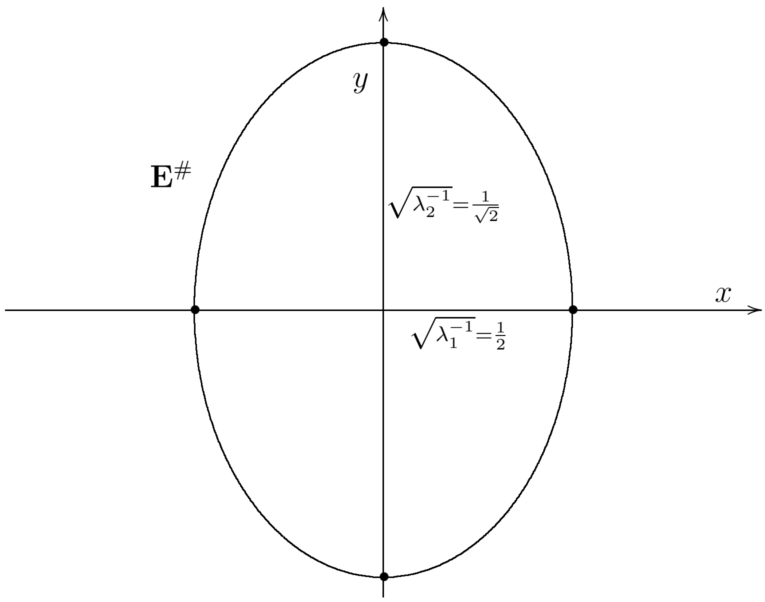
$$\lambda_1 x_1^2 \cdots + \lambda_n x_n^2 = 1.$$

Im definiten Fall $\lambda_1, .., \lambda_n > 0$ ist dies die Gleichung eines **Ellipsoids** im $\mathbb{R}^n$.

Beispiel: Für $S := \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$ ist $M = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$



Die Drehung M um $45°$ liefert $\mathbf{E}^\# : 4x^2 + 2y^2 = 1$:



Hierbei ist $\mathbf{E}$ die Ellipse $3x^2 - 2xy + 3y^2 = 1$. Die Hauptachsen haben die Längen $\sqrt{\lambda_1^{-1}} = \frac{1}{2}$ und $\sqrt{\lambda_2^{-1}} = \frac{1}{\sqrt{2}}$. Es gilt $\lambda_1 = 4$ und $\lambda_2 = 2$. $\mathbf{E}^\#$ geht durch Achsenstreckung aus der Sphäre $\sum x_i^2 = 1$ hervor.

Der Vollständigkeit halber formulieren wir Varianten des obigen Satzes. Die Beweise sind Modifikationen des obigen Argumentes und bleiben dem Leser überlassen.

Satz 2: *Sei $H = H^t \in M_{n,n}(\mathbb{C})$ eine hermitesche Matrix. Dann existiert eine sogenannte unitäre Transformation $T^t = T^{-1} \in Gl(n, \mathbb{C})$, so daß gilt*

$$T \cdot H \cdot T^{-1} = T \cdot H \cdot T^t = Diag(\lambda_1, ..., \lambda_n)$$

mit reellen Zahlen $\lambda_1 \geq ... \geq \lambda_n$. Die Zahlen λ_i sind die Eigenwerte der Matrix H. Schließlich ist H positiv definit genau dann wenn gilt $\lambda_n > 0$.

Eine weitere Verfeinerung dieses Satzes (mit Ausnahme der Definitheitsaussage) zeigt, daß es ausreicht H normal anzunehmen, d.h. $H \cdot H^t = H^t \cdot H$. In der Tat führt man diese Verschärfung durch einen Trick auf den obigen Satz 2 zurück. Man zerlegt dazu die normale Matrix H in eine Summe $H = H_1 + i \cdot H_2$ mit kommutierenden hermiteschen Matrizen H_1 und H_2. Siehe Beispiel 4) und 5) von §51. Man diagonalisiert dann zuerst H_1 durch einen unitären Basiswechsel (Satz 2). Da H_2 mit H_1 kommutiert, bildet H_2 die Eigenräume von H_1 auf sich ab. Wendet man den obigen Satz dann für die Einschränkungen von H_2 auf die Eigenräume von H_1 an, erhält man die gewünschte Verallgemeinerung.

Hermitesche Sylvestertypen: Zum Abschluß betrachten wir die Frage, wann zwei hermitesche Matrizen H_1 und H_2 aus $Gl(n, \mathbb{C})$ durch eine Matrix $M \in Gl(n, \mathbb{C})$ via $H_1 = M \cdot H_2 \cdot M^t$ ineinander übergeführt werden können. Satz 2 zeigt, daß $H = H^t \in Gl(n, \mathbb{C})$ bereits mittels einer unitären Matrix M in die reelle Diagonalmatrix $Diag(\lambda_1, ..., \lambda_n)$ übergeführt werden kann. Für allgemeine $M \in Gl(n, \mathbb{C})$ kann dann sogar

$$M \cdot H \cdot M^t \ = \ Diag(1, 1, ..., 1, -1, -1, ..., -1)$$

erreicht werden. Somit ist der Sylvestertyp der hermiteschen Matrix H gegeben durch die Zahl der Einträge $+1$ resp. -1, also durch die Zahl der positiven respektive negativen Eigenwerte λ_i der Matrix H. Wie im Paragraph §46 zeigt man dann

Satz 3: *Zwei hermitesche invertierbare Matrizen H_1 und H_2 lassen sich genau dann durch eine Matrix $M \in Gl(n, \mathbb{C})$ mittels $H_1 = M \cdot H_2 \cdot M^t$ ineinander überführen, wenn die hermiteschen Sylvestertypen der Matrizen H_1 und H_2 übererreinstimmen.*

54 Beweis des Spektralsatzes

In diesem Abschnitt seien folgende Daten fixiert

- Ein K-Vektorraum V der Dimension n für $K = \mathbb{R}$ oder $\mathbb{C}$.
- Eine positiv definite hermitesche Sesquilinearform $< .,. >$ auf V.
- Ein bezüglich $< v, w >$ selbstadjungierter Endomorphismus $\varphi : V \to V$ mit hermitescher Sesquilinearform $< v, w >_1 = < M(v), w >$.

Die Hilfsfunktion f: Zum Beweis des Spektralsatzes betrachten wir die folgende reellwertige Hilfsfunktion

$$f : V \setminus \{0\} \quad \longrightarrow \quad \mathbb{R}$$

definiert durch (siehe §49 Lemma 1)

$$v \quad \mapsto \quad \frac{< M(v), v >}{< v, v >} \; .$$

Da $< .,. >$ definit ist, wird der Nenner auf dem Definitionsbereich nie Null. Da Zähler und Nenner von $f(v)$ sesquilinear sind, folgt außerdem

$$f(\lambda \cdot v) \; = \; f(v) \quad , \quad \lambda \in K^* \; .$$

Die Funktion $f(v)$ ist somit konstant auf den Geraden durch Null!

Eigenvektoren: Sei λ ein Eigenwert von M zum Eigenvektor $v_0 \neq 0$ aus V. Dann gilt

$$f(v_0) = \frac{< M(v), v >}{< v, v >} = \frac{< \lambda \cdot v, v >}{< v, v >} = \overline{\lambda} \; .$$

Weil f reellwertig ist, ist somit insbesondere auch der Eigenwert $\lambda = \overline{\lambda}$ reell.

Folgerung: *Alle Eigenwerte λ von M sind reell.*

Der Beweis des Spektralsatz benutzt die folgenden drei Aussagen.

1) Die Funktion $f(v)$ nimmt in einem Punkt $v_0 \in V \setminus \{0\}$ ihr Maximum an.

2) Jeder solche Extremwert $v_0 \in V \setminus \{0\}$ von $f(v)$ ist ein Eigenvektor des selbstadjungierten Endomorphismus M mit dem Eigenwert $\lambda = f(v_0)$.

3) Für v_0 wie in Punkt 2) gibt es eine direkte Zerlegung
$$V = K \cdot v_0 \oplus v_0^\perp \ ,$$
wobei M die Unterräume Kv_0 und $v_0^\perp$ jeweils in sich abbildet.

Wir nehmen an 1) und 2) wären bereits gezeigt. Wir zeigen dann 3):

Die direkte Summenzerlegung von V wird genau so gezeigt wie das Lemma von Paragraph §44. Wir können dabei v_0 so normieren, daß $< v_0, v_0 > = 1$ gilt.

Nach 2) ist v_0 ein Eigenvektor. Somit gilt $M(K \cdot v_0) \ \subset \ K \cdot v_0$. Aus
$$< M(v), v_0 > = < v, M^*(v_0) > = < v, M(v_0) > = \lambda < v, v_0 > = 0 \quad , \quad v \in v_o^\perp$$
folgt $M(v_0^\perp) \subset v_0^\perp$. Dies beweist Aussage 3).

Der K-Vektorraum $W = v_0^\perp$ hat kleinere Dimension als V. Die Einschränkung von $< .,. >$ auf den Unterraum W ist wieder eine positiv definite hermitesche Sesquilinearform. Die Einschränkung des Endomorphismus M auf W ist wegen 3) erklärt (und wieder selbstadjungiert). Per Induktion gilt dann der Spektralsatz auf W und liefert dort die gesuchte Orthonormalbasis $b_2, .., b_n$. Ergänzt man um $b_1 = v_0$ erhält man die gesuchte Orthogonalbasis des Spetralsatzes auf V. Dies zeigt den Spektralsatz.

Zum Beweis verbleibt es die Eigenschaften 1) und 2) zu verifizieren: Wir beginnen mit Eigenschaft 2).

Sei $v_0 \neq 0$ aus V ein Extremalpunkt von f, in dem das Maximum angenommen wird. Für einen beliebigen Richtungsvektor $v \neq 0$ aus V ist dann $t \mapsto g(t) = f(v_0 + t \cdot v)$ eine reellwertige Funktion des Parameters $t \in \mathbb{R}$. Es gibt ein offenes Intervall $(-\varepsilon, \varepsilon)$ in $\mathbb{R}$ um Null, derart daß $v_0 + t \cdot v$ für $|t| < \varepsilon$ im Definitionsbereich $V \setminus \{0\}$ der Funktion f liegt!

Als Funktion von t hat $g(t) = f(v_0 + t \cdot v)$ ein Maximum bei $t = 0$. Andererseits ist

$$g(t) \;=\; \frac{t^2 < Mv, v > +2tRe(< Mv, v_0 > + < Mv_0, v_0 >}{t^2 < v, v > +2tRe(< v, v_0 > + < v_0, v_0 >}$$

als Quotient quadratischer Polynome in t differenzierbar im Intervall $(-\varepsilon, \varepsilon)$. Somit ist die folgende Ableitung im Punkt $t = 0$ gleich Null

$$\frac{d}{dt} g(t)|_{t=0} \;=\; \frac{2Re(< Mv, v_0 >) < v_0, v_0 > - < Mv, v_0 > 2Re(< v, v_0 >)}{< v_0, v_0 >^2} \;.$$

Dies liefert die Gleichung

$$Re(< Mv, v_0 >) < v_0, v_0 > \;=\; < Mv_0, v_0 > Re(< v, v_0 >)$$

oder wegen $< v_0, v_0 >= 1$ und $M = M^*$ die Gleichung

$$Re(< v, M(v_0) - \lambda_0 \cdot v_0 >) = 0 \quad , \quad \lambda_0 = \frac{< Mv_0, v_0 >}{< v_0, v_0 >} \;.$$

Die Form $< ., . >$ ist nach Annahme nicht ausgeartet. Da $v \neq 0$ beliebig gewählt werden kann (ersetze insbesondere v durch $i \cdot v$ im Fall $K = \mathbb{C}$), impliziert dies die gewünschte Eigenwertgleichung

$$M(v_0) = \lambda_0 \cdot v_0 \;.$$

Dies zeigt die Behauptung 2).

Bemerkung: *Der so konstruierte Eigenwert λ_0 ist der maximale Eigenwert von M.*

Es verbleibt der Beweis von Aussage 1):

Wir benötigen folgende Fakten aus der Analysis mehrerer Veränderlicher

(I) Eine stetige Funktion nimmt auf einer abgeschlossenen, beschränkten Teilmenge X von $\mathbb{R}^n$ ihr Maximum an.

(II) Polynome auf $\mathbb{R}^n$ sind stetig und Quotienten stetiger Funktionen sind stetig, falls der Nenner keine Nullstellen besitzt.

(III) Ist $h : \mathbb{R}^n \to \mathbb{R}$ eine stetige Funktion, dann ist $X = \{v \in \mathbb{R}^n \mid h(v) = 1\}$ abgeschlossen in $\mathbb{R}^n$.

Zur Erinnerung: Eine Teilmenge des $\mathbb{R}^n$ heißt beschränkt, wenn sie in einem Quader $\prod_{i=1}^n [-R, R]$ für ein geeignetes $R \in \mathbb{R}$ enthalten ist. Eine Teilmenge $A \subset \mathbb{R}^n$ heißt (unter Grenzwertbildung) abgeschlossen, wenn für jede (komponentenweise) in $\mathbb{R}^n$ konvergente Folge $v_k \in A$ der Grenzwert $v \in \mathbb{R}^n$ in A liegt. Eine Funktion $f : U \to \mathbb{R}$ auf einer Teilmenge U heißt stetig, wenn f stetig in allen Punkten v von U ist. f heißt stetig im Punkt $v \in U$ wenn gilt

$$\forall \varepsilon > 0 \; \exists \delta > 0 \quad \|w - v\| < \delta, w \in U \Rightarrow \|f(w) - f(v)\| < \varepsilon \; .$$

(Hierbei ist $\|.\|$ irgend eine Norm auf $\mathbb{R}^n$).

Um obige Sätze der Analysis anwenden zu können, identifizieren wir V mit $\mathbb{R}^n$ (bzw. $\mathbb{R}^{2n}$ im komplexen Fall). Dann ist

$$X \;=\; \{ v \in V \mid \; < v, v >= 1 \; \}$$

beschränkt und abgeschlossen (nach III). Wir begründen dies wie folgt

Zur Beschränkheit: Sei der Einfachheit halber $< .,. >$ die sesquilineare Standardform auf K^n. Dann wählen wir eine Orthonormalbasis $e_1, ..., e_n$ von $< \, , \, >$ auf V (§46) und identifizieren jeden Vektor $v = \sum\limits_{i=1}^{n} x_i e_i$ aus V mit dem Spaltenvektor $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ im $\mathbb{R}^n$. Bezüglich dieser Basiswahl ist

$$X \;=\; \{ v \in V \mid \; < v, v > = 1 \} \;=\; \{ v \in \mathbb{R}^n \mid \; \sum_{i=1}^{n} |x_i|^2 = 1 \}$$

offensichtlich im Quader $\{ v \in \mathbb{R}^n \mid |x_i| \leq 1 \quad i = 1, ..., n \}$ enthalten, also beschränkt. X ist Urbild von $\{1\}$ unter der stetigen Abbildung (Analysisbox II)

$$h(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) \;=\; |x_1|^2 + ... + |x_n|^2 \; .$$

Also ist X abgeschlossen (Analysisbox III).

Aus (II) folgt die Stetigkeit von $f(v)$ auf X. Aus Aussage (I) folgt dann, daß die Funktion $f(v)$ auf X ihr Maximum in einem Punkt $v_0 \in X$ annimmt. Wir behaupten, daß $f(v_0)$ das Maximum der Hilfsfunktion $f(v)$ auf dem ganzen Definitionsbereich $V \setminus \{0\}$ ist. Benutze dazu, daß f konstant auf Geraden in V ist. Dies zeigt die Behauptung 1). Damit ist der Spektralsatz gezeigt – zumindestens im Fall wo $< .,. >$ die sesquilineare Standardform auf K^n ist.

Dieser Spezialfall des Spektralsatzes genügt um den Hauptachsentransformation zu beweisen. Siehe Satz 2 des letzten Paragraphen.

Durch Hauptachsentransformation von H (siehe letzter Paragraph) zeigt man dann, daß für beliebige definite hermitesche Sesquilinearformen $< v, w >= v^t H w$ der Ellipsoid $\mathbf{E}$ aller Punkte $v \in K^n$ mit $< v, v > \leq 1$ beschränkt in K^n ist. Bis auf eine orthogonale resp. unitäre Transformation – welche beschränkte Mengen in beschränkte Mengen überführt! – handelt es sich nämlich um die gestauchte Einheitssphäre $\mathbf{E}^{\#}$. Diese ist offensichtlich beschränkt.

Dieses Argument schließt die noch verbliebene Lücke des Arguments im Fall beliebiger positiv definiter hermitescher Sesquilinearformen auf V. Damit ist der Spektralsatz in voller Allgemeinheit gezeigt.

55 Hermitesche Matrizen und Lorentzgruppe

Punkte $v = (t, x, y, z)$ des reellen Vektorraums $\mathbb{R}^4$ entsprechen Ereignissen in der Raum-Zeit. In der speziellen Relativitätstheorie spielen dann lineare Koordinatenwechsel der Raum-Zeit eine Rolle, welche die indefinite quadratische Form

$$q(v) = t^2 - x^2 - y^2 - z^2$$

festlassen. Die dadurch definierte Lorentzgruppe ist die orthogonale Gruppe der quadratischen Form $S = diag(1, -1, -1, -1)$. Die Untergruppe der speziellen orthogonalen Transformationen nennt man die eigentliche Lorentzgruppe $\mathcal{L}$.

Bewegungen ohne Beschleunigung entsprechen Geraden im $\mathbb{R}^4$. Vom Nullpunkt ausgehende Bewegungen liegen aus physikalischen Gründen im positiven Lichtkegel $\mathcal{P}$ aller Punkte v mit Koordinaten $t > 0$ und

$$t^2 \ > \ r^2 = x^2 + y^2 + z^2 \ .$$

Die Geschwindigkeit der Bewegung – nämlich die Steigung $\beta_v = r/t$ der zugehörigen Gerade – ist hier nämlich kleiner als 1 (Lichtgeschwindigkeit). Analog hat man den rückwertigen Lichtkegel $\mathcal{N} = -\mathcal{P}$. Der eigentliche Lichtkegel $\mathcal{C}$ ist die Menge der Punkte $v \in \mathbb{R}^4$ mit $q(v) = 0$. Die verbleibenden Punkte bilden eine Teilmenge $\mathcal{I}$ des Vektorraums $\mathbb{R}^4$.

Die vom Basisvektor $\sigma_t = (1, 0, 0, 0)$ aufgespannte Gerade beschreibt einen im Ursprung ruhenden 'Bewegung' der Geschwindigkeit Null. Das Orthokomplement $\sigma_t^\perp$ des Vektors σ_t bezüglich der Form S ist der dreidimensionale Euklidsche Raum $\mathcal{H}^0 = \mathbb{R}^3$, welcher von den drei verbleidenden Basisvektoren $\sigma_x, \sigma_y, \sigma_z$ aufgespannt wird. Sei darin $\mathcal{H}^{00}$ der von σ_x, σ_y aufgespannte Teilraum.

Die Abbildungen der (eigentlichen) Lorentzgruppe erhalten die quadratische Form q und bilden daher die Teilmengen $\mathcal{P} \cup -\mathcal{P}$ sowie $\mathcal{C}$ und $\mathcal{I}$ jeweils bijektiv auf sich ab. Zum Beispiel vertauscht die Raum-Zeit Spiegelung $v \mapsto -v$ (eine Element der eigentlichen Lorentzgruppe) die Kegel $\mathcal{P}$ und $\mathcal{N} = -\mathcal{P}$. Ordnet man jeder eigentlichen Lorentzsubstitution ein Vorzeichen ± 1 zu, je nach dem ob $\mathcal{P}$ und $-\mathcal{P}$ vertauscht werden oder nicht, so definiert dies einen surjektiven Gruppenhomomorphismus $N: \ \mathcal{L} \to \{\pm 1\}$, die sogenannte Spinornorm N (siehe Appendix §50) mit Kern $\mathcal{L}^0$ (dem Spinorkern). Offensichtlich gilt $\mathcal{L} = \mathcal{L}^0 \cup -\mathcal{L}^0$.

Wir wollen zeigen, daß die Gruppe $\mathcal{L}^0$ im wesentlichen die Gruppe $Sl(2,\mathbb{C})$ ist. Genauer gesagt gilt

Satz: *Es gilt $\mathcal{L} = \mathcal{L}^0 \cup -\mathcal{L}^0$ und es gibt einen surjektiven Homorphismus von Gruppen*

$$\pi: \; Sl(2,\mathbb{C}) \; \to \; \mathcal{L}^0 \; ,$$

dessen Kern die Untergruppe der Matrizen $\{\pm E\}$ von $Sl(2,\mathbb{C})$ ist.

Bemerkung: π definiert also eine zweiblättrige Überlagerung von $\mathcal{L}^0$!

Zum Beweis identifizieren wir den $\mathbb{R}^4$ mit dem $\mathbb{R}$-Vektorraum $\mathcal{H}$ der hermitesch komplexen 2×2-Matrizen $H = H^t \in M_{2,2,}(\mathbb{C})$ vermöge der Zuordnung von Basisvektoren

$$\sigma_t \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \; , \; \sigma_z \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \; , \; \sigma_x \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \; , \; \sigma_z \mapsto \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \; .$$

Mit anderen Worten der Vektor $v = (t,x,y,z)$ entspricht der hermiteschen komplexen Matrix

$$H = \begin{pmatrix} t+z & x+iy \\ x-iy & t-z \end{pmatrix} \; .$$

Jede hermitesche Matrix H läßt sich eindeutig auf diese Weise schreiben. In der Sprache der hermiteschen Matrizen ist die quadratische Form q durch die Determinante der Matrix H

$$det(H) = t^2 - x^2 - y^2 - z^2$$

gegeben.

Der Lichtkegel $\mathcal{C}$ beschreibt also die Matrizen $H \in \mathcal{H}$ der Determinante Null. Die Zerlegung des Komplements $\mathcal{H}\backslash\mathcal{C}$ in die Teilmengen $\mathcal{P}$, $\mathcal{I}$, $\mathcal{N}$ wird dann jeweils durch den hermiteschen Sylvestertyp der invertierbaren Matrix $H \in \mathcal{H}\backslash\mathcal{C}$ beschrieben. In Termen der Eigenwerte $\lambda_1 \geq \lambda_2$ der hermiteschen Matrix H, welche durch $\lambda_{1/2} = t \pm r$ gegeben sind, bedeutet $H \in \mathcal{P}, \mathcal{I}, \mathcal{N}$ jeweils $\lambda_2 > 0$ resp. $\lambda_1 > 0 > \lambda_2$ resp. $0 > \lambda_1$. Siehe dazu §53.

Weiterhin ist $H \in \mathcal{H}^0$ (d.h. H liegt im 'Raum' $\mathbb{R}^3$) gleichbedeutend mit $Spur(H) = 0$ und $H \in \mathcal{H}^{00}$ (d.h. H liegt in der xy-Ebene) gleichbedeutend mit dem Verschwinden der Diagonale von H.

Konstruktion von π: Für komplexe Matrizen $M \in Gl(2,\mathbb{C})$ definieren wir

$$\pi(M) : H \mapsto MHM^t \quad , \quad H \in \mathcal{H} \ .$$

Offensichtlich ist $\pi(M)$ eine $\mathbb{R}$-lineare Abbildung

$$\pi(M) : \mathcal{H} \to \mathcal{H} \ .$$

Es gilt $\pi(M_1 \cdot M_2) = \pi(M_1) \circ \pi(M_2)$ und $\pi(E) = id_{\mathcal{H}}$. Somit ist π ein Gruppenhomomorphismus. Insbesondere ist $\pi(M^{-1})$ die Umkehrabbildung von $\pi(M)$.

Wegen $det(MHM^t) = |det(M)|^2 \cdot det(H)$ ist somit $\pi(M)$ für $M \in Sl(2,\mathbb{C})$ eine orthogonale Abbildung von $\mathcal{H}$ bezüglich der quadratischen Form $q(.) = det(.)$. $\pi(M)$ liegt dann bereits in der eigentlichen Lorentzgruppe. Es genügt dies für Erzeuger von $Sl(2,\mathbb{C})$ – etwa Elementarmatrizen – zu zeigen. Die Abbildung $\pi : Sl(2,\mathbb{C}) \to \mathcal{L}^0$ ist damit konstruiert. Offensichtlich lässt $\pi(M)$ den Vektor $\sigma_t \in \mathcal{H}$ genau dann fest, wenn M unitär ist: $MM^t = E$.

Zum Beweis des Satzes: Wir überlassen es dem Leser $Kern(\pi) = \{\pm E\}$ zu zeigen. Wir skizzieren aber die Surjektivität der Abbildung $\pi : Sl(2,\mathbb{C}) \to \mathcal{L}^0$. Benutze

a) Für $H \in \mathcal{P}$ mit $det(H) = 1$ existiert $M \in Sl(2,\mathbb{C})$ mit $MHM^t = E$.

b) Für $H \in \mathcal{H}^0$ mit $det(H) = -1$ existiert eine unitäre Matrix $M \in Sl(2,\mathbb{C})$ mit $MHM^t = diag(1,-1)$.

Beide Aussagen folgen unmittelbar aus dem Spektralsatz für hermitesche Operatoren (Satz 2 und Satz 3 aus §53). Dies zu zeigen überlassen wir wieder dem Leser.

Surjektivität von π: Sei g eine Substitution aus $\mathcal{L}^0$. Diese bildet $\sigma_t \in \mathcal{P}$ auf eine Matrix H aus $\mathcal{P}$ ab. Es gilt $det(H) = det(\sigma_t)$, denn g erhält die quadratische Form $q(.) = det(.)$. Wegen a) kann g durch ein $\pi(M) \in \mathcal{L}^0$ so modifiziert werden, daß $H = E$ gilt. Das heißt obdA kann daher zum Beweis im folgenden $g(\sigma_t) = \sigma_t$ angenommen werden.

Im Fall $g(\sigma_t) = \sigma_t$ definiert g eine spezielle orthogonale Abbildung des Orthokomplements $\sigma_t^\perp = \mathcal{H}^0$, also ein Element der <u>Drehgruppe</u> $SO(3,\mathbb{R})$.

Sei nun $\tilde{H} \in \mathcal{H}^0$ das Bild von $\sigma_z \in \mathcal{H}^0$ unter g. Dann gilt $det(\tilde{H}) = det(\sigma_z) = -1$. Nach b) kann g durch ein $\pi(M)$ (M unitär, $det(M) = 1$) modifiziert werden, so daß auch gilt $g(\sigma_z) = \sigma_z$. Damit ist g obdA eine spezielle orthogonale Abbildung der x,y-Ebene, also eine Drehung. Man zeigt aber leicht, daß jede Drehung der x,y-Ebene durch eine Substitution

$$\pi(M) \quad , \quad M = \begin{pmatrix} w & 0 \\ 0 & \overline{w} \end{pmatrix} \in Sl(2,\mathbb{C})$$

mit einer geeigneten komplexen Zahl $w \in C^*$ vom Betrag $w\overline{w} = 1$ realisiert wird. Daraus läßt sich ein beliebiges $g \in \mathcal{L}^0$ als Produkt von drei Substitutionen aus $Bild(\pi)$ schreiben. Da $Bild(\pi)$ eine Gruppe ist, folgt $\mathcal{L}^0 \subset Bild(\pi)$. Somit ist π ist surjektiv.

Der eben skizzierte Beweis zeigt zugleich

Korollar: *Die Einschränkung von π auf $SU(2,\mathbb{C}) \subset Sl(2,\mathbb{C})$ definiert einen <u>surjektiven</u> Gruppenhomomorphismus*

$$\pi : \ SU(2,\mathbb{C}) \to SO(3,\mathbb{R})$$

von der Gruppe $SU(2,\mathbb{C})$ der unitären komplexen 2×2-Matrizen mit Determinante 1 auf die dreidimensionale spezielle orthogonale Gruppe $SO(3,\mathbb{R})$ mit $Kern(\pi) = \{\pm E\}$.

Ringe und Moduln

56 Grundlegende Begriffe

Definition: Ein Ring $(R, +, \cdot)$ ist eine abelsche Gruppe bezüglich $+$ und besitzt eine Multiplikationsabbildung $R \times R \to R$, so daß gilt:

$$\begin{array}{llcll} (I) & (r \cdot s) \cdot t & = & r \cdot (s \cdot t) & \forall r, s, t \in R \\ (II) & (r + s) \cdot t & = & r \cdot t + s \cdot t & \forall r, s, t \in R \\ (III) & r \cdot (s + t) & = & r \cdot s + r \cdot t & \forall r, s, t \in R \end{array}$$

Wir fordern zusätzlich die Existenz eines neutralen Elements $1 = 1_R$ mit der Eigenschaft

$$(IV) \qquad r \cdot 1 = r = 1 \cdot r \qquad \forall r \in R.$$

Aus der geforderten Eigenschaft folgt sofort die Eindeutigkeit dieses neutralen Elementes 1. Wir fordern aber nicht $0 \neq 1$.

Achtung: Im Allgemeinen gilt $r \cdot s \neq s \cdot r$.

Ein Ring R heißt kommutativ, falls $r \cdot s = s \cdot r \qquad \forall r, s \in R$.

Beispiel:

1) Jeder Körper ist ein Ring.

2) $R := \mathbb{Z}$ ist ein Ring.

   Allgemein ist jede bezüglich der Multiplikation abgeschlossene additive Untergruppe eines Körpers ein Ring.

3) $R := M_{n,n}(K)$ mit Matrizenaddition und Matrizenmultiplikation ist ein Ring.

4) Ist R ein Ring, dann ist auch $M_{n,n}(R)$ ein Ring bezüglich der Matrizenaddition und Matrizenmultiplikation.

5) $R := \{0\}$ ist ein Ring.

Polynomringe: Ein wichtiges Beispiel sind die sogenannten Polynomringe.
Sei R ein gegebener Ring. Dann setzen wir

$$R[X] := \{(a_0, a_1, ...) \mid a_i \in R \; \forall i \in \mathbb{N}_0, \; fast \; alle \; a_i = 0\}.$$

Anstelle von $(a_0, a_1, ...)$ schreiben wir auch

$$\sum_{i \geq 0} a_i X^i.$$

Die Addition zweier Elemente ist definiert durch

$$(a_0, a_1, ...) + (b_0, b_1, ...) := (a_0 + b_0, a_1 + b_1, ...).$$

Die Multiplikation zweier Elemente ist definiert durch

$$(a_0, a_1, ...) \cdot (b_0, b_1, ...) := (a_0 \cdot b_0, a_0 \cdot b_1 + a_1 \cdot b_0, ...) = (c_0, c_1, ...),$$

wobei $c_n := \sum_{i=0}^{n} a_i b_{n-i}$.

Lemma 1: *$R[X]$ ist mit $+$ und $\cdot$ ein Ring.*

Beweis: Übungsaufgabe

Definition: Ein Ringhomomorphismus $\varphi : T \to S$ zwischen zwei Ringen R und S ist eine Abbildung mit den Eigenschaften

$$\begin{array}{lcl} \varphi(r_1 + r_2) & = & \varphi(r_1) + \varphi(r_2) \\ \varphi(r_1 \cdot r_2) & = & \varphi(r_1) \cdot \varphi(r_2) \\ \varphi(1_R) & = & 1_S \end{array}$$

Definition: Die Einheiten R^* eines Ringes R sind die Elemente $r \in R$ mit der Eigenschaft:

$$r \in R^* \iff \exists s \in R \; mit \; s \cdot r = 1 = r \cdot s.$$

Beispiel:

1) Sei $R = K$ ein Körper. $K^* = K \setminus \{0\}$.
2) $R = \mathbb{Z}$ $\quad \mathbb{Z}^* = \{1, -1\}$
3) Sei $R = K[X]$ der Polynomring über einem Körper. $R^* = \{(a_0, 0, 0, ...) \mid a_0 \in K^*\}$
4) $R = M_{n,n}(K)$ $\quad R^* = \mathcal{G}\ell(n, K)$

Lemma 2: *R^* ist eine Gruppe bezüglich Multiplikation (die Einheitengrupppe).*

Beweis: $1 \in R^*$ wegen $1 \cdot 1 = 1$.
Nach Definition hat jedes Element $r_i \in R^*$ ein zugehöriges Element $s_i - \in R$ mit $s_i \cdot r_i = 1 = r_i \cdot s_i$. Es folgt $r_1, r_2 \in R^* \Rightarrow r := r_1 \cdot r_2 \in R^*$. Setze $s := s_2 \cdot s_1$.
Dann folgt

$$s \cdot r = (s_2 \cdot s_1) \cdot (r_1 \cdot r_2) = s_2 \cdot (s_1 \cdot r_1) \cdot r_2 = s_2 \cdot 1 \cdot r_2 = s_2 \cdot r_2 = 1$$

$$r \cdot s = (r_1 \cdot r_2) \cdot (s_2 \cdot s_1) = r_1 \cdot (r_2 \cdot s_2) \cdot s_1 = r_1 \cdot 1 \cdot s_1 = r_1 \cdot s_1 = 1$$

Für $r \in R^*$ mit $s \cdot r = 1 = r \cdot s$ ist somit auch $s \in R^*$. s ist inverses Element von R in R^*.

57 R-Moduln

Definition: Sei R ein Ring. Ein R-Modul ist eine abelsche Gruppe $(M, +)$ versehen mit einer zusätzlichen skalaren Multiplikation

$$\begin{array}{lcl} R \times M & \to & M \\ (r, m) & \mapsto & r \cdot m = rm, \end{array}$$

so daß für alle $m, m_1, m_2 \in M,\ r, s \in R$ gilt:

$$\begin{array}{llcl} (I) & (r \cdot s) \cdot m & = & r \cdot (s \cdot m) \\ (II) & (r + s) \cdot m & = & r \cdot m + s \cdot m \\ (III) & r \cdot (m_1 + m_2) & = & r \cdot m_1 + r \cdot m_2 \\ (IV) & 1 \cdot m & = & m. \end{array}$$

Es handelt sich hier um R-Linksmoduln, da die skalare Multiplikation links erfolgt.

Beispiel:

1) Ist $R = K$ ein Körper, dann sind per Definition R-Moduln dasselbe wie K–Vektorräume.

2) Der freie Modul R^n vom Rang n :

   $R^n := \{(r_1, ..., r_n) \mid r_i \in R\}$ mit

$$\begin{array}{rcl} (r_1, ..., r_n) + (r'_1, ..., r'_n) & = & (r_1 + r'_1, ..., r_n + r'_n) \\ r \cdot (r_1, ..., r_n) & = & (r \cdot r_1, ..., r \cdot r_n) \end{array}$$

Eine R-lineare Abbildung $\varphi : M \to N$ zwischen zwei R-Moduln M, N ist eine Abbildung mit der Eigenschaft

$$\varphi(r_1 m_1 + r_2 m_2) = r_1 \varphi(m_1) + r_2 \varphi(m_2).$$

Die Menge aller solcher R-linearen Abbildungen sei $Hom_R(M, N)$.

Bemerkung: Ist speziell R ein kommutativer Ring dann, dann definiert $Hom_R(M, N)$ bezüglich der nachfolgende definierten punktweisen Addition und Skalarmultiplikation wieder einen R-Modul

$$(r\varphi + s\psi)(m) := r\varphi(m) + s\psi(m)\ .$$

Sind nämlich φ, ψ R-linear, dann ist auch die so definierte Abbildung $r\varphi + s\psi$ wieder R-linear. Dies benutzt die Kommutativität des Rings R

$$\begin{aligned}(r\varphi + s\psi)(r_1 m_1 + r_2 m_2) &= r\varphi(r_1 m_1 + r_2 m_2) + s\psi(r_1 m_1 + r_2 m_2)\\ &= rr_1\varphi(m_1) + rr_2\varphi(m_2) + sr_1\psi(m_1) + sr_2\psi(m_2) = a.\end{aligned}$$

Die zweite Gleichheit folgt aus $\varphi, \psi \in Hom_R(M, N)$. Wir hätten gerne

$$\begin{aligned}(r\varphi + s\psi)(r_1 m_1 + r_2 m_2) &= r_1(r\varphi + s\psi)(m_1) + r_2(r\varphi + s\psi)(m_2)\\ &= r_1(r\varphi(m_1) + s\psi(m_1)) + r_2(r\varphi(m_2) + s\psi(m_2))\\ &\overset{(III)}{=} r_1 r\varphi(m_1) + r_1 s\psi(m_1) + r_2 r\varphi(m_2) + r_2 s\psi(m_2) = b.\end{aligned}$$

Die Behauptung $a = b$ folgt, da R kommutativ angenommen worden war. Für nicht kommutatives R ist im allgemeinen $Hom_R(M, N)$ kein R-Modul.

Bemerkung: Folgende Eigenschaften und Definitionen lassen sich unmittelbar vom Vektorraumfall ($R = K$ ein Körper) auf den Fall beliebiger R-Moduln übertragen:

1) Die Komposition R-linearer Abbildungen ist R-linear.

2) Sei M ein R-Modul. Ein R-Untermodul $N \subseteq M$ ist eine abelsche Untergruppe von $(M, +)$ mit

$$r \cdot N \in N \qquad \forall n \in N, \quad \forall r \in R.$$

Insbesondere ist jeder Untermodul ein R-Modul.

3) Kern und Bild jeder R-linearen Abbildung $\varphi : M \to N$ sind R-Untermoduln (und somit insbesondere R-Moduln) von M resp. N. Hierbei ist wie bisher

$$Kern(\varphi) = \{m \in M \mid \varphi(m) = 0\}$$

und

$$Bild(\varphi) = \{\varphi(m) \mid m \in M\}.$$

4) Eine R-lineare Abbildung φ heißt Isomorphismus (von R-Moduln), falls sie ϕ bijektiv ist.

5) Eine R-lineare Abbildung φ ist injektiv genau dann wenn gilt $Kern(\varphi) = 0$.

6) Sei $\varphi : M \to N$ ein Isomorphismus von R-Moduln, dann ist auch φ^{-1} ein Isomorphismus, d.h. ϕ^{-1} ist automatisch wieder R-linear.

<u>Quotientenmoduln</u>: Sei M ein R-Modul und $N \subseteq M$ ein R-Untermodul. Unter einen Quotienten R-Modul von $N \subset M$ verstehen wir ein Paar, bestehend aus einem R-Modul M/N zusammen mit einer R-linearen Abbildung

$$\pi : M \to M/N$$

mit folgenden Eigenschaften

(1) π ist surjektiv.

(2) $Kern(\pi) = N$

<u>Existenz von Quotienten</u>: Die Existenz eines solchen Quotientenmoduls zeigt man wie im Fall der Vektorräume in §15.

Setze hierzu $M/N = \{[m] \mid m \ \in M\}$. Hierbei sei

$$[m] = m + N = \{m + n \mid n \in N\} \ \subseteq \ M.$$

Dann gilt für beliebige $m, m' \in M$:

$$[m] \cap [m'] = \begin{cases} \emptyset & oder \\ [m] \end{cases}$$

M/N ist die Menge aller 'Parallelen' $[m]$ von N in M. Die R-Modulstruktur auf M/N wird erklärt durch

$$r[m] + s[m'] := [rm + sm'].$$

Dies ist wohldefiniert! Definiere dann die Abbildung $\pi : M \to M/N$ durch

$$\pi(m) = [m] \ .$$

Per Konstruktion von M/N ist diese Abbildung π surjektiv und R-linear.

58 Exakte Sequenzen und Komplexe

Eine Sequenz von R-linearen Abbildungen

$$\cdots \longrightarrow K^{-1} \xrightarrow{d_{-1}} K^0 \xrightarrow{d_0} K^1 \xrightarrow{d_1} K^2 \longrightarrow \cdots$$

von R-Moduln K^i $(i \in \mathbb{Z})$ heißt ein **Komplex von R-Moduln**, falls für alle Indices $i \in \mathbb{Z}$ gilt $d_{i+1} \cdot d_i = 0$, d.h.

$$Bild(d_i) \subseteq Kern(d_{i+1}).$$

Definition: Eine Sequenz von R-linearen Abbildungen

$$K' \xrightarrow{\varphi} K \xrightarrow{\psi} K''$$

von R-Moduln heißt exakt (an der Stelle K), wenn gilt

$$Bild(\varphi) = Kern(\psi).$$

Insbesondere gilt dann $\psi \circ \varphi = 0$.

Definition: Ein $Komplex$ heißt exakt, falls er an jeder Stelle exakt ist, d.h. wenn gilt:

$$Bild(d_i) = Kern(d_{i+1}) \qquad \forall i \in \mathbb{Z}.$$

Ein Synonym für exakte Komplexe ist der Begriff lange exakte Sequenz.

Eine kurze, exakte Sequenz von R-Moduln ist gegeben durch:

$$0 \longrightarrow K' \xrightarrow{\varphi} K \xrightarrow{\psi} K'' \longrightarrow 0,$$

wobei φ, ψ R-lineare Abbildungen von R-Moduln K', K, K'' sind, so daß Exaktheit an den drei Stellen K', K, K'' vorliegt.

Einige Erläuterungen zum Begriff der kurzen exakten Sequenz:

1) Eine Sequenz
$$0 \longrightarrow L \xrightarrow{\varphi} L''$$
ist exakt bei L genau dann, wenn φ injektiv ist, d.h. falls $Kern(\varphi) = 0$.

Beweis: Exaktheit bei $L \iff (Bild(0) = 0 = Kern(\varphi)) \iff (Kern(\varphi) = 0)$.

2) Eine Sequenz
$$L' \xrightarrow{\psi} L \longrightarrow 0$$
ist exakt bei L genau dann, wenn ψ surjektiv ist, d.h. falls $Bild(\psi) = L$.

Beweis: Exaktheit bei $L \iff (Bild(\psi) = Kern(0) = L) \iff (Bild(\psi) = L)$.

3) Die Sequenz
$$L' \longrightarrow 0 \longrightarrow K$$
ist immer exakt.

Beweis: trivial.

Behauptung: Gegeben sei eine Surjektion
$$V \overset{\varphi}{\twoheadrightarrow} W,$$
wobei φ R-linear und V, W R-Moduln sind. Dann ist offensichtlich
$$0 \longrightarrow Kern(\varphi) \overset{i}{\hookrightarrow} V \overset{\varphi}{\twoheadrightarrow} W \longrightarrow 0$$
kurz exakt wegen $Bild(i) = Kern(\varphi)$ und der Injektivität der Inklusion i.

Behauptung: Gegeben sei eine Injektion
$$U \overset{\psi}{\hookrightarrow} W,$$
wobei ψ R-linear und U, W R-Moduln sind. Dann ist offensichtlich
$$0 \longrightarrow U \overset{\psi}{\hookrightarrow} W \overset{\pi}{\twoheadrightarrow} W/\psi(U) \longrightarrow 0$$
kurz exakt wegen $Kern(\pi) = \psi(U) = Bild(U)$ und der Surjektivität von π.

59 Ein Fortsetzungssatz

Satz: *Gegeben seien exakte Sequenzen von R-Moduln*

$$\begin{tikzcd}
0 \arrow[r] & M' \arrow[r] \arrow[d, "\varphi"] & M \arrow[r, "d"] \arrow[d, "\psi"] & M'' \arrow[r] \arrow[d, dotted, "\exists!\ \tau"] & 0 \\
0 \arrow[r] & N' \arrow[r] & N \arrow[r, "\tilde{d}"] & N'' \arrow[r] & 0
\end{tikzcd}$$

sowie R-lineare Abbildungen φ, ψ, so daß das linke Quadrat des Diagramms kommutativ ist. Dann existiert <u>eine eindeutig bestimmte R-lineare Abbildung</u>

$$\tau : M'' \to N'' \ ,$$

welche das rechte Quadrat kommutativ macht.

<u>Beweis</u>: Als Ansatz für die gesuchte Abbildung τ nehmen wir

$$\begin{tikzcd}
m \arrow[r, mapsto, "d"] \arrow[d, mapsto, "\psi"] & m'' \arrow[d, dotted, mapsto, "\tau"] \\
\psi(m) \arrow[r, mapsto, "\tilde{d}"] & \tau(m'')
\end{tikzcd}$$

das heißt wir nehmen <u>irgendein</u> Urbild m von m'' – also $d(m) = m''$; ein solches m existiert, da $M \xrightarrow{d} M'' \longrightarrow 0$ exakt ist, somit d surjektiv – und setzen dann

$$\tau(m'') := \tilde{d}(\psi(m)) = (\tilde{d} \circ \psi)(m) \ .$$

Um zu zeigen, daß τ R-linear ist, muß man zuerst zeigen, daß τ nicht von der Wahl des Urbildes m abhängt. Dies wird im nächsten Hilfssatz gezeigt.

Hilfssatz : *$\tilde{d} \circ \psi(m)$ hängt nur ab von $d(m)$.*

<u>Beweis des Hilfssatzes</u>: Es genügt zu zeigen, daß $Kern(d) \subseteq M$ unter der Abbildung $\tilde{d} \circ \psi$ auf Null abgebildet wird.

$$\begin{tikzcd}
 & m' \arrow[r, mapsto] & m \arrow[r, mapsto] & 0 & \\
0 \arrow[r] & M' \arrow[r, "d"] \arrow[d, "\varphi"] & M \arrow[r, "d"] \arrow[d, "\psi"] & M'' \arrow[r] & 0 \\
0 \arrow[r] & N' \arrow[r, "\tilde{d}"] & N \arrow[r, "\tilde{d}"] & N'' \arrow[r] & 0 \\
 & \varphi(m') \arrow[r, mapsto] & \cdots \arrow[r, mapsto] & 0 &
\end{tikzcd}$$

Wegen der Exaktheit bei M impliziert $d(m) = 0$ die Existenz eines $m' \in M'$ mit $d(m') = m$. Wegen der Kommutativität des linken Digrammes gilt:

$$\tilde{d} \circ \psi(m) = \tilde{d} \circ (\psi \circ d)(m') \overset{!}{=} \tilde{d} \circ (\tilde{d} \circ \varphi)(m') \ .$$

Wegen der Exaktheit bei N gilt $\tilde{d} \circ \tilde{d} = 0$ und somit folgt

$$\tilde{d} \circ \psi(m) = 0 \ .$$

q.e.d.

<u>R-Linearität von τ</u>: Betrachte $m_1'', m_2'' \in M''$ und $r_1 \cdot m_1'' + r_2 \cdot m_2'' \in M''$. Wir wollen dann zeigen $\tau(r_1 \cdot m_1'' + r_2 \cdot m_2'') = r_1 \cdot \tau(m_1'') + r_2 \cdot \tau(m_2'')$.

Seien m_1, m_2 Urbilder unter d von m_1'', m_2''. Sei m_3 ein Urbild von $r_1 m_1'' + r_2 m_2''$

$$\begin{tikzcd}
m_1, m_2, m_3 \arrow[r, "d"] \arrow[d, "\psi"] & m_1'', m_2'', r_1 m_1'' + r_2 m_2'' \arrow[d, "\tau"] \\
\psi(m_1), \psi(m_2), \psi(m_3) \arrow[r, "\tilde{d}"] & \tilde{d}\psi(m_1), \tilde{d}\psi(m_2), \tilde{d}\psi(m_3)
\end{tikzcd}$$

Ein weiteres Urbild von $r_1 m_1'' + r_2 m_2''$ ist natürlich auch $r_1 m_1 + r_2 m_2$, denn d ist R-linear

$$d(r_1 m_1 + r_2 m_2) = d(m_3) \ .$$

$(\tilde{d} \circ \psi)$-Anwenden auf die Urbilder gibt links

$$(\tilde{d} \circ \psi)(r_1 m_1 + r_2 m_2) = r_1 \tilde{d}\psi(m_1) + r_2 \tilde{d}\psi(m_2) = r_1 \tau(m_1) + r_2 \tau(m_2)$$

sowie rechts

$$(\tilde{d} \circ \psi)(m_3) = \tau(r_1 m_1 + r_2 m_2).$$

Aus dem letzten Hilfssatz folgt die Gleichheit der rechten Seiten. Somit ist τ R-linear und macht das rechte Quadrat des Diagramms kommutativ.

<u>Eindeutigkeit von τ</u>: Angenommen es gäbe ein weiteres $\tilde{\tau}$, welches das rechte Dreieck kommutativ macht

$$\begin{tikzcd}
M \arrow[r, "d"] \arrow[d, "\varphi"'] \arrow[dr, "\tilde{d} \circ \psi"] & M'' \arrow[d, "\tilde{\tau}"] \\
N \arrow[r, "\tilde{d}"] & N''
\end{tikzcd}$$

Da d surjektiv ist, ist jede Abbildung $\tilde{\tau}$ aber dann gleich τ, d.h. eindeutig bestimmt. Gilt nämlich

$$m \overset{d}{\longmapsto} m'',$$

dann gilt notwendigerweise auch

$$\tilde{\tau} d(m) = \tilde{d} \circ \psi(m)$$

bzw.

$$\tilde{\tau}(m'') = \tilde{d} \circ \psi(m),$$

wobei m irgendein Urbild von m'' ist.

60 Der Isomorphiesatz

Satz: *Gegeben seien R-Moduln V und W und eine R-lineare Abbildung*

$$\varphi : V \to W \ .$$

Sei U ein R-Untermodul von $Kern(\varphi)$ und $i : U \to V$ die Inklusionsabbildung

$$U \subseteq Kern(\varphi),$$

sowie $\pi : V \to Q$ ein Quotient von $U \subseteq V$. Dann existiert genau eine R-lineare Abbildung $\tilde{\varphi}$, welche das folgende Diagramm

$$\begin{tikzcd}
0 \arrow[r] & U \arrow[r, "i"] \arrow[dr, "0"'] & V \arrow[r, "\pi"] \arrow[d, "\varphi"] & Q \arrow[r] \arrow[dl, dotted, "\exists!\tilde{\varphi}"] & 0 \\
& & W & &
\end{tikzcd}$$

kommutativ macht, d.h. für die gilt

$$\varphi = \tilde{\varphi} \circ \pi.$$

Es gilt $Bild(\tilde{\varphi}) = Bild(\varphi)$ und $Kern(\tilde{\varphi}) = \pi\,(Kern(\varphi))$.

Diese Eigenschaft nennt man die universelle Eigenschaft des Quotienten. Sie zeigt, daß Quotienten bis auf Isomorphie eindeutig sind. Ist nämlich $\varphi : V \to W$ selbst ein Quotient von $U \subseteq V$, dann folgt

$$\begin{tikzcd}
V \arrow[r, "\pi"] \arrow[d, "\varphi"] & Q \arrow[dl, dotted, bend left, "\exists!\tilde{\varphi}"] \\
W \arrow[ur, dotted, "\exists!\tilde{\pi}"] &
\end{tikzcd}$$

Aus der Eindeutigkeit der Abbildungen, folgt $\tilde{\pi} \circ \tilde{\varphi} = id_Q$ und $\tilde{\varphi} \circ \tilde{\pi} = id_W$. Zur Begründung: Ist $\phi : V \to W$ gleich zu der Abbildung $\pi : V \to W$, dann ist die eindeutig bestimmte Abbildung $\tilde{\phi}$ des letzten Satzes natürlich die identische Abbildung.

Beweis: Die Sequenzen

$$\begin{tikzcd}
0 \arrow[r] & U \arrow[r, "i"] \arrow[d, "0"] & V \arrow[r, "\pi"] \arrow[d, "\varphi"] & Q \arrow[r] \arrow[dl, "\exists!\tilde{\varphi}"] \arrow[d, dotted, "\exists!\ \tilde{\varphi}"] & 0 \\
0 \arrow[r] & 0 \arrow[r] & W \arrow[r, equal, "id"] & W \arrow[r] & 0
\end{tikzcd}$$

erfüllen die Voraussetzungen des Satzes von §58. Das linke Diagramm ist kommutativ wegen der Annahme $\phi \circ i = 0$. Die Existenz und Eindeutigkeit von $\tilde{\varphi}$ folgt somit aus dem Satz aus §59.
Zu zeigen bleibt: $Bild(\tilde{\varphi}) = Bild(\varphi)$ und $Kern(\tilde{\varphi}) = \pi(Kern(\varphi))$.

$$\begin{aligned}
w \in Bild(\tilde{\varphi}) &\iff \exists [v] \in V/U \ mit\ \tilde{\varphi}([v]) = w \\
&\overset{\pi\ surj.}{\iff} \exists v \in V \ mit\ \tilde{\varphi}\,(\pi(v)) = w \\
&\iff \exists v \in V \ mit\ id \circ \varphi(v) = w \\
&\iff \exists v \in V \ mit\ \varphi(v) = w \\
&\iff w \in Bild(\varphi)
\end{aligned}$$

q.e.d.

$$\begin{aligned}
[v] \in Kern(\tilde{\varphi}) &\iff v \in Kern(\tilde{\varphi} \circ \pi) \quad \forall v \in \pi^{-1}([v]) \\
&\iff v \in Kern(id \circ \varphi) = Kern(\varphi) \quad \forall v \in \pi^{-1}([v]) \\
&\overset{\pi\ anw.}{\iff} [v] = \pi(v) \in \pi\,(Kern(\varphi))
\end{aligned}$$

Isomorphiesatz: *Sei $\varphi : V \twoheadrightarrow W$ eine surjektive R-lineare Abbildung zwischen R-Moduln V, W. Dann definiert $\tilde{\varphi}$ einen Isomorphismus*

$$\boxed{\tilde{\varphi} : V/Kern(\varphi) \xrightarrow{\cong} W}$$

von R-Moduln.

Beweis: Setze $U = Kern(\varphi)$ im Satz. Wegen $Bild(\tilde{\varphi}) = Bild(\varphi) \overset{!}{=} W$ ist $\tilde{\varphi}$ surjektiv. Wegen $Kern(\tilde{\varphi}) = \pi(Kern(\varphi)) = 0$ ist φ injektiv. Somit ist $\tilde{\varphi}$ bijektiv, also ein Isomorphismus.

Kohomologietheorie

61 Kohomologiegruppen eines Komplexes

Gegeben sei ein Komplex $K^\bullet$ von R-Moduln

$$\cdots \longrightarrow K^{-1} \xrightarrow{d_{-1}} K^0 \xrightarrow{d_0} K^1 \xrightarrow{d_1} K^2 \longrightarrow \cdots$$

mit $d_i \circ d_{i-1} = 0$. Wir betrachten dazu folgende abgeleitete R-Moduln:

<u>Für jedes $i \in \mathbb{Z}$ setzen wir</u>:

$$\textbf{i-Zykel:} \qquad Z^i(K^\bullet) = Kern(d_i : K^i \to K^{i+1})$$

und die

$$\textbf{i-Ränder:} \qquad B^i(K^\bullet) = Bild(d_{i-1} : K^{i-1} \to K^i).$$

Wegen $d_i \circ d_{i-1} = 0$ definiert dies R-Moduln mit der Eigenschaft

$$\boxed{B^i(K^\bullet) \subseteq Z^i(K^\bullet) \subseteq K^i}.$$

Definition: *Die <u>i-te Kohomologiegruppe</u> $H^i(K^\bullet)$ des Komplexes ist definiert als Quotienten-R-Modul*

$$H^i(K^\bullet) = Z^i(K^\bullet)/B^i(K^\bullet).$$

Man hat dann per Definition die folgenden kurzen exakten Sequenzen von R-Moduln

$$\begin{array}{ccccccccc} 0 & \to & B^i(K^\bullet) & \to & Z^i(K^\bullet) & \to & H^i(K^\bullet) & \to & 0 \\ & & & & k & \mapsto & [k] & & \end{array}$$

Offensichtlich gilt $H^i(K^\bullet) = 0$ genau dann, wenn $B^i(K^\bullet) = Z^i(K^\bullet)$ ist. Also

Folgerung: $H^i(K^\bullet) = 0 \iff K^\bullet$ *ist exakt an der Stelle i.*

62 Die induzierte Abbildung φ_*

Seien $K^\bullet$ und $L^\bullet$ Komplexe. Eine Komplexabbildung φ von $K^\bullet$ nach $L^\bullet$ ist per Definition ein System $\varphi_n : K^n \to L^n \quad (n \in \mathbb{Z})$ von R-linearen Abbildungen mit der Eigenschaft

$$\varphi_n \circ d_{n-1} = d_{n-1} \circ \varphi_{n-1}$$

für alle $n \in \mathbb{Z}$.

$$\begin{tikzcd}
\vdots \arrow[d] & \vdots \arrow[d] \\
K^{n-1} \arrow[r, "\varphi_{n-1}"] \arrow[d, "d_{n-1}"] & L^{n-1} \arrow[d, "d_{n-1}"] \\
K^n \arrow[r, "\varphi_n"] \arrow[d] & L^{n-1} \arrow[d] \\
\vdots & \vdots
\end{tikzcd}$$

Lemma 1: *Jede Komplexabbildung φ von $K^\bullet$ nach $L^\bullet$ induziert Abbildungen*

$$\varphi_* : H^n(K^\bullet) \to H^n(L^\bullet)$$

für alle $n \in \mathbb{Z}$.

Beweis: Betrachte

$$\begin{tikzcd}
0 \arrow[r] & B^n(K^\bullet) \arrow[r] \arrow[d, "\varphi_n"] & Z^n(K^\bullet) \arrow[r] \arrow[d, "\varphi_n"] & H^n(K^\bullet) \arrow[r] \arrow[d, dotted, "\varphi_*"] & 0 \\
0 \arrow[r] & B^n(L^\bullet) \arrow[r] & Z^n(L^\bullet) \arrow[r] & H^n(L^\bullet) \arrow[r] & 0
\end{tikzcd}$$

Wir verwenden dabei den Satz aus §58, daß man zwischen zwei exakten Sequenzen das linke, kommutative Diagramm eindeutig ergänzen kann durch eine R-lineare Abbildung φ_* zu einem kommutativen rechten Diagramm.

Das linke Diagramm ist kommutativ und von φ_n induziert: Das heißt die Einschränkung von

$$\varphi_n : K^n \to L^n$$

bildet $Z^n(K^\bullet) = Kern(d_n)$ automatisch auf $Kern(d_n) = Z^n(L^\bullet)$ ab, und bildet $B^n(K^\bullet) = Bild(d_{n-1})$ automatisch auf $Bild(d_{n-1}) = B^n(L^\bullet)$ ab.

Begründung:

$$\begin{tikzcd}
\vdots \arrow[d] & \vdots \arrow[d] \\
K^{n-1} \arrow[r, "\varphi_{n-1}"] \arrow[d, "d_{n-1}"] & L^{n-1} \arrow[d, "d_{n-1}"] \\
K^n \arrow[r, "\varphi_n"] \arrow[d, "d_n"] & L^n \arrow[d, "d_n"] \\
K^{n+1} \arrow[r, "\varphi_{n+1}"] \arrow[d] & L^{n+1} \arrow[d] \\
\vdots & \vdots
\end{tikzcd}
\qquad
\begin{tikzcd}
y \arrow[r, mapsto] \arrow[d, mapsto] & \varphi_{n-1}(y) \arrow[d, mapsto] \\
x \arrow[r, mapsto] \arrow[d, mapsto] & \varphi_n(x) \arrow[d, mapsto] \\
0 \arrow[r, mapsto] & 0
\end{tikzcd}$$

Sei $x \in Z^n(K^\bullet) = Kern(d_n)$. Dann folgt $\varphi_n(x) \in Kern(d_n) = Z^n(L^\bullet)$ wegen

$$d_n\varphi_n(x) = \varphi_{n+1}d_n(x) = \varphi_{n+1}(0) = 0 \ .$$

Gilt andererseits $B^n(K^\bullet) = Bild(d_{n-1}) \ni x = d_{n-1}(y)$, so folgt $\varphi_n(x) = \varphi_n d_{n-1}(y) = d_{n-1}(\varphi_{n-1}(y)) \in Bild(d_{n-1}) = B^n(L^\bullet)$. q.e.d.

Damit ist das anfangs benutzte Diagramm erklärt. Wir haben gezeigt, daß die vertikalen Abbildungen φ_n das linke Quadrat kommutativ machen. φ_* ist die einzige Ergänzung nach §58, welche auch das rechte Quadrat kommutativ macht. Es gilt für $[k] \in H^n(K^\bullet)$

$$\varphi_*([k]) = [\varphi_n(k)]$$

für irgendeinen Repräsentaten $k \in Z^n(K^\bullet)$ mit $[k] = \pi(k)$. Das heißt

$$\begin{tikzcd}
0 \arrow[r] & B^n(K^\bullet) \arrow[r] \arrow[dd, "\varphi_n"] & Z^n(K^\bullet) \arrow[r, "\pi"] \arrow[dd, "\varphi_n"] & H^n(K^\bullet) \arrow[r] \arrow[dd, dotted, "\varphi_*"] & 0 \\
& & k \arrow[r, mapsto] \arrow[dd, mapsto] & {[k]} \arrow[dd, mapsto] & \\
0 \arrow[r] & B^n(L^\bullet) \arrow[r] & Z^n(L^\bullet) \arrow[r] & H^n(L^\bullet) \arrow[r] & 0 \\
& & \varphi_n(k) \arrow[r, mapsto] & {[\varphi_n(k)]} &
\end{tikzcd}$$

Lemma 2: *Seien φ und ψ Abbildungen von Komplexen (φ von $K^\bullet$ nach $L^\bullet$ und ψ von $L^\bullet$ nach $M^\bullet$). Dann ist die zusammengesetzte Abbildung $\psi \circ \varphi$ von $K^\bullet$ nach $M^\bullet$ erklärt durch*

$$(\psi\varphi)_n = \psi_n \circ \varphi_n$$

Dies ist eine Komplexabbildung und die induzierte Abbildung

$$(\psi\varphi)_* : H^\bullet(K^\bullet) \to H^\bullet(L^\bullet)$$

ist die Zusammensetzung

$$(\psi\varphi)_* = \psi_* \circ \varphi_*$$

von $\varphi_* : H^\bullet(K^\bullet) \to H^\bullet(M^\bullet)$ und $\psi_* : H^\bullet(M^\bullet) \to H^\bullet(L^\bullet)$.

<u>Beweis</u>: Wir müssen zuerst zeigen, daß $(\psi\varphi)_n$ definiert durch $\psi_n \circ \varphi_n$ eine Komplexabbildung definiert. Gegeben ist

$$\begin{tikzcd}
K^{n-1} \arrow[r, "\varphi_{n-1}"] \arrow[d, "d"] & L^{n-1} \arrow[r, "\psi_{n-1}"] \arrow[d, "d"] & M^{n-1} \arrow[d, "d"] \\
K^n \arrow[r, "\varphi_n"] & L^n \arrow[r, "\psi_n"] & M^n
\end{tikzcd}$$

Zu zeigen ist, daß die zusammengesetzten Quadrate kommutativ sind:

$$\begin{tikzcd}
K^{n-1} \arrow[r, "\psi_{n-1}\varphi_{n-1}"] \arrow[d, "d"] & M^{n-1} \arrow[d, "d"] \\
K^n \arrow[r, "\psi_n\varphi_n"] & M^n
\end{tikzcd}$$

Dies folgt aber aus:

$$\underline{(\psi_n\varphi_n)d} = \psi_n(\varphi_n d) = \psi_n(d\varphi_{n-1}) = (\psi_n d)\varphi_{n-1} = (d\psi_{n-1})\varphi_{n-1} = \underline{d(\psi_{n-1}\varphi_{n-1})}$$

Damit ist gezeigt, daß das System $\psi\varphi = (\psi_n\varphi_n)_{n\in\mathbb{Z}}$ eine Abbildung vom Komplex $K^\bullet$ nach $M^\bullet$ ist.

Zu zeigen bleibt: $\underline{\psi_*\varphi_* = (\psi\varphi)_*}$: Wir betrachten die definierenden Sequenzen für die Kohomologie und die Zusammensetzung von φ_* und ψ_* :

$$\begin{tikzcd}
0 \arrow[r] & B^n(K^\bullet) \arrow[r] \arrow[d, "\varphi_n"] & Z^n(K^\bullet) \arrow[r] \arrow[d, "\varphi_n"] & H^n(K^\bullet) \arrow[r] \arrow[d, dotted, "\varphi_*"] & 0 \\
0 \arrow[r] & B^n(L^\bullet) \arrow[r] \arrow[d, "\psi_n"] & Z^n(L^\bullet) \arrow[r] \arrow[d, "\psi_n"] & H^n(L^\bullet) \arrow[r] \arrow[d, dotted, "\psi_*"] & 0 \\
0 \arrow[r] & B^n(M^\bullet) \arrow[r] & Z^n(M^\bullet) \arrow[r] & H^n(M^\bullet) \arrow[r] & 0
\end{tikzcd}$$

sowie die definierenden Sequenzen für die Abbildung $(\psi\varphi)_*$:

$$\begin{tikzcd}
0 \arrow[r] & B^n(K^\bullet) \arrow[r] \arrow[d, "\psi_n\varphi_n=(\psi\varphi)_n"] & Z^n(K^\bullet) \arrow[r] \arrow[d, "\psi_n\varphi_n=(\psi\varphi)_n"] & H^n(K^\bullet) \arrow[r] \arrow[d, dotted, "(\psi\varphi)_*"] & 0 \\
0 \arrow[r] & B^n(M^\bullet) \arrow[r] & Z^n(M^\bullet) \arrow[r] & H^n(M^\bullet) \arrow[r] & 0
\end{tikzcd}$$

In dem ersten Diagramm ergibt sich durch Aneinanderreihen der beiden rechten kommutativen Diagramme ein kommutatives Diagramm.

$$\begin{tikzcd}
0 \arrow[r] & B^n(K^\bullet) \arrow[r] \arrow[d, "\psi_n\varphi_n"] & Z^n(K^\bullet) \arrow[r] \arrow[d, "\psi_n\varphi_n"] & H^n(K^\bullet) \arrow[r] \arrow[d, dotted, "\psi_*\circ\varphi_*"] & 0 \\
0 \arrow[r] & B^n(M^\bullet) \arrow[r] & Z^n(M^\bullet) \arrow[r] & H^n(M^\bullet) \arrow[r] & 0
\end{tikzcd}$$

Da die rechte kommutative Ergaänzung eindeutig ist, folgt durch Vergleich der beiden letzten Diagramme:

$$\psi_* \circ \varphi_* = (\psi\varphi)_*.$$

63 Die lange exakte Kohomologiesequenz

Seien $K^\bullet, L^\bullet$ und $M^\bullet$ Komplexe von R-Moduln. Wir nehmen an, wir haben Abbildungen φ_n und ψ_n gegeben für alle $n \in \mathbb{Z}$ wie im folgenden Diagramm:

$$\begin{tikzcd}
 & \vdots \arrow[d] & \vdots \arrow[d] & \vdots \arrow[d] & \\
0 \arrow[r] & K^{n-1} \arrow[r, "\varphi_{n-1}"] \arrow[d] & L^{n-1} \arrow[r, "\psi_{n-1}"] \arrow[d] & M^{n-1} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & K^{n} \arrow[r, "\varphi_{n}"] \arrow[d] & L^{n} \arrow[r, "\psi_{n}"] \arrow[d] & M^{n} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & K^{n+1} \arrow[r, "\varphi_{n+1}"] \arrow[d] & L^{n+1} \arrow[r, "\psi_{n+1}"] \arrow[d] & M^{n+1} \arrow[r] \arrow[d] & 0 \\
 & \vdots & \vdots & \vdots &
\end{tikzcd}$$

so daß außerdem gilt:

1) (φ_n) ist eine Komplexabbildung von $K^\bullet$ nach $L^\bullet$, d.h.

$$\varphi_n \circ d = d \circ \varphi_{n-1} \qquad \forall n \in \mathbb{Z}.$$

2) (ψ_n) ist eine Komplexabbildung von $L^\bullet$ nach $M^\bullet$ d.h.

$$\psi_n \circ d = d \circ \psi_{n-1} \qquad \forall n \in \mathbb{Z}.$$

3) für alle $n \in \mathbb{Z}$ sind die waagrechten Sequenzen

$$0 \longrightarrow K^n \xrightarrow{\quad\varphi_n\quad} L^n \xrightarrow{\quad\psi_n\quad} M^n \longrightarrow 0$$

<u>kurze exakte</u> Sequenzen, d.h. insbesondere ist φ_n injektiv und ψ_n surjektiv und es gilt $\psi_n \circ \varphi_n = 0$.

Satz: *Dann gibt es R-lineare Abbildungen $\delta : H^n(M^\bullet) \to H^{n+1}(K^\bullet)$ für alle $n \in \mathbb{Z}$, so daß*

$$\begin{tikzcd}
 & & {} \arrow[dll, "\delta"'] \\
H^n(K^\bullet) \arrow[r, "\varphi_*"] & H^n(L^\bullet) \arrow[r, "\psi_*"] & H^n(M^\bullet) \arrow[dll, "\delta"'] \\
H^{n+1}(K^\bullet) \arrow[r, "\varphi_*"] & H^{n+1}(L^\bullet) \arrow[r, "\psi_*"] & H^{n+1}(M^\bullet) \arrow[dll, "\delta"'] \\
{} & &
\end{tikzcd}$$

eine lange exakte Sequenz von R-Moduln definiert.

Die auftretenden Abbildungen $\delta : H^n(M^\bullet) \to H^{n+1}(K^\bullet)$ heißen *Verbindungs-homomorphismen*.

Bemerkung: In vielen Anwendungen ist nur die Berechnung der Kohomologie-moduln $H^n(\ )$ und die Berechnung der Abbildungen φ_*, ψ_* wichtig. Von δ wird oft nur die Existenz benötigt.

Beweis: 1. Schritt: Konstruktion von δ

Sei $[m] \in H^n(M^\bullet.)$ Dann wähle ein Urbild in der definierenden Sequenz

$$\begin{array}{ccccccccc}
0 & \longrightarrow & Bild(d_{n-1}) & \longrightarrow & Kern(d_n) & \longrightarrow & H^n(M^\bullet) & \longrightarrow & 0 \\
 & & & & m & \longmapsto & [m] & &
\end{array}$$

Dann ist $m \in M^n$, so daß $d(m) = 0$ gilt.

$$\begin{tikzcd}
& \vdots \arrow[d] & & \vdots \arrow[d] & & \vdots \arrow[d] & & & \\
0 \arrow[r] & K^n \arrow[rr, "\varphi_n"] \arrow[dd] & & L^n \arrow[rr, "\psi_n"] \arrow[dd] & & M^n \arrow[rr] \arrow[dd] & & 0 & \\
& & & & l_n \arrow[rr, mapsto] \arrow[dd, mapsto] & & m \arrow[rr, mapsto] \arrow[dd, mapsto] & & 0 \\
0 \arrow[r] & K^{n+1} \arrow[rr, "\varphi_{n+1}"] \arrow[dd] & & L^{n+1} \arrow[rr, "\psi_{n+1}"] \arrow[dd] & & M^{n+1} \arrow[rr] \arrow[dd] & & 0 & \\
& & k_{n+1} \arrow[rr, mapsto] \arrow[dd, mapsto] & & l_{n+1} \arrow[rr, mapsto] \arrow[dd, mapsto] & & d(m) = 0 & & \\
0 \arrow[r] & K^{n+2} \arrow[rr, "\varphi_{n+2}"] \arrow[dd] & & L^{n+2} \arrow[rr, "\psi_{n+2}"] \arrow[dd] & & M^{n+2} \arrow[rr] \arrow[dd] & & 0 & \\
& & ? \arrow[rr, mapsto] & & 0 & & & & \\
& \vdots & & \vdots & & \vdots & & &
\end{tikzcd}$$

Da ψ_n surjektiv ist, gibt es ein Urbild $l_n \in L^n$ mit $\psi_n(l_n) = m$. Wegen $\psi_{n+1}(dl_n) = d\psi_n(l_n) = 0$ und der Exaktheit bei L^{n+1} gilt $l_{n+1} = d(l_n) = \varphi_{n+1}(k_{n+1})$ für ein $k_{n+1} \in K^{n+1}$.

<u>Beh.</u>: $d(k_{n+1}) = 0$

<u>Bew.</u>: Es gilt

$$\varphi_{n+2}(dk_{n+1}) = d\varphi_{n+1}(k_{n+1}) = dl_{n+1} = ddl_n = 0$$

und die Injektivität von φ_{n+2} und $\varphi_{n+2}(dk_{n+1}) = 0$ impliziert $dk_{n+1} = 0$.

<u>Wir machen den Ansatz</u>: $\delta([m]) := [k_{n+1}]$

Willkür trat auf bei der Konstruktion von k_{n+1} an zwei Stellen:

(i) bei der Wahl des Repräsentanten m von $[m]$: $\qquad m' = m + d(m_n)$

(ii) bei der Wahl des Repräsentanten l_n von m : $\qquad l'_n = l_n + \varphi_n(k_n)$

Andere Wahlen bei (i) und (ii) führen zu einem modifizierten Element

$$l_n + \varphi_n(k_n) + d(l_{n-1})$$

an Stelle von l_n. Dies gibt wegen $dd(l_{n-1}) = 0$ als Modifikation von $l_{n+1} = d(l_n)$

$$l_{n+1} + d\varphi_n(k_n) \ = \ l_{n+1} + \varphi_n(dk_n) \ .$$

Somit erhält man an Stelle von k_{n+1} das Element

$$k'_{n+1} = k_{n+1} + d(k_n) \ .$$

Es folgt für die Kohomologieklasse $[k'_{n+1}] = [k_{n+1} + d(k_n)] = [k_{n+1}]$

$$\begin{array}{ccccccccc} 0 & \longrightarrow & Bild(d_n) & \longrightarrow & Kern(d_{n+1}) & \longrightarrow & H^{n+1}(M^\bullet) & \longrightarrow & 0 \\ & & d(k_n) & & k'_{n+1}, k_{n+1} & \longmapsto & [k_{n+1}] & & \end{array}$$

Die Abbildung δ ist also wohldefiniert

$$\boxed{\begin{array}{ccc} H^n(M^\bullet) & \xrightarrow{\delta} & H^{n+1}(K^\bullet) \\ [m] & \longmapsto & [k_{n+1}] \end{array}}$$

und, wie man leicht sieht, R-linear.

<u>2. Schritt</u>: $\delta \circ \psi_* = 0$ auf $H^n(L^\bullet)$

Sei $[m] = \psi_*([l])$ mit $l = l_n \in Z^n(L^\bullet)$. Dann gilt $\delta([m]) = 0$ wegen $l_{n+1} = dl_n = 0$ und somit $k_{n+1} = 0$. Es folgt $\delta \circ \psi_*([l]) = 0$ für alle $[l] \in H^n(L^\bullet)$.

<u>3. Schritt</u>: $\varphi_* \circ \delta = 0$ auf $H^n(M^\bullet)$

$\varphi_*\delta([m]) = \varphi_*([k_{n+1}]) = [\varphi_{n+1}(k_{n+1})] = [l_{n+1}] = 0$, da $l_{n+1} = d(l_n)$ in $B^{n+1}(L^\bullet)$ liegt.

<u>4. Schritt</u>: $\psi_* \circ \varphi_* = 0$ auf $H^n(K^\bullet)$

$\psi_* \circ \varphi_* = (\psi \circ \varphi)_* = 0_* = 0$, da die waagrechten Abbildungen exakt sind und somit $\psi_n \circ \varphi_n = 0$ für alle $n \in \mathbb{Z}$ gilt. Benutze dann Lemma 2 von §62.

<u>5. Schritt</u>: $Kern(\delta) \subseteq Bild(\psi_*)$ in $H^n(M^\bullet)$

Sei $[m] \in Kern(\delta)$. Dann gilt $k_{n+1} = d(k_n) \in B^{n+1}(K^\bullet)$.

$$\begin{tikzcd} k_n \arrow[d, mapsto] & \tilde{l}_n,\ l_n \arrow[r, mapsto] \arrow[d, mapsto] & m \\ k_{n+1} \arrow[r, mapsto] & l_{n+1} & \end{tikzcd}$$

Für $\tilde{l}_n = l_n - \varphi_n(k_n)$ gilt dann auch $\psi_n(\tilde{l}_n) = \psi_n(l_n) = m$. Anderseits gilt jetzt $d(\tilde{l}_n) = 0$, also $\tilde{l}_n \in Z^n(L^\bullet)$. Dies definiert eine Kohomologieklasse $[\tilde{l}_n] \in H^n(L^\bullet)$ und es gilt $[m] = \psi_*([\tilde{l}_n]) \in Bild(\psi_*)$.

<u>6. Schritt</u>: $Kern(\psi_*) \subseteq Bild(\varphi_*)$ in $H^n(L^\bullet)$

Sei $[l_n] \in Kern(\psi_*)$. Dann gilt $\psi_n(l_n) = d(m_{n-1}) \in B^n(K^\bullet)$ und $dl_n = 0$

$$\begin{tikzcd} & l_{n-1} \arrow[r, mapsto] \arrow[d, mapsto] & m_{n-1} \arrow[r, mapsto] \arrow[d, mapsto, "d"] & 0 \\ k_n \arrow[r, mapsto] \arrow[d, mapsto] & l'_n, l_n \arrow[r, mapsto, "\psi_n"] \arrow[d, mapsto] & m_n & \\ ? \arrow[r, mapsto] & 0 & & \end{tikzcd}$$

$l'_n = l_n - d(l_{n-1})$ ist ein modifiziertes Urbild $[l'_n] = [l_n]$ von $[m]$. Wählt man für l_{n-1} ein Urbild von m_{n-1}, so folgt aus $\psi_{n-1}(l_{n-1}) = m_{n-1}$

$$\psi_n(l'_n) = \psi_n(l_n) - \psi_n d(l_{n-1}) = \psi_n(l_n) - dm_{n-1} = m_n - m_n = 0.$$

Wegen der Exaktheit bei L^n folgt daraus $l'_n = \varphi_n(k_n)$ für ein $k_n \in K^n$. Es gilt wegen der Injektivität von φ_{n+1} und wegen $dl'_n = 0$ sogar $d(k_n) = 0$! Also ist k_n in $Z^n(K^\bullet)$ und somit ist $[k_n] \in H^n(K^\bullet)$ eklärt. Es folgt für die Kohomologieklassen $[l_n] = \varphi_*([k_n]) \in Bild(\varphi_*)$.

<u>7. Schritt</u>: $Kern(\varphi_*) \subseteq Bild(\delta)$ in $H^{n+1}(K^\bullet)$

Sei $[k_{n+1}] \in Kern(\varphi_*)$, $\varphi_*([k_{n+1}]) = 0$, d.h. $\varphi_{n+1}(k_{n+1}) = d(l_n) \in B^{n+1}(K^\bullet)$

$$\begin{tikzcd} & l_n \arrow[r, mapsto] \arrow[d, mapsto] & m_n \arrow[d, mapsto] \\ k_{n+1} \arrow[r, mapsto] & l_{n+1} = \varphi_{n+1}(k_{n+1}) \arrow[r, mapsto] & 0 \end{tikzcd}$$

Setze $m_n = \psi_n(l_n)$. Dann gilt

$$d(m_n) = d\psi_n(l_n) = \psi_{n+1}d(l_n) = \psi_{n+1}(l_{n+1}) = \psi_{n+1}\varphi_{n+1}(k_{n+1}) = 0,$$

also ist $m_n \in Z^n(M^\bullet)$ und dies definiert die Kohomologieklasse $[m_n] \in H^n(M^\bullet)$. Dann gilt aber per Definition der Verbindungsabbildung (!)

$$\delta([m_n]) = [k_{n+1}].$$

Daraus folgt $Kern(\varphi_*) \subseteq Bild(\delta)$.

Damit ist der Beweis dieses technischen Satzes abgeschlossen.

64 Das Schlangenlemma

Ein Spezialfall der langen, exakten Kohomologiesequenz aus dem letzten Abschnitt ist das Schlangenlemma

Annahme $(*)$. Seien zwei kurze exakte Sequenzen gegeben

$$\begin{tikzcd}
0 \arrow[r] & K^0 \arrow[r, "\varphi_0"] \arrow[d, "\chi"] & L^0 \arrow[r, "\psi_0"] \arrow[d, "\lambda"] & M^0 \arrow[r] \arrow[d, "\mu"] & 0 \\
0 \arrow[r] & K^1 \arrow[r, "\varphi_0 1"] & L^1 \arrow[r, "\psi_1"] & M^1 \arrow[r] & 0
\end{tikzcd}$$

und vertikale Abbildungen wie im Diagramm. Alle Moduln seien R-Moduln, alle Abbildungen R-linear. Beide Quadrate des Diagramms seien kommutativ.

Dann kann man durch Auffüllen mit Nullmoduln Komplexe $K^\bullet, L^\bullet$ und $M^\bullet$ konstruieren:

$$\begin{tikzcd}
 & \vdots \arrow[d] & \vdots \arrow[d] & \vdots \arrow[d] & \\
0 \arrow[r] & 0 \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & K^0 \arrow[r, "\varphi_0"] \arrow[d, "\chi"] & L^0 \arrow[r, "\psi_0"] \arrow[d, "\lambda"] & M^0 \arrow[r] \arrow[d, "\mu"] & 0 \\
0 \arrow[r] & K^1 \arrow[r, "\varphi_1"] \arrow[d] & L^1 \arrow[r, "\psi_1"] \arrow[d] & M^1 \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & 0 \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & 0 \\
 & \vdots & \vdots & \vdots &
\end{tikzcd}$$

Die Kohomologie des Komplexes $K^\bullet$ ist dann gegeben durch:

$$H^i(K^\bullet) = \begin{cases} 0 & i \neq 0, 1 \\ Kern(\chi) & i = 0 \\ Kokern(\chi) & i = 1 \end{cases}$$

Gleiches gilt für die Komplexe $L^\bullet$ und $M^\bullet$.

Hierbei wurde folgende allgemeine Definition benutzt

Definition: Ist $\chi : K_0 \to K_1$ eine R-lineare Abbildung von R-Moduln, dann ist der $Kokern(\chi)$ definiert durch:

$$Kokern(\chi) = K_1/Bild(\chi).$$

Als Spezialfall der langen, exakten Kohomologiesequenz erhält man – angewandt auf obige Situation – die sogenannte Kern-Kokern-Sequenz:

Lemma: Unter der Annahme $(*)$ von weiter oben gibt es eine exakte Sequenz

$$\begin{tikzcd}
0 \arrow[r] & Kern(\chi) \arrow[r, "\varphi_0"] & Kern(\lambda) \arrow[r, "\psi_0"] & Kern(\mu) \arrow[dll, "\delta"] & \\
& Kokern(\chi) \arrow[r] & Kokern(\lambda) \arrow[r] & Kokern(\mu) \arrow[r] & 0
\end{tikzcd}$$

Dieses Lemma wird aus offensichtlichen Gründen Schlangenlemma genannt.

Algebren

65 K–Algebren

Sei K ein fester Körper oder allgemein ein fest gewählter kommutativer Ring und A ein Ring, welcher K enthält. A kann als K–Modul aufgefaßt werden. Wir nehmen an A sei eine freier R–Modul. Dann besitzt A eine Basis B über K

$$A = K \cdot e_1 \oplus ... \oplus K \cdot e_n$$

oder allgemein $A = \bigoplus_{e_i \in B} K \cdot e_i$, falls $|B| = \infty$.

Definition: A wie oben heißt dann K–**Algebra**, wenn K außerdem im Zentrum $Z(A)$ von A liegt, d.h. wenn gilt

$$\lambda \cdot a = a \cdot \lambda$$

für alle $\lambda \in K$ und alle $a \in A$.

Bemerkung: Die Distributivgesetze für den Ring A besagen, daß die Multiplikation additiv in beiden Variablen ist. Da $K \cdot 1_A$ nach Annahme im Zentrum von A liegt, läßt sich mit dem Distributivgesetz in der Aussage zusammenfassen, daß die Multiplikation

$$A \times A \xrightarrow{m} A$$

eine K–bilineare Abbildung ist. Eine solche genügt es auf den Basiselementen $(e_i, e_j) \in A \times A$ festzulegen. Umgekehrt definieren daher beliebige Vorgaben von Konstanten $C_{ij}^k \in K$ mit $e_i \cdot e_j = \sum_{k=1}^n C_{ij}^k \cdot e_k$ eine K-bilineare Multiplikationsabbildung m auf A. Dies definiert eine K–Algebra, wenn das Assoziativgesetz gilt oder äquivalent

$$\sum_l C_{ij}^l C_{lk}^m = \sum_l C_{il}^m C_{jk}^l \qquad \forall\ i, j, k, m\ ,$$

und wenn ein Einselement existiert.

Beachte: Wegen der Distributivgesetze genügt es, das Assoziativgesetz auf einer K–Basis nachzuweisen!

Eine K–Algebra ist also vollständig bestimmt durch die **Strukturkonstanten** C_{ij}^k

$$e_i \cdot e_j = \sum_{k=1}^{n} C_{ij}^k \cdot e_k.$$

Häufig wählen wir für e_1 das Einselement $e_1 = 1_A$ des Ringes A.

Beispiel: Der Polynomring $K[X]$ ist eine K–Algebra mit der Basis $e_i = X^i$ für $i = 0, 1, 2, ...$ und den Strukturkonstanten $C_{ij}^k = 1$ oder 0 je nachdem, ob $i + j = k$ gilt oder nicht.

Beispiel: (n=2):
In diesem Fall gilt für $A = K \cdot \mathbf{1} \oplus K \cdot e$ allgemein $e^2 = \alpha \cdot \mathbf{1} + \beta \cdot e$. Im Fall $\frac{1}{2} \in K$ kann durch die Basiswahl $\tilde{e} = e - \frac{1}{2}\beta$ angenommen werden, daß $\tilde{e} \cdot \tilde{e} = e^2 - \beta \cdot e + \frac{1}{4}\beta^2 = (\alpha + \frac{1}{4}\beta^2) \in K$. Also obdA $e^2 \in K \cdot 1$.

Definition: $C(a)$ ist die K–Algebra

$$C(a) = K \cdot \mathbf{1} \oplus K \cdot e$$

mit

$$e^2 = -a$$

für gegebenes $a \in K$. Setze $C^+(a) = K \cdot 1$ und $C^-(a) = K \cdot e$.

Bemerkung: Schreibt man $x \cdot 1 + y \cdot e$ in der Form (x, y), dann werden die Rechengesetze von $C(a)$ durch $(x, y) + (x', y') = (x + x', y + y')$ und $(x, y)(x', y') = (xx' - ayy', xy' + yx')$ beschrieben. Man überprüft leicht durch explizites Nachrechnen, daß das Assoziativgesetz sowie die übrigen Ringaxiome erfüllt sind.

Bemerkung: Ist K ein Körper, dann ist $C(a)$ genau dann wieder ein Körper, wenn a kein Quadrat in K ist. Dies zeigt man im Fall des Körpers der komplexen Zahlen. Für $a = 0$ ist $C(0)$ kein Körper, sondern ein Spezialfall der folgenden, allgemeiner definierten Graßmann–Algebren.

66 Graßmann Algebra

Sei $V = K \cdot e_1 \oplus ... \oplus K \cdot e_m$ ein m–dimensionaler freier K–Modul. Die Graßmann–Algebra $A = \Lambda^\bullet(V)$ hat per Definition 2^m Basiselemente

$$e_I \quad , \quad I \subseteqq \{1, ..., m\},$$

welche durch die Teilmenge I der Menge $\{1, ..., m\}$ indiziert werden:

$$\Lambda^\bullet(V) \ = \bigoplus_{I \subseteqq \{1,...,m\}} K \cdot e_I.$$

Wir schreiben oft e_i anstelle von $e_{\{i\}}$ zur Vereinfachung der Schreibweise im Fall der einelementigen Mengen $I = \{i\}$.

Die charakterisierenden Relationen der Graßmann Algebra sind

$$(*) \qquad \begin{array}{ll} e_i \cdot e_j = -e_j \cdot e_i & (1 \leq i, j, \leq m) \\ e_i \cdot e_i = 0 & (1 \leq i \leq m), \end{array}$$

sowie

$$(**) \qquad e_I = e_{i_1} \cdot e_{i_2} \cdots e_{i_r},$$

für

$$I = \{i_1, i_2, ..., i_r\}, \quad i_1 < i_2 < ... < i_r.$$

Diese Formeln legen die Strukturkonstanten der Graßmann Algebra bereits vollkommen fest, wie man sich leicht überlegt: Beispielsweise folgt aus den obigen Bedingungen

$$e_I \cdot e_J = (-1)^{|I| \cdot |J|} e_J \cdot e_I$$

sowie

$$e_I \cdot e_J = \begin{cases} 0 & I \cap J \neq \emptyset \\ \pm e_{I \cup J} & I \cap J = \emptyset. \end{cases}$$

Das Vorzeichen $\pm$ in der letzten Formel ist dabei vollkommen festgelegt als das Signum der Permutation, welche die Zahlen $(i_1, ..., i_r, j_1, ..., j_s)$ mit $i_1 < ... < i_r$ sowie $j_1 < ... < j_s$ in aufsteigende Anordnung überführt. Hierbei sei $I = \{i_1, .., i_r\}$ und $J = \{j_1, ..., j_s\}$.

Zur Existenz von $\Lambda^\bullet(V)$: Es bleibt natürlich zu zeigen, daß die geforderten Bedingungen $(*)$ und $(**)$ tatsächlich einen Ring A mit K–Basis $\{e_I\}$ definieren, welcher insbesondere das Assoziativgesetz erfüllt. Dies kann man natürlich direkt verifizieren. Eleganter und einfacher ist aber die im übernächsten Paragraphen §68 beschriebene Konstruktion mittels eines getwisteten Tensorprodukts.

Bezeichnung:

$$\Lambda^+(V) = \bigoplus_{\substack{I \\ |I| \text{ gerade}}} K \cdot e_I$$

$$\Lambda^-(V) = \bigoplus_{\substack{I \\ |I| \text{ ungerade}}} K \cdot e_I$$

Bemerkung: Für das Produkt in der Graßmann Algebra schreibt man häufig $v \wedge w$ anstelle von $v \cdot w$.

67 Tensorprodukte

Seien V und W zwei K–Moduln. Das Tensorprodukt $V \otimes_K W$ ist ein K–Modul und durch folgende universelle Eigenschaft bis auf Isomorphie eindeutig festgelegt:

Es gibt eine K–bilineare Abbildung

$$\pi : V \times W \longrightarrow V \otimes_K W$$

derart, daß für jede K–bilineare Abbildung $B : V \times W \to U$ in einen K–Modul U eine eindeutig bestimmte K–lineare Abbildung $b : V \otimes_K W \to U$ existiert, welche das Diagram

$$\begin{tikzcd}
V \times W \arrow[rr, "\pi"] \arrow[dr, "B"'] & & V \otimes_K W \arrow[dl, dotted, "\exists! b"] \\
& U &
\end{tikzcd}$$

kommutativ macht.

<u>Bemerkung</u>: Das Tensorprodukt existiert immer. Für zwei freie K–Moduln V und W von endlichem Rang ist dies aber besonders einfach zu sehen. Ist e_1, ..., e_r eine K–Basis von V und e'_1, ..., e'_s eine K–Basis von W, dann definieren wir $V \otimes_K W$ als des freien Modul $K^{r \cdot s}$ mit den Basisvektoren $e_i \otimes_K e'_j$ (formale Bezeichnung). Die Abbildung π wird definiert durch $\pi(e_i, e_j) = e_i \otimes_K e_j$. Für bilineares $B : V \times W \to U$ ist die gesuchte Abbildung b durch folgende Vorgabe auf den Basisvektoren (siehe §11)

$$b(e_i \otimes e'_j) = B(e_i, e'_j).$$

Man überprüft leicht, daß die geforderte universelle Eigenschaft gilt. Insbesondere gilt

$$rang_K(V \otimes_K W) \;=\; rang_K(V) \;\cdot\; rang_K(W).$$

<u>Warnung</u>: $rang_K(V \times W) \;=\; rang_K(V) \;+\; rang_K(W)$.

Tensorprodukte von Algebren:
Seien A und A' zwei K–Algebren, dann ist das Tensorprodukt

$$A \otimes_K A'$$

wieder eine K–Algebra, wenn man die Multiplikation wie folgt definiert

$$(a_1 \otimes_K a_1') \cdot (a_2 \otimes_K a_2') = a_1 a_2 \otimes_K a_1' a_2'.$$

In der Tat ist $1_A \otimes_K 1_{A'}$ das Einselement. Das Assoziativgesetz der K–Algebra ist unmittelbar klar.

Bemerkung: Das Tensorprodukt kommutativer K–Algebren ist offensichtlich wieder kommutativ. Um auch nichtlommutative Algebren zu erhalten betrachten wir alternativ auch folgende getwistete Konstruktion.

68 Getwistete Produkte von Algebren

Eine K–Algebra heißt $\mathbb{Z}_2$–graduierte Algebra, wenn gilt

$$A = A_+ \oplus A_-$$

mit K–Untermoduln A_+, A_- derart, daß

$$(*) \qquad \begin{array}{ccc} A_+ \cdot A_+ \subset A_+ & , & A_- \cdot A_- \subset A_+ \\ A_- \cdot A_+ \subset A_- & , & A_+ \cdot A_- \subset A_-. \end{array}$$

Beispiel: $A = \Lambda^\bullet(V)$ mit $A_\pm = \Lambda^\pm(V)$ oder $C(a)$ mit $C^\pm(a)$.

Die Elemente aus $A_+ \cup A_-$ werden "homogen" genannt. Für homogene Elemente $a \in A_+$ resp. $a \in A_-$ ist der Grad definiert durch $|a| = 0$ resp. $|a| = 1$. Aus den Regeln für die Multiplikation homogener Elemente $(*)$ folgt $|a \cdot b| = |a| + |b| \in \mathbb{Z}_2$ für homogene a und b.

Sind A und A' zwei $\mathbb{Z}_2$–graduierte Algebren, dann kann man dem Tensorprodukt eine **getwistete Ringstruktur** geben vermöge der Definition

$$(a_1 \otimes_K a_1') \cdot (a_2 \otimes_K a_2') = (-1)^{|a_1'||a_2|} a_1 a_2 \otimes_K a_1' a_2'.$$

Hierbei seien $a_i \in A_\pm$ resp. $a_i' \in A_\pm'$ "homogene" Elemente.

Genauer: Die Abbildung M, definiert für die homogenen Elemente wie folgt

$$\begin{array}{rrcl} M: & A \times A' \times A \times A' & \longrightarrow & A \otimes_K A' \\ & (a_1, a_1', a_2, a_2') & \longmapsto & (-1)^{|a_1'||a_2|} a_1 a_2 \otimes_K a_1' a_2', \end{array}$$

ist K–bilinear in (a_1, a_1') und K–bilinear in (a_2, a_2'). Wegen der universellen Eigenschaft des Tensorprodukts faktorisiert somit M über eine Abbildung

$$(A \otimes_K A') \times (A \otimes_K A') \xrightarrow{m} A \otimes_K A'.$$

Diese Abildung m induziert die Multiplikation auf $A \otimes_K^\varepsilon A'$.

Es ist wieder eine leicht Aufgabe zu zeigen, daß diese Multiplikation eine K–Algebra definiert. Diese K–Algebra nennt man das **getwistete Tensorprodukt** $A \otimes_K^\varepsilon A'$ von A und A'. Im übrigen ist $A \otimes_K^\varepsilon A'$ wieder $\mathbb{Z}_2$–graduiert mit der Zerlegung in

$$(A \otimes_K^\varepsilon A')_+ \;=\; (A_+ \otimes_K A'_+) \;\oplus\; (A_- \otimes_K A'_-)$$

und

$$(A \otimes_K^\varepsilon A')_- \;=\; (A_- \otimes_K A'_+) \;\oplus\; (A_- \otimes_K A'_+).$$

Existenz der Eins: $1_A \otimes_K 1_{A'}$ ist ein Einselement von $A \otimes_K^\varepsilon A'$. Dies benützt, daß die Eins einer $\mathbb{Z}_2$–graduierten Algebra stets homogen vom Grad 0 ist.

Nachweis der K–Algebreneigenschaften: Die beiden Distributivgesetze und die K–Algebreneigenschaft besagen gerade, daß die Multiplikationsabbildung K–linear in der ersten und zweiten Variable ist. Wegen der universellen Eigenschaft des Tensorprodukts genügt dazu, daß die Abbildung M K–bilinear in den Variablen (a_1, a'_1) resp. K–bilinear in den Variablen (a_2, a'_2) ist. Es verbleibt also nur das Assoziativgesetz zu zeigen. Wegen der Distributivgesetze, genügt es, das Assoziativgesetz auf Erzeugenden der additiven Gruppe nachzuweisen. Wir können uns daher darauf beschränken

$$\Big[(a_1 \otimes_K^\varepsilon a'_1) \cdot (a_2 \otimes_K^\varepsilon a'_2)\Big] \cdot (a_3 \otimes_K^\varepsilon a'_3) = (a_1 \otimes_K^\varepsilon a'_1) \cdot \Big[(a_2 \otimes_K^\varepsilon a'_2) \cdot (a_3 \otimes_K^\varepsilon a'_3)\Big]$$

zu zeigen für homogene Elemente $a_1, a'_1, a_2, a'_2, a_3, a'_3$.
In der Tat ist das Assoziativgesetz dann äquivalent zu

$$(-1)^{|a'_1||a_2|} \cdot (-1)^{(|a'_1|+|a'_2|)|a_3|} = (-1)^{|a'_1|(|a_2|+|a_3|)} \cdot (-1)^{|a'_2||a_3|},$$

was unmittelbar nachgeprüft werden kann. Beachte:

Für homogene $a_i \in A$ resp. $a'_i \in A'$ gilt $[(a_1 \otimes_K a'_1) \cdot (a_2 \otimes_K a'_2)] \cdot (a_3 \otimes_K a'_3) = (-1)^{|a'_1||a_2|}(a_1 a_2 \otimes_K a'_1 a'_2) \cdot (a_3 \otimes_K a'_3) = (-1)^{|a'_1||a_2|+|a'_1 a'_2||a_3|} a_1 a_2 a_3 \otimes_K a'_1 a'_2 a'_3 = (-1)^{|a'_1||a_2|+|a'_1||a_3|+|a'_2||a_3|} a_1 a_2 a_3 \otimes_K a'_1 a'_2 a'_3$. Berechnet man $(a_1 \otimes_K a'_1) \cdot [(a_2 \otimes_K a'_2) \cdot (a_3 \otimes_K a'_3)]$ erhält man dasselbe Resultat wegen $(-1)^{|a'_1||a_2|+|a'_1||a_2 a_3|} = (-1)^{|a'_2||a_3|+|a'_1||a_3|+|a'_2||a_3|}$. Daraus folgt das Assoziativitätsgesetz für beliebige Elemente von $A \otimes_K^\varepsilon A'$ durch Linearkombinieren.

Schliesslich ist klar, daß $A \otimes_K^\varepsilon A'$ bezüglich der natürlich induzierten Graduierung auf $A \otimes_K^\varepsilon A'$ wieder eine graduierte Algebra wird. Wir überlassen dies als Übungsaufgabe dem Leser.

Beispiel: $A = C(0) \otimes_K^\varepsilon C(0)$ liefert die Graßmann Algebra $\Lambda^\bullet(V)$ für zweidimensionales V. Man erhält als Basis von $A = C(0) \otimes_K^\varepsilon C(0)$ nämlich

$$1 = 1 \otimes_K^\varepsilon 1, \quad e_1 = e \otimes_K^\varepsilon 1, \quad e_2 = 1 \otimes_K^\varepsilon e, \quad e_{\{1,2\}} = e \otimes_K^\varepsilon e.$$

Man überprüft sofort $e_1^2 = e_2^2 = 0$ und $e_1 \cdot e_2 = -e_2 \cdot e_1$ sowie dann $e_{\{1,2\}} = e_1 \cdot e_2$.

69 Iterierte Tensorprodukte

Allgemeiner: Durch Iteration erhält man die Graßmann Algebra

$$\Lambda^\bullet(K \cdot e_1 \oplus ... \oplus K \cdot e_n) = \Lambda^\bullet(K \cdot e_1) \otimes_K^\varepsilon ... \otimes_K^\varepsilon \Lambda^\bullet(K \cdot e_n).$$

Insbesondere folgt

$$\Lambda^\bullet(V \oplus W) = \Lambda^\bullet(V) \ \otimes_K^\varepsilon \ \Lambda^\bullet(W).$$

Lemma: *Die Abbildung*

$$\begin{array}{ccc} A \otimes_K^\varepsilon A' & \longrightarrow & A' \otimes_K^\varepsilon A \\ a \otimes a' & \longmapsto & (-1)^{|a|\cdot|a'|} a' \otimes a \end{array}$$

für homogene $a \in A$, $a' \in A'$ definiert einen Ringisomorphismus

$$A \otimes_K^\varepsilon A' \ \cong \ A' \otimes_K^\varepsilon A$$

von $\mathbb{Z}_2$–graduierten K–Algebren.

Beweis: Betrachte

$$\begin{tikzcd} a_1 \otimes a_1' \arrow[d, mapsto, "(-1)^{|a_1||a_1'|}"'] & \times & a_2 \otimes a_2' \arrow[d, mapsto, "(-1)^{|a_2||a_2'|}"] \arrow[r, mapsto] & a_1a_2 \otimes a_1a_2' \cdot (-1)^{|a_1'||a_2|} \arrow[d, mapsto, "(-1)^{|a_1+a_2|\cdot|a_1'+a_2'|}"] \\ a_1' \otimes a_1 & \times & a_2' \otimes a_2 \arrow[r, mapsto] & a_1'a_2' \otimes a_1a_2 \cdot (-1)^{|a_1||a_2'|} \end{tikzcd}$$

und benutze

$$(-1)^{|a_1||a_1'|} \cdot (-1)^{|a_2||a_2'|} = (-1)^{|a_1||a_2|} \cdot (-1)^{|a_1||a_2'|} \cdot (-1)^{|a_1+a_2|\cdot|a_1'+a_2'|}.$$

Ähnlich zeigt man das Assoziativgesetz des Tensorprodukts

$$(A_1 \otimes_K^\varepsilon A_2) \otimes_K^\varepsilon A_3 = A_1 \otimes_K^\varepsilon (A_2 \otimes_K^\varepsilon A_3).$$

Dies folgert man aus

$$(a_1 \otimes a_1' \otimes a_1'') \cdot (a_2 \otimes a_2' \otimes a_2'') = (-1)^{|a_1''||a_2|+|a_1''||a_2'|+|a_1'||a_2|} a_1 a_2 \otimes a_1' a_2' \otimes a_1'' a_2''$$

jeweils für homogene Elemente in beiden Fällen. Siehe auch die Aufgabe auf dem Übungsblatt.

Schlußfolgerung: Hat man graduierte K–Algebren A_1, ..., A_n, dann kann man iterierte Tensorprodukte bilden ohne auf Klammerung achten zu müssen. Wir schreiben daher kurz

$$A = A_1 \otimes_K^{\varepsilon} A_2 \otimes_K^{\varepsilon} ... \otimes_K^{\varepsilon} A_n \ .$$

Man überlegt sich außerdem leicht, daß die so definierte graduierte Algebra nicht von der Reihenfolge der Tensorproduktbildung abhängt.

70 Die Quaternionenalgebra

Die Quaternionenalgebren sind definiert durch

$$C(a_1, a_2) = C(a_1) \otimes_K^\varepsilon C(a_2).$$

für $a_1, a_2 \in K^*$. Zur Erinnerung $C(a_i) = K \cdot 1 + K \cdot e_i$ mit $e_i^2 = -a_i$ für $i = 1, 2$. Dann ist

$$\begin{aligned} 1 &= 1 \otimes_K^\varepsilon 1 \\ \mathbf{j} &= e_1 \otimes_K^\varepsilon 1 \\ \mathbf{k} &= 1 \otimes_K^\varepsilon e_2 \\ \mathbf{l} &= e_1 \otimes_K^\varepsilon e_2 = \mathbf{j} \cdot \mathbf{k} \end{aligned}$$

eine Basis von $C(a_1, a_2)$. Man erhält

$$\begin{aligned} \mathbf{j} \cdot \mathbf{j} &= -a_1 \cdot 1 \\ \mathbf{k} \cdot \mathbf{k} &= -a_2 \cdot 1 \\ \mathbf{l} \cdot \mathbf{l} &= -a_1 a_2 \cdot 1 \end{aligned}$$

sowie $\mathbf{j} \cdot \mathbf{k} = -\mathbf{k} \cdot \mathbf{j}$, $\mathbf{j} \cdot \mathbf{l} = -\mathbf{l} \cdot \mathbf{j}$ und $\mathbf{k} \cdot \mathbf{l} = -\mathbf{l} \cdot \mathbf{k}$. Für ein beliebiges Element $z \in C(a_1, a_2)$

$$z = \alpha \cdot 1 + \beta \cdot \mathbf{j} + \gamma \cdot \mathbf{k} + \delta \cdot \mathbf{l}$$

definiert man

$$\overline{z} = \alpha \cdot 1 - \beta \cdot \mathbf{j} - \gamma \cdot \mathbf{k} - \delta \cdot \mathbf{l}.$$

Lemma 1: *Es gilt $N(z) := z \cdot \overline{z} = \overline{z} \cdot z$ ist gleich*

$$N(z) = \alpha^2 + a_1 \cdot \beta^2 + a_2 \cdot \gamma^2 + a_1 a_2 \cdot \delta^2$$

und für alle $z_1, z_2 \in C(a_1, a_2)$ gilt

$$\begin{aligned} \overline{z_1 + z_2} &= \overline{z_1} + \overline{z_2} \\ \overline{z_1 \cdot z_2} &= \overline{z_2} \cdot \overline{z_1}. \end{aligned}$$

Wir geben später eine Verallgemeinerung dieser Aussage und lassen daher den Beweis als Übungsaufgabe.

Lemma 2: *$C(a,-a) \cong M_{2,2}(K)$ falls $a \neq 0$.*

Beweis:

$$\mathbf{j} = \begin{pmatrix} 0 & -1 \\ a & 0 \end{pmatrix} \quad \text{und} \quad \mathbf{k} = \begin{pmatrix} 0 & 1 \\ a & 0 \end{pmatrix}$$

erfüllen für $E = 1$

$$\mathbf{j}^2 = -a \cdot E \quad \text{und} \quad \mathbf{k}^2 = +a \cdot E$$

sowie $\mathbf{j} \cdot \mathbf{k} = -\mathbf{k} \cdot \mathbf{j} = diag(-a,a) =: \mathbf{l}$. Dann gilt $\mathbf{l}^2 = -a^2 \cdot E$ und $\mathbf{j} \cdot \mathbf{l} = \mathbf{j} \cdot \mathbf{j} \cdot \mathbf{k} = -\mathbf{j} \cdot \mathbf{k} \cdot \mathbf{j} = -\mathbf{l} \cdot \mathbf{j}$ sowie $\mathbf{k} \cdot \mathbf{l} = \mathbf{k} \cdot \mathbf{j} \cdot \mathbf{k} = -\mathbf{j} \cdot \mathbf{k} \cdot \mathbf{k} = -\mathbf{l} \cdot \mathbf{k}$. Offensichtlich bilden $\mathbf{1}$, $\mathbf{j}$, $\mathbf{k}$ und $\mathbf{l}$ eine K–Basis von $M_{2,2}(K)$.

Beachte: Der Matrixring $M_{2,2}(K)$ erhält durch Lemma 2 eine Struktur einer $\mathbb{Z}_2$–graduierten Algebra durch die Zerlegung:

$$M_{2,2}(K) \;\cong\; C(a,-a) \;\cong\; C^+(a,-a) \oplus C^-(a,-a).$$

Hierbei entsprechen die geraden und ungeraden Anteile den Matrizen vom Typ

$$\begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix} \quad \text{gerade bzw.} \quad \begin{pmatrix} 0 & * \\ * & 0 \end{pmatrix} \quad \text{ungerade.}$$

Lemma 3: *$C(a_1,a_2)$ ist ein* **Schiefkörper** *(alle Körperaxiome sind erfüllt mit Ausnahme des Kommutativgesetzes für die Multiplikation) falls gilt*

$$\alpha^2 + a_1\beta^2 + a_2\gamma^2 + a_1a_2\delta^2 = 0 \quad \Longleftrightarrow \quad (\alpha,\beta,\gamma,\delta) = 0$$

für alle $\alpha,\beta,\gamma,\delta \in K$.

Beispiel: $K = \mathbb{R}$; $a_1 = a_1 = 1$. In diesem Fall erhält man den Schiefkörper $\mathbf{H}$ der **Hamiltonschen** Quaternionen.

71 Tensorprodukte von Matrizen

Sei K ein Körper oder allgemeiner ein kommutativer Ring und sei $V = K^n$ ein freier K–Modul. Die K–bilinearen Abbildungen $\varphi : V \to V$ bilden dann bezüglich Komposition einen Ring, den Endomorphismenring $End_K(V)$. Dieser Ring ist offensichtlich eine K–Algebra mit den Basiselementen $e_{i,j}$ $(i, j = 1, ..., n)$ und den Rechengesetzen

$$e_{i,j} \cdot e_{i',j'} = \begin{cases} 0 & j \neq i' \\ e_{i,j'} & j = i' \end{cases}$$

und dem Einselement $E = e_{1,1} + e_{2,2} + ... + e_{n,n}$. Sei nun $W = K^m$ ein weitere freier K–Modul. Weiterhin seien gegeben Matrizen

$$M \in End_K(V)$$

$$N \in End_K(W).$$

Wir definieren dann eine Matrix

$$M \otimes_K N \in End_K(V \otimes_K W)$$

durch die Vorschrift

$$(M \otimes_K N)(e_i \otimes_K e_j) = M(e_i) \otimes_K N(e_j).$$

Offensichtlich gilt dann allgemeiner

$$(M \otimes_K N)(v \otimes_K w) = M(v) \otimes_K N(w)$$

für alle $v \in V$, $w \in W$. Für $v = \sum \lambda_i e_i$ und $w = \sum \mu_j e_j$ erhält man dies durch Aufsummieren der obigen Gleichungen nach Multiplikation mit $\lambda_i \mu_j$.

Folgerung: $(M_1 \otimes_K N_1) \circ (M_2 \otimes_K N_2) = (M_1 M_2 \otimes_K N_1 N_2)$.

Beweis: Beide Seiten auf Elemente $v \otimes_K w$ (mit $v \in V$, $w \in W$) angewendet liefern dasselbe Bild. Da endliche Summen von Elementen $v \otimes_K w$ das Tensorprodukt $V \otimes_K W$ erzeugen, folgt die Behauptung.

Folgerung: *Die Zuordnung $(M,N) \mapsto M \otimes_K N$ definiert einen K–Algebrenhomomorphismus*

$$\begin{array}{rccc} \rho : & End_K(V) \times End_K(W) & \longrightarrow & End_K(V \otimes_K W) \\ & (M,N) & \longmapsto & M \otimes_K N. \end{array}$$

<u>Beispiel</u>: Sei $V = K^2$, $W = K^2$, $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, $N = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$. Dann ist

$$M \otimes_K N = \begin{pmatrix} \alpha M & \beta M \\ \gamma M & \delta M \end{pmatrix} = \begin{pmatrix} \alpha a & \alpha b & \beta a & \beta b \\ \alpha c & \alpha d & \beta c & \beta d \\ \gamma a & \gamma b & \delta a & \delta b \\ \gamma c & \gamma d & \delta c & \delta d \end{pmatrix}$$

bezüglich der Basis $e_1 \otimes_K e_1$, $e_2 \otimes_K e_1$, $e_1 \otimes_K e_2$ und $e_2 \otimes_K e_2$ von $V \otimes_K W \cong K^4$.

Lemma: *Die Abbildung ρ induziert einen K–Algebrenisomorphismus*

$$\tilde{\rho} : \quad End_K(V) \otimes_K End_K(W) \xrightarrow{\sim} End_K(V \otimes_K W).$$

<u>Beweis</u>: ρ ist K–bilinear in M und N und faktorisiert daher über einen K–Algebrenhomomorphismus $\tilde{\rho}$.

$$\begin{tikzcd} End_K(V) \times End_K(W) \arrow[rr] \arrow[dr, "\rho"'] & & End_K(V) \otimes_K End_K(W) \arrow[dl, dotted, "\exists!\tilde{\rho}"] \\ & End_K(V \otimes_K W) & \end{tikzcd}$$

Um zu zeigen, daß $\tilde{\rho}$ ein Isomorphismus ist, genügt es zu zeigen, daß $\tilde{\rho}$ surjektiv ist. Ein Dimensionsvergleich zeigt nämlich $n^2 \cdot m^2 \overset{!}{=} (n \cdot m)^2$. Aber $\tilde{\rho}$ ist surjektiv, falls ρ surjektiv ist. Die Surjektivität von ρ folgt schließlich aus $\tilde{\rho}(e_{i,j} \otimes_K e_{i',j'}) = e_{ii',jj'}$. Dies macht man sich am besten klar an Hand des obigen Beispiels.

72 *Der getwistete Fall

Seien die Bezeichnungen wie im vorigen Abschnitt. Seien allgemeiner aber

$$\begin{aligned} V &= V_+ \oplus V_- \\ W &= W_+ \oplus W_- \end{aligned}$$

etc. $\mathbb{Z}_2$–graduierte freie K–Moduln. Das heißt $V_\pm = K^{n_\pm}$ und $W_\pm = K^{m_\pm}$.

Die Wahl einer $\mathbb{Z}_2$–Graduierung $V = V_+ \oplus V_-$ induziert eine Zerlegung

$$End_K(V) \; = \; End_K^+(V) \; \oplus \; End_K^-(V) \; .$$

Hierbei sei $End_K^+(V)$ die Algebra aller Endomorphismen $\varphi : V \to V$ ist mit

$$\varphi(V_\pm) \subseteq V_\pm$$

sowie $End_K^-(V)$ die Menge aller K–Endomorphismen $\varphi : V \to V$ mit

$$\varphi(V_\pm) \subseteq V_\mp \; .$$

<u>In Matrixform</u>:

$$End_K^+(V) \; = \; \left\{ \left(\begin{array}{c|c} * & 0 \\ \hline 0 & * \end{array} \right) \right\}$$

beziehungsweise

$$End_K^-(V) \; = \; \left\{ \left(\begin{array}{c|c} 0 & * \\ \hline * & 0 \end{array} \right) \right\}$$

im Sinn von $(n_+ + n_-) \times (m_+ + m_-)$–Blockmatrizen.

Sei nun außerdem auch $W = W_+ \oplus W_-$ graduiert. Beachte, daß dann auch $V \otimes_K^\varepsilon W$ graduiert ist und somit $End_K(V \otimes_K^\varepsilon W)$ im obigen Sinne graduiert ist. Die Frage ist, ob und wie man diese Graduierung in Termen der graduierten K–Algebra $End_K(V)$ und $End_K(W)$ beschreiben kann. Die Antwort lautet

Lemma: *Es gibt einen kanonischen Isomorphismus von $\mathbb{Z}_2$–graduierten K–Algebren*

$$\tilde{\rho}: \quad End_K(V) \otimes_K^\varepsilon End_K(W) \xrightarrow{\sim} End_K(V \otimes_K^\varepsilon W),$$

der vom graduierten Tensorprodukt zweier Matrizen induziert wird.

<u>Beweis</u>: Der Beweis ist analog zum Beweis des Lemmas in §71. In der Tat ist das dort behandelte Lemma der Spezialfall $V = V_+$ und $W = W_+$. Wir beschränken uns daher darauf, die Abbildung $\tilde{\rho}$ anzugeben: $M \in End_K^\pm(V)$ und $N \in End_K^\pm(W)$ bilden wir ab auf den Endomorphismus

$$\tilde{\rho}(M \otimes_K^\varepsilon N) \in End_K(V \otimes_K^\varepsilon W),$$

welcher durch

$$v \otimes_K^\varepsilon w \longmapsto (-1)^{|N|\cdot|v|} M(v) \otimes_K^\varepsilon N(w)$$

gegeben ist für homogenes $v \in V_\pm$ (und $w \in W_\pm$).

73 Cliffordalgebren

Sei K ein Körper und $\frac{1}{2} \in K$, sei V ein endlich dimensionaler K–Vektorraum und $\langle v, w\rangle$ eine symmetrische K–Bilinearform auf V.

Wir betrachten dann K–Algebren A und K–lineare Abbildungen

$$V \xrightarrow{\varphi} A$$

mit der Eigenschaft

$$(*) \qquad \varphi(v) \cdot \varphi(v) + \langle v, v\rangle \cdot 1_A = 0.$$

Erfüllt (φ, A) diese Eigenschaft, dann auch $(\varphi', A') = (f \circ \varphi, A')$,

$$\begin{tikzcd} V \arrow[rr, "\varphi"] \arrow[dr, "\varphi'"'] & & A \arrow[dl, "f"] \\ & A' & \end{tikzcd}$$

wobei $f : A \to A'$ ein beliebiger K–Algebrenhomomorphismus ist.

Satz: *Es gibt ein universelles Paar $(\varphi^{\mathrm{univ}}, A)$ wie oben mit der Eigenschaft $(*)$, so daß jedes andere Paar (φ', A') aus (φ, A) durch Anwenden eines K–Algebrenhomomorphismus $f : A \to A'$ entsteht.*

$$\begin{tikzcd} V \arrow[rr, "\varphi^{\mathrm{univ}}"] \arrow[dr, "\varphi'"'] & & A \arrow[dl, dotted, "\exists! f"] \\ & A' & \end{tikzcd}$$

Definition: Das universelle Paar $(\varphi^{\mathrm{univ}}, A)$ definiert die Cliffordalgebra $A = C(V, \langle\cdot,\cdot\rangle)$ von $(V, \langle\cdot,\cdot\rangle)$. Es wird durch die universelle Eigenschaft bis auf Isomorphie eindeutig charakterisiert.

<u>Bemerkung</u>: Für eine K-lineare Abbildung $\varphi : V \to A$ sind die folgenden Eigenschaften äquivalent

(i) $\varphi(v) \cdot \varphi(v) + \langle v, v\rangle \cdot 1_A = 0$ für alle $v \in V$.

(ii) $\varphi(v) \cdot \varphi(w) + \varphi(w) \cdot \varphi(v) + 2 \cdot \langle v, w\rangle \cdot 1_A = 0$ für alle $v, w \in V$.

(iii) $\varphi(b_i) \cdot \varphi(b_j) + \varphi(b_j) \cdot \varphi(b_i) + 2 \cdot \langle b_i, b_j\rangle \cdot 1_A = 0$ für eine Basis $\{b_i\}$ von V.

Der Beweis erfolgt durch Polarisierung und ist vollkommen analog zum Beweis von Lemma 1 in §45. q.e.d

Da eine K–lineare Abbildung $\varphi : V \to A$ durch die Vorgabe auf einer Basis $b_1, ..., b_n$ von V eindeutig festgelegt ist, betrachten wir Algebren A mit $e_i = \varphi(b_i) \in A$ mit der Eigenschaft

$$e_i \cdot e_j + e_j \cdot e_i + 2 \cdot \langle b_i, b_j\rangle \cdot 1_A = 0.$$

Wählt man für b_1,..., b_n eine Orthonormalbasis der Form $\langle \cdot, \cdot\rangle$, dann gilt $\langle b_i, b_j\rangle = \delta_{ij} \cdot a_i$ für gewisse a_i.

Folgerung: *$A = C(a_1, ..., a_n)$ mit $e_i = 1 \otimes^\varepsilon ... \otimes^\varepsilon 1 \otimes^\varepsilon e \otimes^\varepsilon 1 ... \otimes^\varepsilon 1$ für $i = 1, ..., n$ hat die Eigenschaft*

$$e_i \cdot e_j + e_j \cdot e_i + 2 \cdot a_i \cdot \delta_{ij} \cdot 1_A = 0.$$

Behauptung: $A = C(a_1, ..., a_n)$ mit $\varphi(b_i) = e_i$ hat die gewünschte universelle Eigenschaft. (Insbesondere ist für die Nullform $< .,. >= 0$ die Cliffordalgebra gerade die Graßmann Algebra).

Beweis: Man reduziert den Beweis mit Induktion nach n auf den trivialen Fall $n = 1$ mit Hilfe der folgenden Aussage: Seien A_1, A_2, A' $\mathbb{Z}_2$-graduierte K-Algebren und $f_i : A_i \to A'$ Grad erhaltende K-Algebrenhomomorphismen. Diese setzten sich zu einem Grad erhaltenden K-Algebrenhomomorphismus $f : A_1 \otimes^\varepsilon A_2 \to A$ fort durch $f(a_1 \otimes a_2) = f_1(a_1) \cdot f_2(a_2)$ falls $f_1(a_1) \cdot f_2(a_2) = (-1)^{|a_1||a_2|} f_2(a_2) f_1(a_1)$ gilt für alle homogenene Elemente $a_i \in A_i$. q.e.d

Satz: *Ist ψ eine isometrische K–lineare Abbildung*

$$\psi : V \longrightarrow W,$$

bezüglich symmetrischer Bilinearformen $\langle\cdot,\cdot\rangle_V$ resp. $\langle\cdot,\cdot\rangle_W$ auf V und W, d.h.

$$\langle\psi(v),\psi(v')\rangle_W = \langle v,v'\rangle_V ,$$

dann existiert ein eindeutig bestimmter K–Algebrenhomomorphismus $C(\psi)$ zwischen den Cliffordalgebren $C(V,\langle\cdot,\cdot\rangle_V)$ und $C(W,\langle\cdot,\cdot\rangle_W)$, welcher folgendes Diagramm kommutativ macht:

$$\begin{tikzcd} V \arrow[r, "\psi"] \arrow[d, "\varphi^{\mathrm{univ}}"'] & W \arrow[d, "\varphi^{\mathrm{univ}}"] \\ C(V,\langle\cdot,\cdot\rangle_V) \arrow[r, dotted, "\exists!\ C(\psi)"] & C(W,\langle\cdot,\cdot\rangle_W). \end{tikzcd}$$

Beweis: Ist ψ eine isometrische K–lineare Abbildung $\psi : V \to W$, dann liefert die Zusammensetzung $\varphi = \varphi^{\mathrm{univ}} \circ \psi$ eine K–linearen Abbildung von V in die K–Algebra $A = C(W,\langle,\rangle_W)$ mit der Eigenschaft

$$\begin{aligned} \varphi(v)\cdot\varphi(v) \quad &+ \quad \langle v,v\rangle_V \cdot 1_A \\ &= \quad \varphi^{\mathrm{univ}}(\psi(v))\cdot\varphi^{\mathrm{univ}}(\psi(v)) + \langle\psi(v),\psi(v)\rangle_W \cdot 1_A \\ &= \quad 0. \end{aligned}$$

Somit existiert wegen der universellen Eigenschaft von $(\varphi, C(V,\langle,\rangle_V))$ ein eindeutiger Algebrenhomomorphismus $C(\psi) : C(V,\langle,\rangle_V) \to C(W,\langle,\rangle_W)$, welcher das obige Diagramm kommutativ macht. q.e.d

Korollar: *Für Isometrien $\phi : U \to V$ und $\psi : V \to W$ (bezüglich gewisser symmetrischer K-Bilinearformen $< .,. >_U, < .,. >_V$ und $< .,. >_W$ auf U,V,W gilt*

$$C(\psi\circ\phi) = C(\psi)\circ C(\phi)\ .$$

Beweis: Durch Aneinandersetzen der Diagramme für ϕ und ψ des obigen Satz, erhält man ein kommutatives Diagramm

$$\begin{tikzcd} U \arrow[r, "\psi\circ\phi"] \arrow[d, "\varphi^{\mathrm{univ}}"'] & W \arrow[d, "\varphi^{\mathrm{univ}}"] \\ C(U,\langle\cdot,\cdot\rangle_U) \arrow[r, dotted, "C(\psi)\circ C(\phi)"] & C(W,\langle\cdot,\cdot\rangle_W). \end{tikzcd}$$

Wegen der Eindeutigkeit eines solchen Diagram, ist der Homomorphismus $C(\psi)\circ C(\phi)$ von K-Algebren daher notwendiger Weise gleich $C(\psi\circ\phi)$. Damit ist das Korollar bewiesen.

Im Spezialfall, wo $\langle\cdot,\cdot\rangle$ die Nullform auf V ist, ist die Cliffordalgebra gerade die Graßmann Algebra $\Lambda^\bullet(V)$ und

$$\begin{tikzcd} V \arrow[rr, "\varphi"] \arrow[dr, equal] & & \Lambda^\bullet(V) \\ & \Lambda^1(V) \arrow[ur, hook] & \end{tikzcd}$$

ist die kanonische Abbildung auf die Schicht vom Grad 1.

Im Spezialfall der Nullform $< .,. >\equiv 0$ ist jede K–lineare Abbildung $\psi : V\to W$ eine Isometrie und man erhält als Spezialfall das

Korollar: *Ist $\psi : V \to W$ eine K–lineare Abbildung, dann existiert genau ein K–Algebrenhomomorphismus*

$$\Lambda^\bullet(\psi) : \Lambda^\bullet(V) \longrightarrow \Lambda^\bullet(W)$$

mit der Eigenschaft

$$\begin{tikzcd} \Lambda^\bullet(V) \arrow[r, dotted, "\Lambda^\bullet(\psi)"] & \Lambda^\bullet(W) \\ \Lambda^1(V) \arrow[u, hook] \arrow[r, "\psi"] & \Lambda^1(W). \arrow[u, hook] \end{tikzcd}$$

sowie $\Lambda^\bullet(\psi\circ\phi) = \Lambda^\bullet(\psi)\circ\Lambda^\bullet(\phi)$.

Im Fall der Graßmann Algebren hat man außerdem die Besonderheit, daß der Ringhomomorphismus $\Lambda^\bullet(\psi)$ die einzelnen Schichten der Graßmann Algebra in sich abbildet. Das heißt die Abbildung $\Lambda^\bullet(\psi)$ zerfällt in eine ganze Kollektion von Abbildungen

$$\Lambda^i(\psi):\ \Lambda^i(V)\to\Lambda^i(W) \quad , \quad \forall i\in\mathbb{N}\ .$$

Zum Beweis genügt, daß $\Lambda^i(V)$ erzeugt wird von den i-fachen Produkten von Vektoren aus V. Diese werden auf Summen von i-fachen Produkten von Vektoren aus W, also nach $\Lambda^i(W)$ abgebildet. Wir schreiben im folgenden $v\wedge w$ für das Produkt in der Graßmann Algebra.

74 Die Determinante

Wir betrachten jetzt die Abbildung $\Lambda^\bullet(\psi)$ des vorigen Abschnitts in einem Spezialfall. Sei dazu

$$V = W = K^n$$

und sei

$$\psi : K^n \to K^n$$

ein Endomorphismus. Die induzierten K-Abbildungen

$$\Lambda^i(\psi) : \Lambda^i(K^n) \to \Lambda^i(K^n)$$

lassen sich bezüglich der Basis $e_I, |I| = i$ wieder als Matrix schreiben. Die Koeffizienten dieser Matrix lassen sich durch Unterdeterminanten der Matrix von ψ beschreiben. Wir diskutieren hier nur den wichtigsten Spezialfall $i = n$.

Beispiel: Wir beginnen mit dem Fall $n = 2$. Die Abbildung ψ wird durch eine Matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

beschrieben. Offensichtlich ist $\Lambda^0(\psi) = id$, da $e_\emptyset = 1$ in $e_\emptyset = 1$ übergeführt wird. Auf $K^2 = \Lambda^1(K^2)$ ist $\Lambda(\psi)$ gerade ψ selbst. $\Lambda^2(k^2) = Ke_{\{1,2\}}$ wird von $e_{\{1,2\}} = e_1 \wedge e_2$ erzeugt. Da $\Lambda^\bullet(\psi)$ ein multiplikativ ist (ein Ringhomomorphismus!), folgt

$$\Lambda^\bullet(\psi)(e_1 \cdot e_2) = \Lambda^\bullet(\psi)(e_1) \cdot \Lambda^\bullet(\psi)(e_2) = (a \cdot e_1 + c \cdot e_2) \wedge (b \cdot e_1 + d \cdot e_2) \ .$$

Aus $e_1 \wedge e_1 = e_2 \wedge e_2 = 0$ und $e_2 \wedge e_1 = -e_1 \wedge e_2$ folgt dann durch distributives Ausmultiplizieren

$$\begin{aligned} &= (ad - bc) \cdot e_1 \wedge e_2 \\ &= det(\psi) \cdot e_1 \wedge e_2 \ . \end{aligned}$$

Allgemeiner

Lemma: *Es gilt für alle Endomorphismen $\psi : k^n \to k^n$*

$$\Lambda^n(\psi) \ = \ det(\psi) \ .$$

Beweis: Zu zeigen ist $\Lambda^n(\psi)(e_{\{12..n\}}) = det(\psi) \cdot e_{\{12..n\}}$. Wie im Fall $n = 2$ gilt

$$\Lambda^\bullet(\psi)(e_1 \wedge e_2 \wedge ... \wedge e_n) = \Lambda^\bullet(\psi)(e_1) \wedge \Lambda^\bullet(\psi)(e_2) \wedge ... \wedge \Lambda^\bullet(\psi)(e_n)$$

$$= \sum_{j_1=1}^{n} M_{j_1 1} e_{j_1} \wedge \sum_{j_2=1}^{n} M_{j_2 2} e_{j_2} \wedge ... \wedge \sum_{j_n=1}^{n} M_{j_n n} e_{j_n} \ .$$

Wegen $e_i \wedge e_i = 0$ kann man alle Terme weglassen, für die $(j_1, .., j_n)$ nicht paarweise verschieden sind. Diese Indizes entsprechen eindeutig den Permutationen σ aus der symmetrischen Gruppe S_n vermöge der Zuordnung $j_i = \sigma(i)$ für $i = 1, ..., n$. Es ergibt sich

$$\sum_{\sigma \in S_n} sign(\sigma) \cdot \prod_{k=1}^{n} M_{\sigma(k)k} \cdot e_1 \wedge ... \wedge e_n$$

$$= det(\psi) \cdot e_1 \wedge ... \wedge e_n \ .$$

Das Vorzeichen $sign(\sigma)$ ergibt sich dabei aus den Graßmann Regeln beim Umordnen des Produktes $e_{\sigma(1)} \wedge ... \wedge e_{\sigma(n)} = sign(\sigma) \cdot e_1 \wedge ... \wedge e_n$ in aufsteigende Reihenfolge wie im Fall $n = 2$.

Ausblick: Somit erweisen sich die im letzten Paragraph bewiesenen Formeln

$$\Lambda^i(\psi \circ \phi) = \Lambda^i(\psi) \circ \Lambda^\bullet(\phi)$$

als eine weitreichende Verallgemeinerung der Determinanten Multiplikationsformel. Denn $\phi : U \to V$ und $\psi : V \to W$ waren dabei beliebige R-lineare Abbildungen (für einen kommutativen Ring R), und diese Formeln gelten für alle $i = 1, 2, 3, ...$

75 Der Koszulkomplex

Seien f_1,..., f_n K–lineare Abbildungen $f_i : R \longrightarrow R$ einer kommutativen K–Algebra R, welche miteinander kommutieren:

$$f_i \circ f_j = f_j \circ f_i \qquad , \quad \forall i, j.$$

Dann definiert man den Koszulkomplex $K(f_1, ..., f_n)$ durch die K-Vektorräume

$$\Lambda^i(R^n) = \bigoplus_{|I|=i} R \cdot e_I$$

mit den K-linearen Abbildungen $d_i : \Lambda^i(R^n) \longrightarrow \Lambda^{i+1}(R^n)$, welche für $r \in R$ auf den Elementen $r \cdot e_I$ durch

$$d_i(r \cdot e_I) \ = \ \sum_{\nu=1}^{n} f_\nu(r) \cdot e_\nu \wedge e_I$$

und sonst durch K-lineare Fortsetzung definiert werden. Dies definiert einen Komplex von K-Vektorräumen, denn offensichtlich gilt $d_{i+1} \circ d_i = 0$ wegen

$$d_{i+1} \circ d_i(r \cdot e_I) \ = \ \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} f_\mu(f_\nu(r)) \cdot e_\mu \wedge e_\nu \wedge e_I \ = \ 0 \ .$$

Benutze $e_\mu \wedge e_\nu + e_\nu \wedge e_\mu = 0$ sowie $(f_\mu \circ f_\nu)(r) = (f_\nu \circ f_\mu)(r)$. Im Fall $n = 2$ erhält man nämlich $(f_1 \circ f_1)(r) \cdot e_1 \wedge e_1 + (f_1 \circ f_2)(r) \cdot e_1 \wedge e_2 + (f_2 \circ f_1)(r) \cdot e_2 \wedge e_1 + (f_2 \circ f_2)(r) \cdot e_2 \wedge e_2 = [(f_1 \circ f_2)(r) - (f_2 \circ f_1)(r)] \cdot e_1 \wedge e_2 = 0$. Der allgemeine Fall $n \geq 2$ geht genauso.

Beispiel: $(n = 1)$ Hier ist $K^0 = K^1 = R$, $f = f_1$ und der Komplex $K(f_1)$ kann mit dem Komplex

$$R \xrightarrow{f} R$$

identifiziert werden.

Beispiel: $(n = 2)$ Hier ist $K^0 = K^2 = R$ und $K^1 = R$ und der Komplex $K(f_1, f_2)$ kann mit dem Komplex

$$\begin{array}{ccccc} R & \overset{(f_1,f_2)}{\longrightarrow} & R^2 & \overset{f_2-f_1}{\longrightarrow} & R \\ r & \longmapsto & (f_1(r), f_2(r)) & & \\ & & (r_1, r_2) & \longmapsto & f_2(r_2) - f_1(r_1) \end{array}$$

identifiziert werden.

<u>Übungsaufgabe</u>: Sei $R = K[x,y] = (K[x])[y]$ der Polynomring in zwei Variablen und K eine Körper. Berechne die Kohomologiegruppen dieses Komplexes $K(f_1, f_2)$ in den folgenden beiden Fällen:

1) f_1 resp. f_2 ist Multiplikation mit x respektive y.

2) f_1 resp. f_2 ist Ableiten nach x respektive nach y.

76 Differentialformen

Ist $\mathcal{U} \subseteqq \mathbb{R}^n$ eine offene Teilmenge, dann ist der Ring der $\mathbb{R}$-wertigen unendlich oft partiell diffenzierbaren Funktionen auf $\mathcal{U}$

$$R = C^\infty(\mathcal{U}, \mathbb{R})$$

eine kommutative $\mathbb{R}$–Algebra.

Die partiellen Ableitungen $\partial_i = \frac{\partial}{\partial x_i}$, $i = 1, ..., n$ definieren $\mathbb{R}$–lineare Abbildungen $\partial_i : R \to R$ mit der Eigenschaft

$$\partial_i \circ \partial_j = \partial_j \circ \partial_i$$

(Symmetrie der Hessematrix).

Setzt man daher $K = \mathbb{R}$ und $f_i = \partial_i$ mit den Bezeichnungen des vorigen Abschnitts, so erhält man den Differentialformenkomplex der **alternierenden Differentialformen** auf $\mathcal{U}$.

$$\begin{array}{rcl} A^i(\mathcal{U}, \mathbb{R}) & = & \Lambda^i(C^\infty(\mathcal{U}, \mathbb{R})^n) \\ \omega & = & \sum_{|I|=i} f_I(x_1, ..., x_n) \cdot dx_I. \end{array}$$

Beachte die Schreibweise dx_ν resp. dx_I anstelle von e_ν resp. e_I. Dies ist motiviert durch folgende Formel, welche sich für die Abbildung $d = d_0$ aus der Definition des letzten Paragraphen ergibt

$$\begin{array}{ccl} A^0(\mathcal{U}, \mathbb{R}) & \longrightarrow & A^1(\mathcal{U}, \mathbb{R}) \\ f(x_1, ..., x_n) & \longmapsto & df = \sum_{\nu=1}^n \left(\frac{\partial}{\partial x_\nu} f\right)(x_1, ..., x_n) \cdot dx_\nu. \end{array}$$

Poincare Lemma (ohne Beweis): *Ist $\mathcal{U}$ eine sternförmige offene Teilmenge von $\mathbb{R}^N$, dann ist der Komplex $A^\bullet(\mathcal{U}, \mathbb{R})$ der alternierenden Differntialformen auf $\mathcal{U}$ exakt mit Ausnahme im Grad $i = 0$.*

Definition: Eine offene Teilmenge U von $\mathbb{R}^n$ heißt sternförmig, falls es einen Punkt $x_0 \in U$ gibt so daß mit $x \in U$ auch die Verbindungsgerade $\overline{xx_0}$ in U liegt.

Bemerkung: $A^\bullet(\mathcal{U}, \mathbb{R})$ ist ein Ring bezüglich der Tensorproduktalgebrenstruktur auf $A^\bullet(\mathcal{U}, \mathbb{R}) = C^\infty(\mathcal{U}, \mathbb{R}) \otimes_{\mathbb{R}} \bigwedge^\bullet(\mathbb{R}^n)$. Die Multiplikation in diesem Ring wird durch das $\wedge$-Produkt gegeben. Man zeigt dann leicht die folgende verallgemeinerte **Produktregel**

$$d(\omega \wedge \eta) \ = \ d(\omega) \wedge \eta + (-1)^{|\omega||\eta|} \cdot \omega \wedge d(\eta)$$

für $\omega \in A^i(\mathcal{U}, \mathbb{R})$ und $\eta \in A^j(\mathcal{U}, \mathbb{R})$ mit $|\omega| = i$, $|\eta| = j$.

Die ist eine formale Folgerung aus der Eigenschaft $f_i(r \cdot s) = f_i(r) \cdot s + r \cdot f_i(s)$ der partiellen Ableitungen (Produktformel).

77 Pullbacks

Seien $\mathcal{U} \subsetneqq \mathbb{R}^n$ und $\mathcal{V} \subsetneqq \mathbb{R}^m$ offene Teilmengen und sei $\varphi : \mathcal{V} \to \mathcal{U}$ eine unendlich oft diffenzierbare Abbildung. Für jede Funktion $f \in C^\infty(\mathcal{U}, \mathbb{R})$ ist dann der Pullback $\varphi^*(f) = f \circ \varphi$

$$\begin{tikzcd} \mathcal{V} \arrow[rr, "\varphi"] \arrow[dr, "\varphi^*(f)"'] & & \mathcal{U} \arrow[dl, "f"] \\ & \mathbb{R} & \end{tikzcd}$$

mit $\varphi^*(f) \in C^\infty(\mathcal{V}, \mathbb{R})$ definiert.

Die Zuordnung $f \mapsto \varphi^*(f)$ definiert einen $\mathbb{R}$–Algebrenhomomorphismus

$$\varphi^* : C^\infty(\mathcal{U}, \mathbb{R}) \longrightarrow C^\infty(\mathcal{V}, \mathbb{R}) \ ,$$

den sogenannten Pullback von Funktionen. Es gilt offensichtlich für zusammengesetzte Abbildungen

$$(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$$

wegen $(\varphi \circ \psi)^*(f) = f \circ (\varphi \circ \psi) = (f \circ \varphi) \circ \psi = \psi^*(f \circ \varphi) = \psi^*(\varphi^*(f))$.

Lemma: (ohne Beweis) *Den oben definierte Pullback kann man auf eine einzige Weise zu einer Abbildung*

$$\varphi^* : A^\bullet(\mathcal{U}, \mathbb{R}) \longrightarrow A^\bullet(\mathcal{V}, \mathbb{R})$$

fortsetzen mit folgenden Eigenschaften

(i) φ^* *ist eine Komplexabbildung, d.h.*

$$\varphi^*(d(\omega)) = d(\varphi^*(\omega)) \quad , \quad \varphi^*(A^i(\mathcal{U}, \mathbb{R})) \subsetneqq A^i(\mathcal{V}, \mathbb{R}).$$

(ii) φ^* *ist ein* $\mathbb{R}$*–Algebrenhomomorphismus, welcher im Grad* $i = 0$ *auf* $A^0(\mathcal{U}, \mathbb{R}) = C^\infty(\mathcal{U}, \mathbb{R})$ *mit dem Pullback von Funktionen übereinstimmt.*

<u>Hinweis</u>: Da $A^\bullet(\mathcal{U},\mathbb{R})$ als Ring von $C^\infty(\mathcal{U},\mathbb{R})$ und den dx_1, ..., dx_n erzeugt wird, legt Bedingung (ii) den Pullback φ^* fest, wenn man alle $\varphi^*(dx_\nu)$ kennt. Beachte weiterhin: Ist $\pi_\nu : \mathcal{U} \to \mathbb{R}$ die ν–te Koordinatenfunktion, welche jedem Punkt $x \in \mathcal{U} \subset \mathbb{R}^n$ seine ν-te Koordinate x_ν zuordnet, dann gilt $A^1(\mathcal{U},\mathbb{R}) \ni d(\pi_\nu) = \sum_{i=1}^n \frac{\partial}{\partial x_i}(\pi_\nu)dx_i = dx_\nu$ wegen $\partial x_i(\pi_\nu) = \delta_{i\nu}$. Nennt man die Abbildung π_ν suggestiv x_ν, dann schreibt sich dies

$$d(x_i) = dx_i \quad , \quad (i = 1, ..., n) \ .$$

Wegen Bedingung (i) ist jedenfalls dann für $\varphi = (\varphi_1, ..., \varphi_n)$ (mit $\varphi_\nu \in C^\infty(\mathcal{V},\mathbb{R})$)

$$\varphi^*(dx_\nu) = d(\varphi^*(\pi_\nu)) = d(\varphi_\nu) \ .$$

Daher legt (i) und (ii) den Pullback auf eindeutige Weise fest.

Die funktorielle Eigenschaft der Graßmann Algebra (§73) zeigt umgekehrt, daß es einen solchen Pullback φ^* tatsächlich gibt.

<u>Bemerkung 1</u>: Wegen obiger Formel $dx_i = d(x_i)$ schreibt man allgemeiner für $\omega \in A^i(\mathcal{U},\mathbb{R})$ oft nur $d\omega$ statt $d(\omega)$.

<u>Bemerkung 2 (Verallgemeinerte Kettenregel)</u> : Die Eindeutigkeit der Pullbacks (im Sinn des obigen Lemmas) und der oben erläuterte Spezialfall des Pullbacks von Funktionen führen allgemein zu der Formel

$$(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$$

für den Pullback von alternierenden Differentialformen unter zusammengesetzten Abbildungen.

Grundlagen der Arithmetik

78 Euklidsche Ringe

Wir kehren nun zurück zu den kommutativen Ringen. Jeder Ring wird im folgenden stillschweigend als kommutativ angenommen.

Definition: Ein Ring heißt nullteilerfrei, falls gilt $0 \neq 1$ und wenn aus $ab = 0$ entweder $a = 0$ oder $b = 0$ folgt.

Beispiel:

1) Jeder Körper ist nullteilerfrei.
2) $\mathbb{Z}$ ist nullteilerfrei.
3) Der Polynomring $K[X]$ für Körper ist nullteilerfrei.
4) Jeder Teilring eines nullteilerfreien Ringes ist nullteilerfrei (z.B. $\mathbb{Z} \subseteq \mathbb{Q}$).

Definition: Ein Euklidscher Ring R ist ein kommutativer, nullteilerfreier Ring, versehen mit einer Gradfunktion

$$g : R \setminus \{0\} \to \mathbb{N}_{\geq 0},$$

so daß gilt:

(i) $g(ab) \geq g(a) \qquad \forall a, b \in R \setminus \{0\}$

(ii) Für jedes Paar $a, b \in R$ mit $b \neq 0$ existieren $m, r \in R$ mit $a = mb + r$, wobei entweder $r = 0$ oder $g(r) < g(b)$ gilt.

$$\boxed{a = mb + r \qquad g(r) < g(b) \textit{ oder } r = 0}$$

Beispiele:

1. Jeder Körper K ist für $g(a) := 1,\ \forall a \in K$ ist ein Euklidscher Ring.

2. $R = \mathbb{Z}$ ist nullteilerfrei. Sei $g(r) = |r|$. Für jedes Paar ganzer Zahlen $a, b \in \mathbb{Z} \setminus \{0\}$ gilt

$$g(ab) = |ab| = |a||b| \geq |a| = g(a)$$

wegen $|b| \geq 1$. Jede ganze Zahl $a \in \mathbb{Z}$ schreibt sich in der Form

$$a = mb + r$$

mit $r \in [0, b[$ und $m \in \mathbb{Z}$ ($b \neq 0$ aus $\mathbb{Z}$ gegeben). Es gilt entweder $r = 0$ oder $g(r) = |r| \leq |b| - 1 < b = g(b)$.

3. Sei K ein Körper und $R = K[X]$. Setze

$$g(a_0 + a_1 X + \cdots + a_n X^n) = \max\{n \in \mathbb{N} \mid a_n \neq 0\},$$

falls das Polynom nicht identisch Null ist, also nicht alle $a_i = 0$ sind.

$$g\left(\sum_{i=1}^{n} a_i X^i\right)$$

heißt Grad des Polynoms $P(X) = \sum\limits_{i=1}^{n} a_i X^i$.

In einer der Übungsaufgaben der Vorlesung LAI war nachzuweisen, daß $K[X]$ bezüglich g ein Euklidscher Ring ist.

4. Setze
$$R = Z[i] := \{a + bi \mid a, b \in \mathbb{Z}\} \subseteq \mathbb{C}.$$
R heißt der Ring der ganzen Gauß'schen Zahlen. R ist ein Ring wegen

$$\begin{aligned}(a + bi) + (a' + b'i) &= (a + a') + (b + b')i \\ (a + bi) \cdot (a' + b'i) &= (aa' - bb') + (ab' + a'b)i\end{aligned}$$

Da R ein Unterring des Körpers $\mathbb{C}$ ist, ist R nullteilerfrei. Setze

$$g(a + bi) = N(a + bi) = a^2 + b^2.$$

Beweis von Axiom (i) für $R = \mathbb{Z}[i]$:

$$g(ab) = N(ab) = N(a) \cdot N(b) \geq N(a) = g(a) \qquad \forall a, b \in \mathbb{C} \setminus \{0\},$$

denn es gilt $N(b) \geq 1$ für jede von Null verschiedene ganze Gaußsche Zahl.

<u>Beweis von Axiom (ii)</u>: Gegeben seien ganze Gaußsche Zahlen a, b mit $b \neq 0$. Gesucht sind m und r mit $a = m \cdot b + r$ und $g(r) < g(b)$ oder $r = 0$.

Betrachte $z = \frac{a}{b} \in \mathbb{C}$. Dann gibt es ein $m \in \mathbb{Z}[i]$ so, daß gilt

$$z - m := \tilde{r} \ \in \ Q$$

für den Quader

$$Q = \{x + iy \in \quad \mid |x| \leq 1/2, |x| \leq 1/2\} \ .$$

Jede komplexe Zahl z kann also durch Translation um eine geeignete Gaußsche ganze Zahl m in das Innere des Einheitskreises (um den Nullpunkt) gebracht werden. Für alle solche Punkte gilt dann offensichtlich $N(\tilde{r}) < 1$ oder $\tilde{r} = 0$ hat, denn der Quader Q ist im Inneren des Einheitskreis enthalten.

Setze $r = b \cdot \tilde{r}$. Dann folgt $a - mb = r$ mit $N(r) = N(b) \cdot N(\tilde{r}) < N(b)$ oder $r = 0$. q.e.d.

<u>Bemerkung</u>: Ähnlich zeigt man, daß der Ring der Eisensteinzahlen $\mathbb{Z}[\rho] = \mathbb{Z} + \mathbb{Z} \cdot \rho$, welcher durch Hinzunahme einer dritten Einheitswurzel $\rho \in \mathbb{C}$ entsteht, ein Euklidscher Ring ist. Für den Ring $R = \mathbb{Z} + \mathbb{Z} \cdot \sqrt{-5}$ ist dies aber nicht mehr richtig.

79 Ideale

Ein Maß für die Komplexität eines kommutativen Ringes R sind die R-Untermoduln des Ringes. Solche Untermoduln heißen Ideale des Ringes R.

Ist $a \in R$ beliebig, dann definieren die Vielfachen von a

$$(a) = \{r \in R \mid \exists s \in R \; mit \; r = s \cdot a\} \subseteq R$$

ein solches Ideal. Man nennt (a) das Hauptideal gebildet zum Element a.

Die R-Modul-Eigenschaft von (a) ist klar. Sind nämlich $m_1, m_2 \in (a)$, dann gilt $m_1 = s_1 \cdot a$ und $m_2 = s_2 \cdot a$. Also folgt

$$m_1 + m_2 = s_1 \cdot a + s_2 \cdot a = (s_1 + s_2) \cdot a \in (a),$$

denn $(s_1 + s_2) \in R$. Weiterhin

$$\lambda m_1 = \lambda(s_1 \cdot a) = (\lambda s_1) \cdot a \in (a),$$

denn $\lambda s_1 \in R$.

Beispiele:

1) Sei $R = \mathbb{Z}$, dann bildet die Teilmenge der geraden Zahlen einen $\mathbb{Z}$-Untermodul von $\mathbb{Z}$. Dieser Untermodul ist das von Element $a = 2$ erzeugte Hauptideal (2).

2) $(n) \subseteq \mathbb{Z}$ für $n \in \mathbb{Z}$ ist die Teilmenge aller durch n teilbaren Zahlen, falls $n \neq 0$ in $\mathbb{Z}$ und gleich $\{0\}$, falls $n = 0$ ist.

3) Sei $R = K$ ein Körper. Jedes Ideal I ist entweder $I = K$ oder $I = \{0\}$.

Begründung: Ideale von $R = K$ sind K-Untervektorräume I von K^1. Wir wissen aber nach dem Basisergänzungssatz

$$dim_K(I) \leq dim_K(K) = 1 \; .$$

Also ist entweder $dim_K(I) = 0$ und somit $I = \{0\}$, oder aber $dim_K(I) = 1$ und dann gilt $I = K$. Siehe §13).

Satz: *Jedes Ideal in einem Euklidschen Ring ist ein Hauptideal.*

Beweis: Sei $I \subseteq R$ ein R-Untermodul, d.h. I ist ein Ideal von R. Nach der Definition eines Untermoduls ist I nicht leer. Also entweder $I = \{0\}$, oder es existiert ein weiteres Element $a \neq 0$ in I). Sei oBdA $I \neq \{0\}$ und

$$m = \min\{g(a) \mid 0 \neq a \in I\} \in \mathbb{N}_{\geq 0}.$$

Wähle ein $a \in I$ mit $grad(a) = m$.

Behauptung: $I = (a) = \{sa \mid s \in R\}$, insbesondere ist I also ein Hauptideal.

Beweis: Sei b aus I. Zu zeigen ist: $\exists m \in R$ mit $b = m \cdot a$.
Wegen (ii) gilt

$$b = ma + r \textit{ mit } r = 0 \textit{ oder } g(r) < g(a).$$

Es folgt wegen $r = b - ma$ dann $r \in I$, da b und ma in I liegen. Wegen $g(r) < g(a)$ und da $g(a)$ minimal ist, folgt $r = 0$. Also ist b ein Vielfaches von a wie behauptet: $I \subseteq (a)$.

Andererseits gilt $a \in I \Rightarrow m \cdot a \in I$ für alle $m \in R$.. Somit $(a) \subseteq I$. Es folgt $I = (a)$.

Definition: Ein Hauptidealring ist ein kommutativer, nullteilerfreier Ring mit Einselement, in dem jedes Ideal ein Hauptideal ist.

Folgerung: *Jeder Euklidsche Ring ist ein Hauptidealring.*

80 Elementare Teilertheorie

Definition: Ein Ring R heißt Integritätsbereich, wenn R ein kommutativer nullteilerfreier Ring ist.

Kürzungslemma: *In einem Integritätsbereich folgt aus* $x \cdot a = y \cdot a$ *und* $a \neq 0$ *die gekürzte Identität* $x = y$.

Beweis: Man erhält $(x - y) \cdot a = 0$ und wegen der Nullteilerfreiheit von R und $a \neq 0$ also $x - y = 0$. Das heißt $x = y$.

Quotientenkörper: Die Menge der Äquivalenzklassen von formalen Brüchen x/y (mit $y \neq 0$) von Elementen aus einem Integritätsbereich R bilden mit den üblichen Rechen- und Kürzungsregeln einen Körper $K(R)$, den sogenannten **Quotientenkörper** von R. (Zwei formale Brüche x/y und x'/y' heißen äquivalent, wenn $xy' = x'y$ gilt; aus dem Kürzungslemma folgt die Transitivität dieser Relation!). Der Integritätsbereich R läßt sich mit dem Teilring von $K(R)$ aller Brüche der Gestalt $x/1$ identifizieren.

Sei R ein Integritätsbereich.

Definition: Für Elemente $a, b \in R$ sagen wir a teilt b (und schreiben $a|b$), falls eine der folgenden drei äquivalenten Aussagen gilt:

(i) $b = m \cdot a$ für ein $m \in R$ (b ist ein Vielfaches von a)

(ii) $b \in (a)$

(iii) $(b) \subseteq (a)$

Beweis: (i) $\Leftrightarrow$ (ii) $\Rightarrow$ (iii) ist trivial. Umgekehrt (iii) $\Rightarrow$ (i) folgt aus: $b = 1 \cdot b \in (b) \subseteq (a)$. Also ist $b = m \cdot a$ für ein $m \in R$.

Beispiel:

1) $1|a$ für alle $a \in R$ wegen $(a) \subseteq (1) = R$.

2) $a|0$ für alle $a \in R$ wegen $(0) \subseteq (a)$.

3) $a|b$ und $b|c \Rightarrow a|c$ wegen $b = ma, c = m'b \Rightarrow c = (m'm)a$.

Lemma: *Für von $a \neq 0$ und $b \neq 0$ aus R sind äquivalent:*

(i) $(a) = (b)$

(ii) $a = r \cdot b$ *für eine Einheit* $r \in R^*$

(iii) $b = r' \cdot a$ *für eine Einheit* $r' \in R^*$

Beweis: (ii) $\Rightarrow$ (i): Sei $r \in R^*$, d.h. es gibt ein $r' \in R$ mit $rr' = r'r = 1$. Aus $a = rb$ folgt $(a) \subseteq (b)$. Andererseits gilt $r' \cdot a = r' \cdot r \cdot b = b$. Der umgekehrte Schluß liefert dann $(b) \subseteq (a)$, und somit $(a) = (b)$. (iii) $\Rightarrow$ (i) folgt dann aus Symmetrie!

(i) $\Rightarrow$ (ii), (i) $\Rightarrow$ (iii): Aus (i) folgt $a = r \cdot b = r \cdot (r' \cdot a) = rr' \cdot a$ für gewisse r, r' aus R. Das Kürzungslemma impliziert wegen $a \neq 0$

$$1 = rr' = r'r \ .$$

Somit sind r und r' wie behautet Einheiten des Rings R.

Definition: Ein Teiler a von b heißt echter Teiler, falls gilt

$$(b) \underset{\neq}{\subset} (a) \underset{\neq}{\subset} R.$$

Beispiel: $1|b$, aber 1 ist kein echter Teiler von b. Genauso gilt $b|b$, aber b ist kein echter Teiler von b.

Bemerkung: Sei $a \neq 0$. Dann ist a als Teiler von b mit

$$b = m \cdot a$$

genau dann ein echter Teiler von b, wenn weder a noch m eine Einheit von R ist.

Zum Beweis: $(b) = (a)$ ist nach obigem Lemma äquivalent zu $b = r \cdot a$ für ein $r \in R^*$. Andererseits gilt $b = m \cdot a$ nach Annahme. Das Kürzungslemma impliziert $r = m$, somit ist m eine Einheit im Fall $(b) = (a)$. Ist umgekehrt $m \in R^*$ eine Einheit, dann folgt $(b) = (a)$ nach obigem Lemma.

Analog dazu ist $(a) = R = (1)$ äquivalent zur Aussage $a \in R^*$.

Definition: Ein Element $p \in R$ mit $p \neq 0$ und $p \notin R^*$ heißt Primelement, falls p eine der folgenden äquivalenten Eigenschaften besitzt

(i) p besitzt keine echten Teiler

(ii) Für jede Zerlegung $p = u \cdot v$ von p folgt $u \in R^*$ oder $v \in R^*$.

(iii) Für jedes Hauptideal I mit $(p) \subseteq I \subseteq R$ gilt $I = (p)$ oder $I = R$.

Beispiele:

1) In $R = K[X]$ ist jedes lineare Polynome $X - a_0$ ein Primelement.
2) $X^2 - 2aX + a^2$ ist kein Primelement in $K[X]$.
3) Ist $K = \mathbb{R}$, dann ist $X^2 + 1$ ein Primelement in $\mathbb{R}[X]$.
4) In $R = \mathbb{Z}$ ist ein Element n Primelement genau dann, wenn $|n|$ eine Primzahl ist.

81 Teilertheorie in Hauptidealringen

Von nun an werde zusätzlich zu der Annahme des vorigen Abschnitts angenommen, daß R ein Hauptidealring ist.

Für gegebene Elemente $a_1, ..., a_n$ des Ringes R setzen wir

$$(a_1, ..., a_n) = \left\{ \sum_{i=1}^{n} m_i a_i \mid m_i \in R \right\}.$$

Diese Teilmenge von R ist unter Addition und außerdem unter Skalarmultiplikation mit Elementen aus R abgeschlossen. Also ist $(a_1, ..., a_n)$ ein R-Untermodul von R oder mit anderen Worten ein Ideal. Dies gilt für jeden kommutativen Ring. Da aber R nach Voraussetzung jetzt ein Hauptidealring ist, existiert jetzt aber ein $a \in R$ mit

$$(a_1, ..., a_n) = (a).$$

Die Zahl $a \in R$ ist bis auf eine Einheit eindeutig bestimmt, und man nennt sie $ggT(a_1, ..., a_n)$ oder größter gemeinsamer **T**eiler der Zahlen $a_1, ..., a_n$ aus R.

Analog definiert man $kgV(a_1, ..., a_n) = b$ mittels des Durchschnittideals

$$(a_1) \cap ... \cap (a_n) = (b) \; .$$

Bemerkung: Aus der Charakterisierung von Primelementen im letzten Paragraph sind in einem Hauptidealring folgende Aussagen äquivalent

(i) p ist ein Primelement

(ii) Für jedes I mit $(p) \subseteq I \subseteq R$ gilt $I = (p)$ oder $I = R$.

Lemma 1: *Sei p ein Primelement eines Hauptidealrings und $q \in R$ sei beliebig. Dann sind äquivalent:*

(i) $p|q$

(ii) $(p,q) = (p)$

<u>Beweis</u>: (i) $\Rightarrow$ (ii): $p|q \Rightarrow (q) \subseteq (p) \Rightarrow (p,q) \subseteq (p)$ per Definition des Ideals (p,q). Andererseits ist $(p) \subseteq (p,q)$ immer richtig. Also $(p,q) = (p)$.

(ii) $\Rightarrow$ (i): $(p,q) = (p) \Rightarrow 0 \cdot p + 1 \cdot q = m \cdot p \Rightarrow q = m \cdot p \Rightarrow p|q$

Lemma 2: (Teilerlemma) *Sei p ein Primelement in einem Hauptidealring. Dann gilt:*

$$\textit{Aus } p|q\tilde{q} \textit{ folgt } p|q \textit{ oder } p|\tilde{q}.$$

<u>Beweis</u>: Sei $I = ggT(p,q)$. Dann gilt $(p) \subset I \subset R$. Wegen obiger Bemerkung ist also entweder $ggT(p,q) = (1) = R$ oder $ggT(p,q) = (p)$. Im letzten Fall $ggT(p,q) = (p)$ folgt $p|q$ (letztes Lemma). OBdA können wir daher $ggT(p,q) = 1$ annehmen. Also existieren $\alpha, \beta \in R$ mit

$$\alpha \cdot p + \beta \cdot q = 1 \ .$$

Daraus folgt wegen $q\tilde{q} = mp$

$$\tilde{q} = (\alpha \cdot p + \beta \cdot q) \cdot \tilde{q} = \alpha \cdot p \cdot \tilde{q} + \beta \cdot m \cdot p = (\alpha \cdot \tilde{q} + \beta \cdot m) \cdot p,$$

also $p|\tilde{q}$ wie behauptet.

<u>Bemerkung</u>: Der <u>Gaußsche</u> Ring $\mathbb{Z} + \mathbb{Z} \cdot i$ und der <u>Eisensteinsche</u> Ring $\mathbb{Z} + \mathbb{Z} \cdot \rho$ mit $\rho^3 = 1, Im(\rho) > 0$ sind beides euklidsche Ringe, somit Hauptidealringe und das Teilerlemma gilt in diesen Ringen. Dies ist nicht richtig für den Ring $R = \mathbb{Z} + \mathbb{Z} \cdot \sqrt{-5}$ wegen $(1+2\sqrt{-5})(1-2\sqrt{-5}) = 3 \cdot 7$. Man kann zeigen, daß 3, 7, $1+2\sqrt{-5}$ und $1-2\sqrt{-5}$ Primelemente in R sind, aber 3 weder $1+2\sqrt{-5}$ noch $1-2\sqrt{-5}$ teilt. Betrachte die Normen! Der Ring $\mathbb{Z} + \mathbb{Z} \cdot \sqrt{-5}$ ist also kein Hauptidealring!

82 Primzahlzerlegung

Sei R ein kommutativer Hauptidealring.

Satz (Eindeutigkeit der Zerlegung in Primelemente): *Jede Zerlegung $a = p_1 \cdot ... \cdot p_r$ einer Zahl $0 \neq a \in R$ in Primelemente $p_1, ..., p_r$ bestimmt die Primelemente $p_i (i = 1, ..., r)$ eindeutig bis auf die Reihenfolge und bis auf Abänderungen derselben durch Einheiten.*

Beweis: Sei $q = p_1 \cdot ... \cdot p_r = q_1 \cdot ... \cdot q_s$ (p_i, q_j Primelemente). Wir schließen durch Induktion nach $n = \max(r, s)$. OBdA sei $r \geq s$ und $r \geq 1$. Aus der Relation folgt insbesondere:

$$p_1 | q_1 \cdot ... \cdot q_s.$$

Durch Iteration von Lemma 2 aus dem letzten Abschnitt folgt $p_1 | q_i$ für ein $i, 1 \leq i \leq s$. Dies impliziert wiederum $(p_1) = (q_i)$, da q_i als Primelemente keine echte Teiler besitzt. Es folgt $p_1 = \epsilon \cdot q_i$ mit einer Einheit ϵ und man erhält aus dem Kürzungslemma

$$(\epsilon p_2) \cdot ... \cdot p_r \ = \ q_1 \cdot ... \cdot \cancel{q_i} \cdot ... \cdot q_s \ .$$

Die Zahl der Primfaktoren dieser Identität ist $\max(r - 1, s - 1) < \max(r, s)$. Mit Hilfe der Induktionsannahme folgt daraus der Satz.

Sei R wie bisher ein Hauptidealring. Dann gilt

Satz: (Existenz der Primzahlzerlegung) *Jedes Element $r \in R, r \neq 0$ besitzt eine Zerlegung $r = p_1 \cdot ... \cdot p_s$ in ein Produkt von endlich vielen Primelementen, oder r ist eine Einheit.*

Korollar: *In einem Hauptidealring besitzt jede von Null verschiedene Nichteinheit eine "eindeutige" Zerlegung in Primelemente.*

Zum Beweis der Existenz einer Primelementzerlegung benötigen wir folgendes

Lemma: *Sei R ein Hauptidealring. Dann wird jede aufsteigende Kette von Idealen*

$$(a_1) \subseteq \cdots \subseteq (a_n) \subseteq (a_{n+1}) \subseteq \cdots$$

stationär, d.h. es gibt einen Index i_0, so daß für alle $i \geq i_0$ gilt

$$(a_i) = (a_{i+1}).$$

Beweis: Setze $I := \bigcup_{i=1}^{n} (a_i) \subseteq R$. Offensichtlich ist I ein R-Untermodul von R. Da R ein Hauptidealring ist, gilt $I = (a)$ für ein $a \in R$. Wegen

$$a \in I = \bigcup_{i=1}^{n} (a_i)$$

existiert dann ein i_0, so daß $a \in (a_{i_0})$.

$$a = m \cdot a_{i_0} \in (a_{i_0})$$

Jetzt folgt

$$I = (a) \subseteq (a_{i_0}) \subseteq (a_i) \subseteq I$$

für alle $i \geq i_0$. Also $I = (a_i)$ für $i \geq i_0$. Damit ist das Lemma bewiesen.

Beweis von der Existenz der Primelementzerlegung: Sei r ein von Null verschiedene Nichteinheit in R. Sei

$$r = \prod_{i=1}^{n} r_i^{(n)} \quad r_i^{(n)} \in R, r_i^{(n)} \notin R^*$$

eine Zerlegung von r der Länge n. Sind nicht alle $r_i^{(n)}$ Primfaktoren, dann existiert ein i (oBdA $i = n$), so daß $r_i^{(n)}$ eine weitere Zerlegung in das Produkt zweier Nichteinheiten besitzt.

$$r_n^{(n)} = r_n^{(n+1)} \cdot r_{n+1}^{(n+1)}$$

Setzt man $r_i^{(n+1)} = r_i^{(n)}$, dann erhält man eine Zerlegung wie oben von der Länge $n+1$. Dieser Prozeß bricht nach endlich vielen Schritten ab, denn andernfalls wäre

$$(r) = \left(r_1^{(1)}\right) \underset{\neq}{\subset} \left(r_2^{(2)}\right) \underset{\neq}{\subset} \left(r_3^{(3)}\right) \underset{\neq}{\subset} \cdots$$

eine nichtstationäre Kette von Teilern. Dies ist nicht möglich, wie im letzten Lemma gezeigt wurde. Bricht der Prozeß dagegen ab, ist die Zerlegung die gewünschte Zerlegung von r in Primelemente.

83 Der Chinesische Restsatz

Seien $a, b \in R$ mit $ggT(a,b) = (1) = (R)$, das heißt a, b seien teilerfremd. Dann existieren per Definition $\alpha, \beta \in R$ mit

$$\alpha a + \beta b = 1.$$

Setze $A = \alpha a$ und $B = \beta b$. Dann gilt

$$A + B = 1 \quad \text{und} \quad A \in (a) \quad B \in (b) \ .$$

Andererseits gilt für den $kgV(a,b) \ = \ (a) \cap (b)$ dann außerdem

$$kgV(a,b) \ = \ (a \cdot b) \ ,$$

denn die Teilerfremdheit von a und b impliziert wegen der Existenz und Eindeutigkeit der Primfaktorzerlegung in R, daß eine Zahl z durch a und durch b teilbar ist genau dann, wenn z durch $a \cdot b$ teilbar ist. Dies liefert die folgende nichttriviale Identität

Lemma: *Für teilerfremde Zahlen a, b in einem Hauptidealring gilt*

$$(a) \cap (b) = (a \cdot b) \ .$$

Daraus folgern wir jetzt

Satz: *Für teilerfremde Zahlen a, b in einem Hauptidealring R gilt*

$$R/(a \cdot b) \ \cong \ R/(a) \oplus R/(b).$$

Beweis: Betrachte die R-lineare Abbildung

$$\varphi : \ R \ \to \ R/(a) \oplus R/(b)$$

welche durch Restklassenbildung wie folgt

$$\varphi(r) = ([r], [r]) = (r + (a), r + (b))$$

definiert ist. Offensichtlich gilt

0) φ ist R-linear.

1) $Kern(\varphi) = (a) \cap (b)$ beziehungsweise auch $Kern(\varphi) = (a \cdot b)$ nach obigem Lemma, da nach Annahme a und b teilerfremd sind.

2) φ ist surjektiv: Seien A und B gewählt wie am Anfang des Paragraphen! Dann gilt $\varphi(A) = ([0],[1])$, denn $A \in (a)$ und $A = 1 - B \in [1]$. Analog zeigt man $\varphi(B) = ([1],[0])$.

   Es folgt $\varphi(r_2 A + r_1 B) = (r_2[0], r_2[1]) + (r_1[1], r_1[0]) = ([r_1],[r_2])$. Das heißt jedes Element $([r_1],[r_2]) \in R/(a) \oplus R/(b)$ liegt im Bild von φ.

Der Isomorphiesatz aus §60 impliziert wegen 2) daher wie behauptet

$$R/(a) \oplus R/(b) \;\cong\; R/Kern(\varphi) \;\overset{1)}{=}\; R/(a \cdot b)\ .$$

Beispiel: $\mathbb{Z}/(6) \cong \mathbb{Z}/(3) \oplus \mathbb{Z}/(2)$

Korollar: *Sei* $a = \prod_{i=1}^{r} p_i^{n_i} \cdot \epsilon$ *mit einer Einheit* $\epsilon \in R^*$ *und seien paarweise teilerfremde Primelemente* $p_1, ..., p_r$ *gegeben. Dann gilt*

$$R/(a) \;\cong\; \bigoplus_{i=1}^{r} R/\left(p_i^{n_i}\right).$$

84 Äquivalenz von Matrizen

Sei R ein kommutativer Ring. Eine Matrix $U \in M_{n,n}(R)$ mit Einträgen in R heißt invertierbar, wenn eine Umkehrmatrix $U^{-1} \in M_{n,n}(R)$ mit Einträgen in R existiert mit

$$U \circ U^{-1} = U^{-1} \circ U = E \ .$$

U ist invertierbar genau, dann wenn $det(R)$ eine Einheit von R ist. Dies folgt unmittelbar aus der Cramerschen Regel.

Die Menge der invertierbaren Matrizen in $M_{n,n}(R)$ bezeichnen wir mit

$$\mathcal{G}\ell(n, R) \ .$$

Matrizen in $\mathcal{G}\ell(n, R)$ bilden eine Gruppe bezüglich der Matrixmultiplikation.

Bemerkung: Für $U \in \mathcal{G}\ell(n, R)$ gilt auch $\Lambda^i(U) \in \mathcal{G}\ell(N, R)$ für die Graßmann Potenzen $\Lambda^i(U)$. Nämlich $\Lambda^i(U^{-1})$ ist invers zu $\Lambda^i(U)$ und hat Koeffizienten in R. Siehe §86.

Definition: Zwei Matrizen $M_1, M_2 \in M_{n,m}(R)$ heißen äquivalent, wenn invertierbare Matrizen $U \in \mathcal{G}\ell(n, R)$ und $V \in \mathcal{G}\ell(m, R)$ existieren mit der Eigenschaft

$$M_1 = U \cdot M_2 \cdot V.$$

Man schreibt dann $M_1 \sim M_2$. Dies verallgemeinert die Definition aus §25.

Wie man leicht nachprüfen kann, gilt

$$M_1 \sim M_1$$

$$M_1 \sim M_2 \Rightarrow M_2 \sim M_1$$

$$M_1 \sim M_2, M_2 \sim M_3 \Rightarrow M_1 \sim M_3.$$

Der Beweis ist vollkommen analog zum Beweis des bereits behandelten Körperfalls. Siehe §25. Auch der nächste Abschnitt ist eine Verallgemeinerung von §25.

85 Elementarteiler (Existenz)

Satz: *Sei R ein Hauptidealring. Dann ist jede Matrix $M \in M_{n,m}(R)$ äquivalent zu einer Matrix $\mathbb{E}$ in Normalform:*

$$\mathbb{E} = \left(\begin{array}{ccc|c} e_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & e_n & \end{array}\right) \quad bzw. \quad \mathbb{E} = \left(\begin{array}{ccc} e_1 & & 0 \\ & \ddots & \\ 0 & & e_m \\ \hline & & \\ & 0 & \end{array}\right)$$

$$m \geq n \qquad\qquad n \geq m$$

für gewisse $e_1, ..., e_r \in R$ mit der Eigenschaft

$$e_1 | e_2 | \cdots | e_{r-1} | e_r \qquad r = \min\{n, m\}.$$

Die Zahlen $e_1, ..., e_r$ heißen <u>*Elementarteiler*</u> *der Matrix M. Sie sind bis auf Einheiten in R^* eindeutig durch die Matrix M bestimmt.*

<u>Bemerkung</u>: Der Fall $e_i = 0$ ist möglich. Beachte, daß jede Zahl $m \in R$ Null teilt. Die $e_i = 0$ stehen daher am Ende.

<u>Beweis des Satzes</u>: Wir führen den Beweis nur in dem Spezialfall durch, wo R ein Euklidscher Ring ist. Sei $g : R \to \mathbb{N}_{\geq 0}$ die Gradfunktion von R.
Invertierbare $n \times n$-Matrizen mit Einträgen in R sind die Elementarmatrizen:

1) $$Diag(a_1, ..., a_n) = \begin{pmatrix} a_1 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & a_n \end{pmatrix} \text{ mit } a_1, ..., a_n \in R^*$$

2) $$E_{\nu,\mu}(\lambda) = \begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 \\ 0 & 1 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ 0 & \vdots & \ddots & & & & \vdots & 0 \\ \vdots & \vdots & & \ddots & & & \vdots & \vdots \\ \vdots & \vdots & & & \ddots & \lambda & \vdots & \vdots \\ 0 & \vdots & & & & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & 1 \end{pmatrix} \text{ mit } \lambda \in R \text{ und } \nu \neq \mu$$

3) Permutationsmatrizen $P(\sigma) = (P(\sigma)_{ij})$

$$P(\sigma)_{ij} = \begin{cases} 1 & j = \sigma(i) \\ 0 & j \neq \sigma(i) \end{cases} \text{ für } \sigma \in S_n$$

Dies zeigt man wie in §24 im Körperfall.

Linksmultiplikationen mit solchen Matrizen (beziehungsweise Rechtsmultipliaktionen mit den entsprechenden $m \times m$-Matrizen) zeigt wie in §31, daß durch folgende Operationen Matrizen in äquivalente Matrizen übergeführt werden:

1. Multiplikation der Spalten (bzw. Zeilen) mit Einheiten
2. Addition des λ-fachen der ν-ten Zeile (Spalte) zur μ-ten Zeile (Spalte) $(\nu \neq \mu)$
3. Permutation der Spalten (Zeilen).

Der Beweis des Elementarteilersatzes erfolgt durch Ausräumen der gegebenen Matrix M mit Hilfe der Operationen 1.-3. g hat Werte in $\mathbb{N}_{\geq 0}$. Wir wählen unter allen Matrizen, welche zur ursprünglichen Matrix äquivalent sind, eine für die $g(a_{11})$ minimal ist (insbesondere ist dann $a_{11} \neq 0$ für $M \neq 0$; aber die Nullmatrix hat bereits Normalform).
OBdA können wir also M selbst als eine dieser minimalenMatrizen annehmen.

Folgerung: *Jede der Zahlen a_{1i} und a_{i1} ist durch a_{11} teilbar.*

Beweis: $a_{11} \nmid a_{1i} \Rightarrow a_{1i} = m \cdot a_{11} + r$ mit $r \neq 0$ und $g(r) < g(a_{11})$. Addiert man das $-m$-fache der ersten Spalte zur i-ten Spalte und vertauscht dann die erste und die i-te Spalte,

$$\begin{pmatrix} a_{11}\cdots a_{1i}\cdots \\ \\ * \end{pmatrix} \xrightarrow{2.\ \lambda=-m} \begin{pmatrix} a_{11}\cdots \overbrace{a_{1i}-m\cdot a_{11}}^{=r}\cdots \\ \\ * \end{pmatrix}$$

$$\xrightarrow{3.} \begin{pmatrix} r\cdots \\ \\ * \end{pmatrix} g(r) < g(a_{11})$$

erhält man einen Widerspruch zur Minimalität von $g(a_{11})$. q.e.d. Analog schließt man für a_{i1}.

Durch Umformungen mittels der Operationen 1.,2. und 3. können wir also jetzt die Matrix überführen in die Gestalt

$$\left(\begin{array}{c|c} a_{11} & 0 \\ \hline 0 & * \end{array}\right)$$

<u>Behauptung</u>: *Jeder Eintrag der Matrix* $\left(\begin{array}{c|c} a_{11} & 0 \\ \hline 0 & * \end{array}\right)$ *ist durch* a_{11} *teilbar.*

<u>Beweis</u>: Angenommen $a_{11} \not| a_{ij} \Rightarrow a_{ij} = m\cdot a_{11}+r$ mit $r \neq 0$ und $g(r) < g(a_{11})$. Die Operationen

$$\begin{pmatrix} a_{11} & \cdots & 0 & \cdots \\ \\ 0 & & a_{ij} & \\ \\ \end{pmatrix} \xrightarrow{2.\ \lambda=1} \begin{pmatrix} a_{11} & \cdots & a_{11} & \cdots \\ \\ 0 & & a_{ij} & \\ \\ \end{pmatrix}$$

$$\xrightarrow{2.\ \lambda=-m} \begin{pmatrix} a_{11} & \cdots & a_{11} & \cdots \\ \\ * & & r & \\ \\ \end{pmatrix}$$

$$\xrightarrow{3.} \begin{pmatrix} r & * \\ \\ * & * \\ \\ \end{pmatrix}$$

liefern dann wegen $g(r) < g(a_{11})$ einen Widerspruch zur Minimalität. Somit ist die letzte Behauptung bewiesen.

Es folgt

$$\begin{pmatrix} a_{11} & 0 \\ & \\ 0 & * \end{pmatrix} = a_{11} \cdot \begin{pmatrix} 1 & 0 \\ & \\ 0 & N \end{pmatrix}$$

mit einer Matrix $N \in M_{n-1,m-1}(R)$. Per Induktion hat N oBdA Normalform. Setzt man $e_1 = a_{11}$, erhält man:

$$\left(\begin{array}{cccc|c} e_1 & & & 0 & \\ & e_1e_1' & & & \\ & & e_1e_2' & & 0 \\ & & & \ddots & \\ 0 & & & & \end{array}\right) \qquad m \geq n.$$

Setzt man nun $e_i = e_1 e_{i-1}'$ für $i = 2, ..., n$, dann erhält man:

$$\left(\begin{array}{cccc|c} e_1 & & & 0 & \\ & e_2 & & & \\ & & \ddots & & 0 \\ 0 & & & e_n & \end{array}\right) \qquad m \geq n.$$

a_{11} ist der erste Elementarteiler.

Der Beweis der Eindeutigkeit der Elemetarteilerzerlegung wird im nächsten Abschnitt gegeben. Vorab sei aber das folgende Verfahren zur Bestimmung der Elementarteiler mitgeteilt, welches auch die Eindeutigkeit derselben impliziert.

Ist $M = (i_{\nu\mu})$, dann gilt

$$\begin{array}{rcl} e_1 & = & ggT(i_{11}, ..., i_{nm}) \\ e_1 \cdot e_2 & = & ggT(det \text{ aller } 2 \times 2 \text{ Teilmatrizen von } M\ ) \\ e_1 \cdot e_2 \cdot e_3 & = & ggT(det \text{ aller } 3 \times 3 \text{ Teilmatrizen von } M\ ) \\ \vdots & & \\ e_1 \cdot ... \cdot e_n & = & ggT(det \text{ aller } n \times n \text{ Teilmatrizen von } M\ ) \end{array}$$

Der Beweis dieser Aussage sei dem ambitionierten Leser als Übungsaufgabe überlassen. Hierbei sind die $\nu \times \nu$ Teilmatrizen zu verstehen als diejenigen Matrizen, die durch Streichen von $n - \nu$ Zeilen und $n - \nu$ Spalten entstehen!

Hinweis: Die Aussage ist klar, wenn M in Normalform vorliegt. Benutze $\Lambda^i(U) \circ \Lambda^i(V) = \Lambda^i(U \circ V)$ und $U \in \mathcal{G}\ell(n,R), V \in \mathcal{G}\ell(m,R) \Rightarrow \Lambda^i(U) \in \mathcal{G}\ell(N,R), \Lambda^i(V) \in \mathcal{G}\ell(M,R)$ um den allgemeinen Fall auf diesen Fall zurückzuführen.

Beispiel: $M = \begin{pmatrix} 5 & 3 \\ 2 & 2 \end{pmatrix}$ hat Elementarteiler $e_1 = 1$, $e_2 = det(M) = 4$.

86 Elementarteiler (Eindeutigkeit)

In diesem Abschnitt sei R wieder ein Hauptidealring.

Für U in $M_{n,n}(R)$ gilt offensichtlich $det(U) \in R$. Eine Matrix U mit dieser Eigenschaft heißt unimodular.

Für eine Matrix U in $M_{n,n}(R)$ schreiben wir $U \in \mathcal{G}\ell(n,R)$, wenn eine zu U inverse Matrix U^{-1} in $M_{n,n}(R)$ existiert. In diesem Fall ist wegen $det(U) \cdot det(U^{-1}) = 1$ die Zahl $det(U)$ eine Einheit in R^*, also U unimodular. Die Umkehrung gilt auch, insbesondere bilden die unimodularen Matrizen eine Gruppe

Lemma: *Es gilt $U \in \mathcal{G}\ell(n,R) \Longleftrightarrow det(U) \in R^*$ für $U \in M_{n,n}(R)$.*

Dies ergibt sich unmittelbar aus der Formel für die inverse Matrix U^{-1} in Termen der adjungierten Matrix $\tilde{U} \in M_{n,n}(R)$.

Lemma: *Für $M, N \in M_{n,m}(R)$ seien*

$$(a) = ggT(M_{ij}) \quad und \quad (b) = ggT(N_{ij})$$

die ggT's der Matrixkoeffizienten. Sind M und N äquivalente Matrizen, dann gilt

$$(a) = (b) \ .$$

Beweis: Es gilt $M = a \cdot M_1$ für ein $M_1 \in M_{n,m}(R)$. Aus $N = UMV = a \cdot UM_1V$, folgt $a|b$, da UM_1V Koeffizienten in R besitzt. Wegen Symmetrie folgt genauso $b|a$, also $(a) = (b)$.

Sei nun $M \in M_{n,m}(R)$ eine beliebige Matrix. Bezeichne $r := \min\{n,m\}$. Nach §85 ist M äquivalent zu einer Elementarteilermatrix

$$M \sim \mathbb{E} = \begin{pmatrix} e_1 & & 0 \\ & \ddots & \\ 0 & & e_r \end{pmatrix} \quad \text{mit } e_1 \mid ... \mid e_r \text{ und } r = \min\{n,m\}.$$

Folgerung: *Es gilt $\Big(ggT(M_{11}, ..., M_{nm})\Big) = (e_1)$.*

<u>Die Graßmannkonstruktion</u>: $M \in M_{n,m}(R)$ ist eine R–lineare Abbildung

$$M : R^m \to R^n$$

zugeordnet: $M(b_i) = \sum_{j=1}^n M_{ji} b_j$ für $i = 1, ..., m$. Hierbei seien b_i bzw. b_j die kanonischen Basisvektoren des R^m bzw. R^n. Die Graßmannkonstruktion (für $k = 0, 1, 2, ...$) induziert R-lineare Abbildungen

$$\Lambda^k(R^m) \stackrel{\Lambda^k(M)}{\longrightarrow} \Lambda^k(R^n) ,$$

so daß gilt

$$\Lambda^k(id) = id \quad \text{und} \quad \Lambda^k(M_1 \cdot M_2) = \Lambda^k(M_1) \cdot \Lambda^k(M_2) .$$

Bezüglich der induzierten R-Basen b_I lassen sich diese Abbildungen wieder durch Matrizen beschreiben, welche Koeffizienten in R besitzen. Dies folgt aus $\Lambda^k(M)(b_I) = \Lambda^k(M)(b_{i_1} \wedge ... \wedge b_{i_k}) = M(b_{i_1}) \wedge ... \wedge M(b_{i_k}) = \sum_{j_1=1}^n M_{j_1 i_1} b_{j_1} \wedge ... = \sum_{J \subset \{1,...,n\}} \Lambda^k(M)_{IJ} b_J$ durch ausmultiplizieren und zusammenfassen.

<u>Beispiel</u>: Für $n = m = 4, k = 2$ und $M = diag(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ ist $\Lambda^2(M)$ gleich

$$\begin{array}{cc}
\begin{array}{cccccc} b_1 \wedge b_2 & b_1 \wedge b_3 & b_1 \wedge b_4 & b_2 \wedge b_3 & b_2 \wedge b_4 & b_3 \wedge b_4 \end{array} & \\
\begin{pmatrix}
\lambda_1\lambda_2 & 0 & 0 & 0 & 0 & 0 \\
0 & \lambda_1\lambda_3 & 0 & 0 & 0 & 0 \\
0 & 0 & \lambda_1\lambda_4 & 0 & 0 & 0 \\
0 & 0 & 0 & \lambda_2\lambda_3 & 0 & 0 \\
0 & 0 & 0 & 0 & \lambda_2\lambda_4 & 0 \\
0 & 0 & 0 & 0 & 0 & \lambda_3\lambda_4
\end{pmatrix} &
\begin{array}{c} b_1 \wedge b_2 \\ b_1 \wedge b_3 \\ b_1 \wedge b_4 \\ b_2 \wedge b_3 \\ b_2 \wedge b_4 \\ b_3 \wedge b_4 \end{array}
\end{array}$$

denn $b_i \wedge b_j \mapsto \lambda_i \cdot b_i \wedge \lambda_j \cdot b_j = \lambda_i\lambda_j \cdot b_i \wedge b_j$. Allgemeiner

Lemma: *Für $M = diag(\lambda_1, .., \lambda_r)$ ist $\Lambda^k(M)$ eine Diagonalmatrix mit den Diagonaleinträgen $\prod_{i \in I} \lambda_i$ an der Diagonalenstelle $I \subset \{1, ..., n\}$ mit $|I| = k$.*

Lemma: *Wenn $U : R^n \to R^n$ invertierbar in $M_{n,n}(R)$ ist, ist $\Lambda^k(U) : \Lambda^k(R^n) \to \Lambda^k(R^n)$ invertierbar in $M_{\binom{n}{k},\binom{n}{k}}(R)$.*

<u>Beweis</u>: Weil U invertierbar ist, gibt es ein $U_1 \in M_{n,n}(R)$ mit $U_1 U = U U_1 = id$. Es folgt $\Lambda^k(U_1) \cdot \Lambda^k(U) = \Lambda^k(U_1 \cdot U) = \Lambda^k(id) = id$. Also ist $\Lambda^k(U_1)$ invers zu $\Lambda^k(U)$, und ist ganzzahlig weil $\tilde{U}$ ganzzahlig ist.

Aus diesem Lemma und der Formel $\Lambda^k(UMV) = \Lambda^k(U) \cdot \Lambda^k(M) \cdot \Lambda^k(V)$ ergibt sofort als

Folgerung: *Sind M und N äquivalent, dann auch $\Lambda^k(M)$ und $\Lambda^k(N)$.*

Nach diesen Vorbereitungen kommen wir nun zu der entscheidenden Aussage dieses Abschnitts

Satz: *Für die Elementarteiler e_i einer Matrix M gilt*

$$\prod_{i=1}^{k} e_i \;=\; ggT(\textit{Koeffizienten von } \Lambda^k(M)) \;.$$

Insbesondere sind diese Produkte (für $k = 1, 2, .., r$) und damit auch die Elementarteiler $e_1, .., e_r$ selber bis auf Einheiten eindeutig durch die Matrix M bestimmt.

Beweises des Satzes: Da wie nun gezeigt $\Lambda^k(\mathbb{E})$ und $\Lambda^k(M)$ äquivalent sind, haben sie denselben Koeffizienten ggT. Aus den oben gegebenen Formeln für die Einträge der Diagonalmatrix $\Lambda^k(\mathbb{E})$ und der Teilerbedingung $e_1|e_2...$ ergibt sich der Koeffizienten ggT von $\Lambda^k(\mathbb{E})$ als die Zahl $\prod_{i=1}^{k} e_i = ggT_{I,|I|=k}(\prod_{i \in I} e_i)$.

Hauptidealringe und ihre Moduln

87 Kern, Bild, Kokern

Genau so wie im Körperfall kann man nun auch lineare Gleichungssysteme mit Koeffizienten in einem Hautidealring R betrachten und für ein solches Gleichungssystem nach Lösungen mit Koeffizienten in R fragen. Wird das Gleichungssystem durch die Matrix $M \in M_{m,n}(R)$ beschrieben, hat man analog zum Körperfall eine zugeordnete R-lineare Abbildung

$$\begin{aligned} M : R^n &\longrightarrow R^m \\ x &\longmapsto M \cdot x. \end{aligned}$$

Wie im Körperfall ist die Lösungsmenge des linearen Gleichungssystems entweder leer oder ist von der Gestalt $x_0 + Kern(M)$. Es stellt sich dann natürlich die Frage nach der Struktur der R-Moduln $Kern(M) \subset R^n$ und $Bild(M) \subset R^m$? Es sind Untermoduln von freien Moduln. Tatsächlich sind beide Moduln selbst wieder freie R-Moduln. Den Beweis kann man zum Beispiel dadurch führen, daß man auf den Fall reduziert, wo M eine Elementarteilermatrix ist. Dieser Fall ist, wie wir sehen werden, von besonderem Interesse.

Die Aussage folgt aber auch aus dem im Anhang bewiesenen allgemeinen Satz. Wir erhalten in jedem Fall

Satz: *Für ein R-lineares Gleichungssystem $M : R^n \to R^m$ mit $M \in M_{m,n}(R)$ gilt*

- *$Kern(M)$ ist ein freier R-Modul vom Rang $\leq n$*
- *$Bild(M)$ ist ein freier R-Modul vom Rang $\leq m$.*

Frage: Wann hat nun $M \cdot x = b$ überhaupt eine Lösung (zu gegebenem $b \in R^m$)? Was ist das Hindernis dafür?

Antwort: Die Lösbarkeit des inhomogenen Gleichungssystems ist aus tautologischen Gründen äquivalent zum Verschwinden der Restklasse $[b]$ im $Kokern(M) = R^m/Bild(M)$. Der R–Modul $Kokern(M)$ hängt aber nur von M ab und nicht von b. Man nennt ihn daher auch den ¨Hindernisraum" oder "Obstruktionsraum". Dieser Kokern ist im allgemeinen aber kein freier R-Modul mehr, enthält aber interessante Informationen über das Gleichungssystem.

Beispiel: Für $M\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x-y \end{pmatrix}$ ist der Hindernisraum $\simeq \mathbb{Z}/2\mathbb{Z}$.

In der Tat hängt der Hindernisraum (bis auf Isomorphie von R-Moduln) nur von den Elementarteilern von M ab. Im obigen Beispiel sind die Elementarteiler gerade 1 und 2, und dies gibt dann $\mathbb{Z}/2\mathbb{Z}$ für den Hindernisraum. Die Rechnung zeigt aber mehr, und erlaubt es diesen Modul in Termen der Elementarteiler direkt auszurechnen:

Sei

$$M = U \cdot \mathbb{E} \cdot V$$

mit $U \in GL(m,R)$, $V \in \mathcal{G}\ell(n,R)$ und Elementarteilermatrix $\mathbb{E}$. Beachte

$v \in Kern(M) \Leftrightarrow M \cdot v = 0 \Leftrightarrow U \cdot \mathbb{E} \cdot V \cdot v = 0 \Leftrightarrow \mathbb{E} \cdot V \cdot v = 0 \Leftrightarrow V(v) \in Kern(\mathbb{E})$. Also folgt

Lemma: *Die Abbildung*

$$\begin{array}{ccc} Kern(M) & \xrightarrow{\sim} & Kern(\mathbb{E}) \\ v & \longmapsto & V(v) \end{array}$$

ist ein R–linearer Isomorphismus.

Beispiel: Für $\mathbb{E} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ gilt $Kern(\mathbb{E}) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| 2 \cdot x = 0 \right\}$.

Somit ist $Kern(\mathbb{E})$ wegen der Nullteilerfreiheit von R gleich dem von dem 2. und 3. Basisvektor aufgespannten freien R-Untermodul vom R^3, also isomorph zu R^2.

Bemerkung: Im allgemeinen folgt genauso $Kern(\mathbb{E}) \simeq R^l$, wobei

$$l = \begin{cases} \#(\text{Elementarteiler } e_i = 0) & \text{falls } n \leq m \\ \#(\text{Elementarteiler } e_i = 0) + n - m & \text{falls } n \geq m. \end{cases}$$

Den Rang des freien Moduls $Kern(M)$ kann man daher daher direkt von den Elementarteilern von M ablesen.

Lemma: *Man hat R-Modulisomorphismen* $Bild(M) \cong Bild(\mathbb{E})$ *und*

$$Kokern(M) \cong Kokern(\mathbb{E})$$

in Termen der Elementarteilermatrix $\mathbb{E}$ *von* M*.*

Beweis: Für das Bild und Kokern gilt $Kokern(M) = R^m/Bild(M)$ sowie

$$Bild(M) = Bild(U \cdot \mathbb{E} \cdot V) = U \cdot \mathbb{E} \cdot \underbrace{V(R^n)}_{=R^n} = (U\mathbb{E})(R^n) \ .$$

Man erhält einen Isomorphismus $U : Bild(\mathbb{E}) = \mathbb{E}(R^n) \to Bild(M)$ mit Umkehrabbildung

$$U^{-1} : (U\mathbb{E})(R^n) \to \mathbb{E}(R^n) \ .$$

Dies ergibt das Diagramm

$$\begin{tikzcd}
0 \arrow[r] & (U\mathbb{E})(R^n) \arrow[r] \arrow[d, "\simeq"', "U^{-1}"] & R^m \arrow[r] \arrow[d, "\simeq"', "U^{-1}"] & Kokern(M) \arrow[r] \arrow[d, dotted, "\simeq"', "\exists!"] & 0 \\
0 \arrow[r] & \mathbb{E}(R^n) \arrow[r] & R^m \arrow[r] & R^m/\mathbb{E}(R^n) \arrow[r] & 0
\end{tikzcd}$$

Die R-Isomorphismen U^{-1} induzieren einen kanonisch bestimmten R-Isomorphismus zwischen den R-Moduln $Kokern(M)$ und $Kokern(\mathbb{E}) = R^m/\mathbb{E}(R^n)$.

Appendix

Satz: *Sei R ein Hauptidealring. Dann ist ein R Untermodul M des freien R–Moduls R^n wieder ein freier R–Modul, und sein Rang ist $\leq n$.*

Beweis: (Induktion nach n) Der Fall $n = 1$ ist im wesentlichen gerade die Annahme, daß R ein Hauptidealring ist. Für den Induktionsschritt betrachte die Abbildungen $i : R^{n-1} \to R^n$ und $pr_n : R^n \to R$

$$i\begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ 0 \end{pmatrix} \qquad \text{und} \qquad pr_n \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_n \ .$$

Es gilt $M \subset R^n$ nach Annahme. Sei I das Bild von M unter pr_n. Als R–Untermodul von R ist $I = pr_n(M) \subset R$ ein Ideal. Weil R Hauptidealring ist, gibt es ein $a \in R$ mit $I = (a)$. ObdA ist $a \neq 0$, da sonst $M \subset R^{n-1} = Kern(pr_n)$ und wir per Induktion schließen. Nun hat man folgendes kommutative Diagramm, dessen oberste Zeile offensichtlich (!) exakt ist

$$\begin{tikzcd}
0 \arrow[r] & R^{n-1} \arrow[r, "i"] & R^n \arrow[r, "pr_n"] & R \arrow[r] & 0 \\
0 \arrow[r] & i(R^{n-1}) \cap M \arrow[r] \arrow[u, hook] & M \arrow[r] \arrow[u, hook] & (a) \arrow[r] \arrow[u, hook] & 0
\end{tikzcd}$$

Die untere Zeile ist ebenfalls exakt: Die Exaktheit bei $i(R^{n-1}) \cap M$ und bei (a) ist klar. Zur Exaktheit in der Mitte

$$Kern\,(pr_n|_M) = M \cap Kern\,(pr_n) = M \cap i(R^{n-1}) \ .$$

Per Induktion ist nun $i(R^{n-1}) \cap M \subset i(R^{n-1}) \simeq R^{n-1}$ frei vom Rang $\leq n-1$. $(a) = R \cdot a \simeq R$ ist als R-Modul frei vom Rang 1 mit Basiselement a.

Der Satz folgt daher aus dem folgenden Lemma

Lemma: *Sei $0 \to F_1 \xrightarrow{\varphi} M \xrightarrow{\pi} F_2 \to 0$ eine kurze exakte Sequenz von R–Moduln. Sind F_1, F_2 freie R–Moduln vom Rang n_1 resp. n_2 (aus $\mathbb{N}$), dann ist M frei vom Rang $n_1 + n_2$.*

Beweis: Sei b_1, ..., b_{n_1} eine Basis von F_1 und $\tilde{b}_1$, ..., $\tilde{b}_{n_2}$ eine Basis von F_2. Da π surjektiv ist, existieren b_{n_1+1}, ..., $b_{n_1+n_2} \in M$ mit $\pi(b_{n_1+i}) = \tilde{b}_i$. Das Lemma folgt dann aus der

Behauptung: $\varphi(b_1)$, ..., $\varphi(b_{n_1})$,b_{n_1+1}, ..., $b_{n_1+n_2}$ bilden eine Basis von M.

Beweis: Lineare Unabhängigkeit: Wenn für $\lambda_i \in R$ gilt

$$0 = \lambda_1\varphi(b_1) + ... + \lambda_{n_1}\varphi(b_{n_1}) + \lambda_{n_1+1}b_{n_1+1} + ... + \lambda_{n_1+n_2}b_{n_1+n_2},$$

so folgt durch Anwenden von π

$$0 = 0 + \lambda_{n_1+1}\tilde{b}_1 + ... + \lambda_{n_1+n_2}\tilde{b}_{n_2}.$$

Weil die $\tilde{b}_i$ linear unbhängig sind, folgt daraus $0 = \lambda_{n_1+1} = ... = \lambda_{n_1+n_2}$.
Also ist die ursprüngliche Relation von der Gestalt $0 = \sum_{i=1}^{n} \lambda_i\varphi(b_i) = \varphi\left(\sum_{i=1}^{n_1} \lambda_i b_i\right)$.
Da φ injektiv ist, gilt $0 = \sum_{i=1}^{n_1} \lambda_i b_i$, Wegen der linearen Unabhängigkeit der b_i hat das zur Folge, daß auch die restlichen $\lambda_1 = ... = \lambda_{n_1} = 0$ verschwinden.

Erzeugendensystem: Sei $m \in M$ beliebig. Weil $\tilde{b}_i$ ein Erzeugendensystem von F_2 bilden, hat man für geeignete $\lambda_i \in R$

$$F_2 \ni \pi(m) = \lambda_1\tilde{b}_1 + ... + \lambda_{n_2}\tilde{b}_{n_2} = \sum_{i=1}^{n_2} \lambda_i\pi(b_{n_1+i}) = \pi\left(\sum_{i=1}^{n_2} \lambda_i b_{n_1+i}\right).$$

Also $\pi(m - \sum_{i=1}^{n_2} \lambda_i b_{n_1+i}) = 0$ oder

$$m - \sum_{i=1}^{n_2} \lambda_i b_{n_1+i} \ \in \ Kern(\pi) = \varphi(F_1)\ .$$

Die b_i bilden ein Erzeugendensystem von F_1. Also gibt es $\mu_i \in R$ mit

$$m - \sum_{i=1}^{n_2} \lambda_i b_{n_1+i} = \varphi(\sum_{j=1}^{n_1} \mu_j b_j) = \sum_{j=1}^{n_1} \mu_j\varphi(b_j)\ .$$

Folglich gilt $m = \sum_{j=1}^{n_1} \mu_j\varphi(b_j) + \sum_{i=1}^{n_2} \lambda_i b_{n_1+i}$. q.e.d.

Bemerkung: Betrachtet man in diesem Appendix – für ein fest gewähltes Primelement π von R – nur noch R-Moduln M mit der Eigenschaft $\pi \cdot M = 0$ und ersetzt 'freier R-Modul vom Rang n' durch 'R-Modul vom Typ $(R/\pi R)^n$', dann übertragen sich sowohl die Aussage des Satzes als auch sein Beweis.

88 Struktursätze

Satz: *Jede endliche abelsche Gruppe G ist isomorph zu einem (direkten) Produkt von zyklischen Gruppen.*

Variante: *(mittels chinesischem Restsatz) G ist sogar isomorph zu einem Produkt von Gruppen $\mathbb{Z}/p^r\mathbb{Z}$ mit Primzahlen $p \in \mathbb{N}$ (und $r \in \mathbb{N}$).*

Dies sind Spezialfälle im Fall $R = \mathbb{Z}$ eines in diesem Abschnitt formulierten allgemeinen Struktursatzes für endlich erzeugte R-Moduln über einem Hauptidealring.

Definition:

1) Ein R–Modul heißt endlich erzeugt, wenn er ein endliches Erzeugendensystem hat.

2) Ein R–Modul heißt zyklisch, wenn er von einem einzigen Element erzeugt wird.

Präsentationen: Sei m_1,...,$m_n \in M$ ein endliches Erzeugendensystem eines R–Moduls M. Die Abbildung

$$\begin{array}{ccc} R^n & \xrightarrow{\pi} & M \\ \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} & \longmapsto & r_1m_1 + ... + r_nm_n \end{array}$$

ist eine wohldefinierte R-lineare Abbildung. Weil $\{m_1, .., m_n\}$ den R–Modul M erzeugen, ist π surjektiv. Für $F := Kern(\pi)$ erhält man die exakte Sequenz

$$0 \longrightarrow F \xrightarrow{i} R^n \xrightarrow{\pi} M \longrightarrow 0.$$

Da R ein Hauptidealring ist, ist F als Untermodul von R^n freier Modul vom Rang $r \leq n$, d.h. es gibt einen Isomorphismus $\delta : F \simeq R^r$. Betrachte

$$\mathbb{M} = i \circ \delta^{-1} : R^r \simeq F \hookrightarrow R^n \ .$$

Als Zusammensetzung injektiver Abbildungen ist $\mathbb{M}$ wieder injektiv. Der Kokern von $\mathbb{M}$ ist isomorph zu M, denn

$$\begin{tikzcd}
0 \arrow[r] & F \arrow[r, "i"] \arrow[d, "\simeq"', "\delta"] & R^n \arrow[r, "\pi"] \arrow[d, equal] & M \arrow[r] \arrow[d, dotted, "\simeq"', "\exists!"] & 0 \\
0 \arrow[r] & R^r \arrow[r, "\mathbb{M}"'] & R^n \arrow[r] & Kokern(\mathbb{M}) \arrow[r] & 0
\end{tikzcd}$$

ist kommutativ mit exakten Zeilen. Also hat man den

Präsentationssatz: *Jeder endlich erzeugte R-Modul M über einem Hauptidealring R ist isomorph zu einem Modul $Kokern(\mathbb{M})$ für eine geeignete Matrix $\mathbb{M} \in M_{n,r}(R)$ ($n \geq r \in \mathbb{N}$ geeignet).*

Andererseits wissen wir, daß $\mathbb{M}$ ähnlich ist zu seiner Elementarteilermatrix $\mathbb{E}$ und daß gilt $Kokern(\mathbb{M}) \simeq Kokern(\mathbb{E})$. Also stellt sich die Frage, wie der Kokern aussieht für Matrizen der Gestalt

$$\mathbb{E} = \begin{pmatrix} e_1 & & 0 \\ & \ddots & \\ 0 & & e_r \\ & 0 & \end{pmatrix} \in M_{n,r}(R).$$

Die zugehörige Abbildung $\mathbb{E} : R^r \to R^n$ ist

$$\mathbb{E}(\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix}) = \begin{pmatrix} e_1x_1 \\ \vdots \\ e_rx_r \\ 0 \end{pmatrix} .$$

<u>Behauptung</u>:

$$Kokern(\mathbb{E}) = R^n/Bild(\mathbb{E}) \simeq R/e_1R \oplus ... \oplus R/e_rR \oplus \underbrace{R/0 \oplus ... \oplus R/0}_{n-r \text{ Kopien}} .$$

<u>Beweis</u>: Die Abbildung

$$\begin{array}{ccc} R^n & \overset{p}{\longmapsto} & R/e_1R \oplus ... \oplus R/0R \\ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \longmapsto & ([x_1], ..., [x_n]) \end{array}$$

ist surjektiv und hat als Kern den Untermodul

$$\left\{ \left. \begin{pmatrix} e_1 \cdot x_1 \\ \vdots \\ e_r \cdot x_r \\ 0 \end{pmatrix} \right| x_i \in R \right\} = Bild(\mathbb{E}).$$

Aus der Eindeutigkeit des Quotienten folgt daher

$$R/e_1R \oplus ... \oplus R/0R \;\simeq\; R^n/Bild(\mathbb{E}) \;.$$

Bemerkung: $R/0R = R$ ist zyklisch und R/e_iR ist zyklisch.

Struktursatz: *Jeder endlich erzeugte R–Modul M über einem Hauptidealring R ist isomorph zu einer direkten Summe von zyklischen R-Moduln*

$$M \simeq \bigoplus_{i=1}^{n} R/a_iR$$

für $a_i \in R$ mit $a_1 \mid ... \mid a_n$. Hierbei können und wollen wir außerdem annehmen, daß keine der Zahlen a_i eine Einheit in R ist.

Im nächsten Abschnitt wird gezeigt, daß dann n und die $a_i, i = 1, .., n$ bis auf Einheiten durch (die Isomorphieklasse von) M eindeutig festgelegt sind. Wir nennen die Zahlen a_i daher auch die Elementarteiler des R-Moduls M.

89 Beweis der Eindeutigkeit

Wir wollen zeigen, daß die Elementarteiler eines endlich erzeugten R-Moduls über einem Hauptidealring R bis auf Einheiten eindeutig bestimmt sind.

Gibt es einen Isomorphismus von R-Moduln

$$N = \bigoplus_{i=1}^{n} R/(a_i) \quad \cong \quad N' = \bigoplus_{j=1}^{m} R/(b_j)$$

sowie $a_1 \mid ... \mid a_n$ und $b_1 \mid ... \mid b_m$, müssen wir $n = m$ sowie $(a_i) = (b_i)$ zeigen für alle $i = 1, .., n$. Nach Annahme ist keine der Zahlen a_i, b_j eine Einheit.

Aus Symmetriegründen können wir N und N' vertauschen und annehmen

$$r = \#\{a_i \neq 0\} \ \geq \ r' = \#\{b_j \neq 0\} \ .$$

Sei k das Minimum der Zahl k der Primteiler (mit Vielfachheiten) aller $a_i \neq 0$ bzw. $b_j \neq 0$. Wir benutzen Induktion nach k. Der Induktionsanfang $k = 0$ ist der Fall $a_1 = .. = a_n = 0$ bzw. $b_1 = ... = b_m = 0$. Somit ist der Modul $N \cong N'$ frei und alle a_i und b_j sind Null. Wegen $N = R^n$ und $N' = R^m$ und dem Satz aus dem Appendix von §87 folgt dann $m \leq n$ (und dann $n = m$ aus Symmetrie).

Die Strategie: Im Fall $k > 0$ wählen wir einen Primteiler π von a_1 und betrachten den Untermodul $N[\pi] \subset N$ aller π-Torsionselemente

$$N[\pi] \ = \ \{x \in N \mid \pi \cdot x = 0\} \ .$$

Es ist klar, daß ein Isomorphismus $\psi : N \to N'$ von R-Moduln einen Isomorphismus der π-Torsionsmoduln $\psi : N[\pi] \to N'[\pi]$ induziert. Wir erhalten somit

$$\begin{tikzcd}
0 \arrow[r] & N[\pi] \arrow[r] \arrow[d, "\cong"] & N \arrow[r] \arrow[d, "\cong"] & N/N[\pi] \arrow[r] \arrow[d, dotted, "\cong"] & 0 \\
0 \arrow[r] & N'[\pi] \arrow[r] & N' \arrow[r] & N'/N'[\pi] \arrow[r] & 0
\end{tikzcd}$$

Die Strategie ist jetzt die folgende: Wir setzen die Elementarteiler von N in Beziehung zu denen von $N[\pi]$ und $N/N[\pi]$ und zeigen per Induktion, daß $N/N[\pi] \cong N'/N'[\pi]$ sowie $N[\pi] \cong N'[\pi]$ dieselben Elementarteiler besitzen.

Berechnung von $N[\pi]$: Unsere Moduln N sind direkte Summen von Moduln $N = \bigoplus_i N_i$. Somit ist klar $N[\pi] = \bigoplus_i N_i[\pi]$. Für die Berechnung können wir uns somit auf den Fall zyklischer Moduln

$$M = R/aR$$

beschränken. Hier gibt es drei Fälle.

Erster Fall (π ist teilerfremd zu a). Dann folgt wegen $1 = \alpha \cdot \pi + \beta \cdot a$ sowie $a \cdot m = \pi \cdot m = 0$, $m \in M[\pi]$ sofort $m = 1 \cdot m = 0$ für alle Elemente $m \in M[\pi]$. Also folgt

$$M[\pi] = 0 \ .$$

Zweiter Fall (π teilt $a = 0$). Dann ist $M = R$ und wegen der Nullteilerfreiheit

$$M[\pi] = 0 \ .$$

Dritter Fall (π teilt $a \neq 0$). Dann gilt $M[\pi] \ = \ \{m = r \cdot \tilde{a} \ mod \ (a) \mid r \in R\}$ für die Zahl $\tilde{a} = a/\pi$. Somit ist $M[\pi]$ ein zyklischer R-Modul erzeugt von der Restklasse $[\tilde{a}]$. Die Abbildung $R \to M[\pi]$, welche $r \in R$ auf $r \cdot [\tilde{a}]$ abbildet, ist surjektiv und hat als Kern das Ideal (π). Aus dem Isomorphisatz folgt daher im dritten Fall

$$M[\pi] \ \cong \ R/(\pi) \ .$$

Korollar: *Ist M ein endlich erzeugter R-Modul über einem Hauptidealring, dann ist $M[\pi]$ wieder endlich erzeut als R-Modul. Es gilt $M[\pi] = M$ genau dann wenn gilt*

$$M \ \cong \ (R/\pi R)^\rho$$

Die natürliche Zahl ρ ist durch die Isomorphieklasse von M eindeutig bestimmt.

Beweis: Alle Aussagen folgen im zyklischen Fall aus den obigen Fallunterscheidungen. Im allgemeinen resultieren sie dann – mit Ausnahme der Eindeutigkeit von ρ – aus dem Struktursatz für endlich erzeugte R-Moduln durch Zerlegung in zyklische Moduln. Die Eindeutigkeit von ρ folgt aus der Bemerkung im Appendix von §87. q.e.d.

Wir kehren nun zu unserem Induktionsschluß zurück:

Zur Erinnerung: $N[\pi] \cong N'[\pi]$ und $N/N[\pi] \cong N'/N'[\pi]$.

Aus $N/N[\pi] = \bigoplus_i N_i/N_i[\pi]$ und den obigen Fallunterscheidungen folgt

$$N[\pi] \ \cong \ (R/\pi R)^\rho \qquad , \qquad \rho = \#\{i \mid a_i \neq 0 \ , \ \pi | a_i\}$$

sowie

$$N/N[\pi] \ \cong \ \bigoplus_{i=1}^n R/(\frac{a_i}{\pi}) \ .$$

Beachte, daß für die Indizes i mit $a_i = \pi$ die a_i/π zu Einheiten werden, und die dazu gehörigen Summanden Null werden.

Es genügt die letzte Formel für $N/N[\pi]$ im zyklischen Fall nachzurechnen: Für zyklisches $N_i = R/(a_i)$ ist wegen $\pi | a_i$ der R-Modul $N_i[\pi]$ erzeugt von $\tilde{a}_i = a_i/\pi$ und somit $N_i/N_i[\pi] \cong R/(\tilde{a}_i)$.

Nach Wahl von π ist aber $N[\pi] \neq 0$ und $r = \rho$. Somit folgt aus $N[\pi] = (R/\pi R)^r$ und $N'[\pi] \cong N[\pi]$ dann $N'[\pi] = (R/\pi)^{\rho'} \neq 0$ mit $\rho' = r$ (siehe letztes Korollar!). Wegen

$$r \geq r' \geq \rho' = r$$

folgt somit dann $r = r' = \rho = \rho'$. Aus der analogen Formel für $N'[\pi]$

$$\rho' = \#\{j \mid b_j \neq 0, \ \pi | b_j\}$$

folgt somit $\pi | b_j$ für alle $b_j \neq 0$. Damit gilt auch die zur obigen Formel analoge Formel für $N'/N'[\pi]$.

Aus dem Isomorphismus $N/N[\pi] \cong N'/N'[\pi]$ folgt daher

$$\bigoplus_{i=1}^n R/(\frac{a_i}{\pi}) \ \cong \ \bigoplus_{j=1}^m R/(\frac{b_j}{\pi})$$

und per Induktion (!) die Gleichheit all der Elementarteiler, welche jeweils von (π) verschieden sind. Benutzt man die dadurch erhaltene Information, schließt man dann die noch fehlende Gleichheit der Anzahl der Elementarteiler vom Typ (π) aus $r = r'$.

90 Der $K[X]$–Modul eines Endomorphismus

Der Polynomring $R = K[X]$ über einem kommutativen Körper K wird durch den Polynomgrad zu einem euklidischen Ring.

Einsetzungen: Sei a ein Element einer K–Algebra A mit 1_A. Dann gibt es genau einen K–Algebrenhomomorphismus $\varphi = \varphi_a$

$$\varphi : \ K[X] \longrightarrow A$$

mit der Eigenschaft $\varphi(X) = a$.

Dies ist sehr einfach zu sehen! $\varphi(1) = 1_A$, $\varphi(X^2) = a^2$ ... und die K-Linearität erzwingen für $p(X) = \sum \lambda_i \cdot X^i \in K[X]$

$$\varphi(p(X)) = \varphi(\sum_{i=0}^{n} \lambda_i \cdot X^i) = \sum_{i=0}^{n} \lambda_i \cdot a^i \ .$$

Daß die so definierte Abbildung φ_a wohldefiniert ist und ein Ringhomomorphismus, ist leicht nachzurechnen.

Der Modul (V, M): Sei $A = End(V)$ für einen K-Vektorraum V und $a = M$ für einen Endomorphismus M von V. Durch obigen Einsetzungshomomorphismus wird V zu einem $K[X]$–Modul, den wir mit (V, M) bezeichnen. Die Modulmultiplikation von $p(X) = \sum_{i=0}^{n} \lambda_i \cdot X^i \ \in \ K[X]$ auf V wird gegeben durch

$$\begin{array}{ccl} K[X] \times M & \longrightarrow & M \\ (p(X), v) & \longmapsto & P(a)v = \sum_{i=0}^{n} \lambda_i \cdot a^i(v) \end{array}$$

Beispiel: Sei $V = K^2$ und $M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Für $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in K^2$

$$\begin{array}{rcl} 1 \cdot v & = & v \\ X \cdot v & = & M \cdot v = \begin{pmatrix} x_2 \\ 0 \end{pmatrix} \\ X^2 \cdot v & = & M^2 v \ = 0 \cdot v = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ & \vdots & \\ X^n \cdot v & = & 0 \qquad (n \geq 2) \end{array}$$

M hat zwei K–Erzeuger e_1, e_2 (die Standardbasisvektoren von K^2). Also ist M endlich erzeugt als K–Modul und erst recht als $K[X]$–Modul. Wegen $X \cdot e_2 = e_1$ ist der Modul sogar ein zyklischer $K[X]$–Modul erzeugt von e_2. Nach dem Struktursatz gilt daher $M \cong K[X]/a(X)$. Man zeigt leicht $a(X) = X^2$. Also $M \cong K[X]/X^2$, was als K–Vektorraum isomorph ist zu $K + K \cdot X$ ist.

Ist der K-Vektorraum V endlich dimensional, dann ist offensichtlich (V, M) als $K[X]$–Modul endlich erzeugt. Nach dem Stuktursatz für Moduln über dem Hauptidealring $R = K[X]$ in §88 gilt: Jeder endlich erzeugte R–Modul ist isomorph zu einem R–Modul der Gestalt

$$R/(a_1) \oplus R/(a_2) \oplus ... \oplus R/(a_r)$$

mit $a_1 \mid a_2 \mid ... \mid a_r \in R$. Wir können diesen Struktursatz auf die $K[X]$–Moduln (V, M) anwenden.

Bemerkung: Für $a = 0$ ist $R/(a) = R$. Im Fall $R = K[X]$ ist dies unendlichdimensional als K–Vektorraum. Da (V, M) aber endlichdimensional als K-Vektorraum ist, kommt somit das Polynom $a_i = 0$ in der Zerlegung in zyklische Moduln nicht vor.

Korollar: *Sei V ein endlich dimensionaler K-Vektorraum, $M \in End(V)$ und (V, M) der zugeordnete $K[X]$-Modul. Dann gibt es nichtverschwindende, obdA normierte Polynome $a_i = a_i(X) \in K[X]$ mit $a_1 \mid ... \mid a_r$ vom Grad $n_i \geq 1$, so daß gilt*

$$(V, M) \simeq \bigoplus_{i=1}^{n} R/(a_i).$$

Ein Polynom heißt normiert, wenn es von der Form ist

$$a_i(X) = X^{n_i} + (\textit{Polynom vom Grade } < n_i) \; .$$

91 Zyklische Moduln

Sei im folgenden $a(X) = X^n + \alpha_1 X^{n-1} + ... + \alpha_n$ ein normiertes Polynom mit Koeffizienten in K vom Grad $n \geq 1$.

Die Restklassen der Monome in dem Quotientenmodul $W = K[X]/(a(X))$ des Polynomrings $K[X]$ nach dem von $a(X)$ aufgespannten Hauptideal, sind Mengen der Form

$$\begin{aligned} [1] &= 1 + a(X) \cdot K[X] \\ [X] &= X + a(X) \cdot K[X] \\ &\vdots \\ [X^{n-1}] &= X^{n-1} + a(X) \cdot K[X] \end{aligned}$$

Lemma: *Faßt man $W = K[X]/(a(X))$ als K–Vektorraum auf, dann bildet $[1]$, $[X]$, $[X^2]$, ... $[X^{n-1}]$ eine K–Basis von W.*

Beweis: (Erzeugendensystem) Sei $[p(X)] \in W$ mit Repräsentant $p(X) \in K[X]$. Division mit Rest gibt Polynome r und m mit $p(X) = m(X) \cdot a(X) + r(X)$ und $grad(r) < grad(a)$ oder $r \equiv 0$ gilt. Für $r \equiv 0$, d.h. $p(X) = m(X) \cdot a(X)$, ist $[p(X)] = [0]$; oder $r(X) = \rho_1 X^{n-1} + ... + \rho_n$ hat Grad kleiner als $n = grad(a)$. Dann liegt $[p(X)] = [r(X)] = \rho_1[X^{n-1}] + ... + \rho_n[1]$ im K-Aufspann der Vektoren $[X^{n-1}]$, ..., $[1]$.

Lineare Unabhängigkeit: Für λ_0, ..., $\lambda_{n-1} \in K$ so daß

$$\begin{aligned} \lambda_0[1] + \lambda_1[X] + ... + \lambda_{n-1}[X^{n-1}] &= 0 \\ \Leftrightarrow \quad [\lambda_0 + \lambda_1 X + ... + \lambda_{n-1} X^{n-1}] &= 0 \\ \Leftrightarrow \quad \lambda_{n-1} X^{n-1} + ... + \lambda_1 X + \lambda_0 &= m(X) \cdot a(X) \end{aligned}$$

(für ein $m \in K[X]$) folgt aus Gradgründen $m(X) = 0$ (wegen $grad(a(X)) = n$) und damit $\lambda_0 = ... = \lambda_{n-1} = 0$.

Betrachte die K-lineare Selbstabbildung von $W = K[X]/(a(X))$

$$\begin{array}{rcl} \varphi : W & \longrightarrow & W \\ v & \longmapsto & X \cdot v \end{array}$$

Die Matrix zu φ bezüglich der Basis $([1], [X], [X^2], ..., [X^{n-1}])$ besitzt die sogenannte Jordan–Normalform

$$\begin{array}{cc} & \begin{array}{cccccc} [1] & [X] & [X^2] & \cdots & & [X^{n-1}] \end{array} \\ M = & \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -\alpha_n \\ 1 & 0 & 0 & \cdots & 0 & -\alpha_{n-1} \\ 0 & 1 & 0 & \cdots & 0 & -\alpha_{n-2} \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & & 0 & -\alpha_2 \\ 0 & 0 & \cdots & & 1 & -\alpha_1 \end{pmatrix}, \end{array} \quad \begin{array}{c} [1] \\ [X] \\ [X^2] \\ \vdots \\ [X^{n-2}] \\ [X^{n-1}] \end{array}$$

denn $\varphi([X^i]) = X \cdot [X^i] = [X^{i+1}]$ für $i = 0, 1, ..., n-2$, und

$$\begin{aligned} \varphi([X^{n-1}]) &= [X^n] = [a(X) - \alpha_1 X^{n-1} - ... - \alpha_n] \\ &= -\alpha_1 [X^{n-1}] - ... - \alpha_{n-1}[X] - \alpha_n [1] \ . \end{aligned}$$

Bemerkung: Wir haben also dem zyklischen $K[X]$-Modul W eine Matrix M zugehördnet. Diese beschreibt die Operation von $X \in K[X]$ auf dem $W = K[X]/(a(X))$ zugrundeliegenden K-Vektorraum. Offensichlich gilt daher

$$W \cong (K^n, M)$$

mit den Bezeichnungen des letzten Paragraphen.

Lemma 1: *Das charakteristische Polynom der Matrix M zu φ ist*

$$\chi_M(t) = t^n + \alpha_1 t^{n-1} + ... + \alpha_{n-1} t + \alpha_n.$$

Lemma 2: $M^n + \alpha_1 M^{n-1} + ... + \alpha_{n-1} M + \alpha_n E \equiv 0$

Beweis von Lemma 1: Laplaceentwicklung nach der ersten Spalte.

Beweis von Lemma 2: Zu zeigen ist, daß $X^n + \ldots + \alpha_n = a(X)$ die Nullabbildung ist auf $W = k[X]/(a(X))$. Dazu rechnet man nach, daß Multiplizieren mit $a(X)$ auf W die Nullabbildung ist: $a(X) \cdot [P] = [a(X) \cdot P] = [0]$.

Wir erhalten als Anwendung den folgenden

Satz: (Cayley–Hamilton) *Sei $M \in M_{n,n}(K)$, K ein Körper und $\chi(t)$ das charakteristische Polynom von M. Dann gilt*

$$\chi(M) = 0.$$

Beispiel: Für $M = \begin{pmatrix} 2 & 3 \\ 5 & 1 \end{pmatrix}$ gilt $\chi(t) = t^2 - 3t - 13$. Also

$$M^2 - 3M - 13E = \begin{pmatrix} 19 & 9 \\ 15 & 16 \end{pmatrix} - \begin{pmatrix} 6 & 9 \\ 15 & 3 \end{pmatrix} - \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} .$$

Beweis des Satzes: Sei $V = K^n$. Für den $K[X]$-Modul (V, M) gilt

$$(V, M) \simeq \bigoplus_{i=1}^{r} \underbrace{k[X]/(a_i(X))}_{W_i} \simeq \bigoplus_{i=1}^{r} (W_i, M_i) ,$$

wobei die M_i in Jordan–Normalformen sind. Die Matrix M beschreibt die Skalarmultiplikation mit X. Bezüglich der Zerlegung in zyklische Moduln hat M daher als Matrix Blockmatrizengestalt $M = Diag(M_1, \ldots, M_r)$. Für Blockdiagonalmatrizen M (§38) gilt $\chi_M(t) = \chi_{M_1}(t) \cdots \chi_{M_r}(t)$. Weil $\chi_{M_i}(M_i) \equiv 0$ auf W_i und wegen $\chi_{M_i}(t) \mid \chi_M(t)$ folgt $\chi_M(M_i) = 0$ für $i = 1, \ldots, r$. Also ist $\chi_M(M) = 0$ auf allen W_i und somit auch auf V.

92 Jordan–Normalform

Gegeben seien Endomorphismen $M : V \to V$ bzw. $M' : V' \to V'$ von endlichdimensionalen K–Vektorräumen V, V', sowie die dazu gebildeten $K[X]$–Moduln (V, M) und (V', M').

Frage: Was bedeutet in der Sprache der K-Vektorräume, daß eine Abbildung

$$\varphi : (V, M) \to (V', M')$$

$K[X]$–linear ist?

Antwort: $\varphi : V \to V'$ ist einerseits notwenig K–linear und zusätzlich gilt $\varphi(Xv) = X\varphi(v)$ für alle $v \in V$, das heißt

$$\varphi \circ M = M' \circ \varphi \ ,$$

oder als kommutatives Diagramm geschrieben

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d, "M"'] & V' \arrow[d, "M'"] \\ V \arrow[r, "\varphi"'] & V'. \end{tikzcd}$$

Ist umgekehrt $\varphi : V \to V'$ eine K–lineare Abbildung, so daß obiges Diagramm kommutiert, dann definiert φ eine $k[X]$–lineare Abbildung zwischen den $k[X]$–Moduln (V, M) und (V', M').

Begründung: Aus der K–Linearität von φ sowie $\varphi(Mv) = M'\varphi(v)$ (für alle $v \in V$) folgt

$$\varphi(\sum_i \lambda_i X^i v) \overset{\text{Def.}}{=} \varphi(\sum_i \lambda_i M^i v) = \sum_i \lambda_i M'^i \varphi(v) = (\sum_i \lambda_i X^i)\varphi(v) \ .$$

Somit folgt $\varphi(rv) = r\varphi(v)$ für alle $r \in K[X]$.

Folgerung: Die Zerlegung $(V, M) \simeq \bigoplus_{i=1}^{r}(V_i, M_i)$ von $k[X]$–Moduln in zyklische Moduln bedeutet daher in der Sprache der K-linearen Algebra die Existenz einer K–linearen bijektiven Abbildung φ, welche das Diagramm kommutativ macht

$$\begin{tikzcd} V \arrow[r, "\varphi", "\sim"'] \arrow[d, "M"'] & \bigoplus_{i=1}^{r} V_i \arrow[d, "\bigoplus_i M_i"] \\ V \arrow[r, "\sim", "\varphi"'] & \bigoplus_{i=1}^{r} V_i . \end{tikzcd}$$

Wenn man Basen B_i von V_i wählt um V_i mit $K^{\dim(V_i)}$ und M_i mit einer Matrix zu idendifizieren, so hat die $\bigoplus_i M_i$ zugeordnete Matrix Blockdiagonalgestalt. Das heißt, wir haben via φ letzlich damit auch eine Basis von V gefunden, bezüglich der M Blockdiagonalgestalt bekommt.

Satz: *Sei $K = \mathbb{C}$ und sei $M \in M_{n,n}(\mathbb{C})$. Dann existiert eine Basiswechselmatrix $T \in \mathcal{G}\ell(n, K)$, so daß gilt*

$$TMT^{-1} = \begin{pmatrix} M_1 & & 0 \\ & \ddots & \\ 0 & & M_s \end{pmatrix}$$

mit (sogenannten) Jordanblöcken der Gestalt

$$M_i = \begin{pmatrix} \lambda_i & 0 & 0 & \cdots & & 0 \\ 1 & \lambda_i & 0 & \cdots & & 0 \\ 0 & 1 & \lambda_i & \cdots & & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & & 1 & \lambda_i \end{pmatrix} \quad , \qquad \lambda_i \in \mathbb{C}.$$

<u>Beweis</u>: Sei $V = K^n$. Nach dem Struktursatz von §88 ist der Modul (V, M) als $K[X]$–Modul direkte Summe von zyklischen Moduln. Mit dem chinesischen Restsatz kann man die zyklischen Summanden als $K[X]$–Moduln weiter zerlegen in $K[X]$-Moduln der Gestalt $\bigoplus_i K[X]/p_i^{e_i}$, wobei $p_i(X)$ die Primpolynome sind

(obdA paarweise teilerfremd). Zusammengefaßt gibt es also Primpolynome $\pi_i = \pi_i(X)$ in $K[X]$, so daß

$$(V, M) \overset{K[X]}{\simeq} \bigoplus_{i=1}^{s} K[X]/(\pi_i^{e_i}) \ .$$

Jetzt wird ausgenützt, daß für $K = \mathbb{C}$ der Fundamentalsatz der Algebra gilt. Somit hat jedes normierte Primpolynom $\pi(X) \in \mathbb{C}[X]$ die Gestalt $\pi(X) = X - \lambda$ für ein geeignetes $\lambda \in \mathbb{C}$.

Also kann man sich nun auf Moduln der Gestalt $\mathbb{C}[X]/((X-\lambda)^n)$ konzentrieren. Eine seiner Basen (als $\mathbb{C}$–Vektorraum) besteht aus den Restklassen $[1]$, $[X]$, $[X^2]$, ..., $[X^{n-1}]$. Also ist auch $[1]$, $[X-\lambda]$, $[(X-\lambda)^2]$, ..., $[(X-\lambda)^{n-1}]$ eine Basis, wie man leicht zeigt. Bezüglich dieser Basis wird die Multiplikation mit X durch folgende $n \times n$–Matrix beschrieben

$$\begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 1 & \lambda & 0 & \cdots & 0 \\ 0 & 1 & \lambda & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \lambda \end{pmatrix},$$

denn der i–te Basisvektor $[(X-\lambda)^i]$ wird abgebildet auf

$$\begin{aligned} X \cdot [(X-\lambda)^i] &= (X-\lambda) \cdot [(X-\lambda)^i] + \lambda \cdot [(X-\lambda)^i] \\ &= [(X-\lambda) \cdot (X-\lambda)^i] + \lambda \cdot [(X-\lambda)^i] \\ &= [(X-\lambda)^{i+1}] + \lambda \cdot [(X-\lambda)^i]. \end{aligned}$$

Behauptung: *Auf der Diagonalen der Jordan–Normalform einer Matrix $M \in M_{n,n}(\mathbb{C})$ stehen die Eigenwerte von M.*

Beweis: Mit den Bezeichnungen des Satzes ist das charakteristische Polynom

$$\begin{aligned} \chi_M(t) &= det(M - t \cdot E) = det(TMT^{-1} - t \cdot E) \\ &= det(\begin{pmatrix} \lambda_1 - t & & 0 \\ & \ddots & \\ * & & \lambda_s - t \end{pmatrix}) \\ &= (\lambda_1 - t)^{n_1} \cdots (\lambda_s - t)^{n_s}. \end{aligned}$$