LINEARE ALGEBRA II

ao.Univ.Prof. Dr. Michael Drmota

Sommersemester 2006

Inhaltsverzeichnis

Kapitel 8

Polynome und Körpererweiterungen

8.1 Polynome über Ringe und Körpern

8.1.1 Der Polynomring $R[x]$

Sei R ein Ring. Unter einem **Polynom** $f(x)$ über R in der **Unbestimmten** x versteht man üblicherweise eine formale Summe der Form

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n$$

mit $n \geq 0$ und $a_0, a_1, \ldots, a_n \in R$. Die Elemente $a_0, a_1, \ldots, a_n$ heißen **Koeffizienten** von $f(x)$.

Diese Definition ist allerdings *ungünstig*, wenn man mit ihr *exakt* weiterarbeiten möchte. Beispielsweise müßte man, um Gleichheit zu definieren, folgendermaßen vorgehen:

> Zwei Polynome $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ und $g(x) = b_0 + b_1 x + \cdots + b_m x^m$ heißen gleich, i.Z. $f(x) = g(x)$, wenn die entspechenden Koeffizienten gleich sind, d.h. (für $m \leq n$) gilt $a_0 = b_0$, $a_1 = b_1$, …, $a_m = b_m$, und $a_j = 0$ für $m < j \leq n$.

Es ist viel günstiger, Polynome als Spezialfall formaler Potenzreihen

$$\sum_{n \geq 0} a_n x^n$$

zu betrachten, wobei verlangt wird, daß die Menge

$$\{n \in \mathbb{N} \,|\, a_n \neq 0\}$$

endlich ist. In dieser Schreibweise sind zwei Polynome genau dann gleich, wenn alle Koeffizienten übereinstimmen.

Definition 8.1 *Sei R ein Ring. Eine formale Reihe der Form*

$$\sum_{n \geq 0} a_n x^n$$

mit $a_n \in R$ $(n \geq 0)$ *heißt* **formale Potenzreihe** *über* R. *Die Menge aller formalen Potenzreihen über* R *wird durch* $R[[x]]$ *bezeichnet.*

Eine formale Potenzreihe

$$\sum_{n\geq 0} a_n x^n$$

mit der Eigenschaft, daß nur endlich viele a_n *von* 0 *verschieden sind, heißt* **Polynom** *über* R. *Die Menge aller Polynome über* R *wird durch* $R[x]$ *bezeichnet.*

Es lassen sich auch leicht Grad eines Polynoms und Rechenoperationen definieren.

Definition 8.2 *Sei* R *ein Ring und*

$$f(x) = \sum_{n\geq 0} a_n x^n \in R[x]$$

ungleich dem Nullpolynom, so heißt

$$\mathrm{grad}(f(x)) := \max\{n \in \mathbb{N} \,|\, a_n \neq 0\}$$

Grad *von* $f(x)$. *Der Koeffizient* $a_{\mathrm{grad}(f(x))} \neq 0$ *von* $f(x)$ *heißt* **führender Koeffizient** *oder* **Hauptkoeffizient** *von* $f(x)$.

Sind alle Koeffizienten a_j *eines Polynoms* $f(x)$ $a_j = 0$, *so bezeichnet man* $f(x)$ *als* **Nullpolynom**, *das auch durch* 0 *bezeichnet wird.*

Das Nullpolynom und die Polynome vom Grad 0 *heißen* **konstante Polynome**. *Sie haben die Form* $f(x) = a_0$.

Ist R *ein Ring mit* 1, *so heißt ein Polynom* **normiert** *oder* **monisch**, *wenn der führende Koeffizient* 1 *ist.*

Definition 8.3 *Sei* R *ein Ring und* $f(x) = \sum_{n\geq 0} a_n x^n$, $g(x) = \sum_{n\geq 0} b_n x^n \in R[x]$. *Die* **Summe** *von* $f(x)$ *und* $g(x)$ *ist durch*

$$f(x) + g(x) := \sum_{n\geq 0} (a_n + b_n) x^n$$

definiert, und das **Produkt** *von* $f(x)$ *und* $g(x)$ *durch*

$$f(x)g(x) := \sum_{n\geq 0} \left(\sum_{j=0}^{n} a_j b_{n-j} \right) x^n.$$

Man beachte, daß Summe und Produkt von Polynomen wieder Polynome sind. Für die Grade gilt

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

$R[x]$ bildet mit diesen Operationen sogar wieder einen Ring.

Satz 8.4 *Für jeden Ring* $\langle R,+,\cdot\rangle$ *ist* $\langle R[x],+,\cdot\rangle$ *wieder ein Ring, der sogenannte* **Polynomring** *über R. Die konstanten Polynome* $f(x)=a_0$ *(mit* $a_0 \in R$*) bilden einen Unterring von* $R[x]$*, der in natürlicher Weise zu R isomorph ist.*

Dabei werden folgende Eigenschaften vererbt:

1. *R ist Ring mit* $1 \Longrightarrow R[x]$ *ist Ring mit* 1.
2. *R ist kommutativ* $\Longrightarrow R[x]$ *ist kommutativ.*
3. *R ist nullteilerfrei* $\Longrightarrow R[x]$ *ist nullteilerfrei.*

Insbesondere ist der Polynomring eines Integritätsbereichs wieder ein Integritätsbereich. Hier gilt übrigens auch

$$\mathrm{grad}(f(x)g(x)) = \mathrm{grad}(f(x)) + \mathrm{grad}(g(x)).$$

In vielen Fällen reicht es auch aus, Polynome über einem Körper K zu betrachten. Da jeder Körper auch Integritätsring ist, ist der Polynomring $K[x]$ sicherlich auch ein Integritätsbereich. $K[x]$ ist aber sicherlich kein Körper, da das Polynom x kein multiplikatives Inverses besitzt. Nur die Polynome $f(x)=a_0\neq 0$ von Grad 0 sind invertierbar.

Polynomringe über einem Körper tragen automatisch auch eine Vektorraumstruktur. Ist $\lambda\in K$ und $f(x)\in K[x]$, so kann λ als konstantes Polynom interpretiert werden, und es ist natürlich, die Skalarmultiplikation $\lambda f(x)$ durch die Polynommultiplikation $\lambda f(x)$ zu definieren.

Satz 8.5 *Sei K ein Körper. Dann bildet* $K[x]$ *einen abzählbar unendlichdimensionalen Vektorraum über K. Die Menge*

$$\mathbf{B} := \{1, x, x^2, x^3, \ldots\}$$

ist eine Basis.

8.1.2 Was ist x?

Die oben verwendeten Begriffe *Unbestimmte x*, *formale Summe*, *formale Potenzreihe* bedürfen natürlich einer gewissen Erläuterung. Im folgenden soll eine mögliche Interpretation für x etc. im Fall eines kommutativen Rings R mit 1 gegeben werden. Zunächst geben wir eine alternative Definition für Polynome.

Definition 8.6 *Sei R ein Ring. Dann wird durch*

$$R[x] := \{(a_n)_{n\in\mathbb{N}} \,|\, (\forall n\in\mathbb{N} : a_n\in R) \wedge (|\{n\in\mathbb{N}\,|\,a_n\neq 0\}| < \aleph_0)\}$$

die Menge der Polynome über R definiert, d.h. ein Polynom über R ist eine Folge $(a_n)_{n\in\mathbb{N}} = (a_0, a_1, \ldots)$ *von Elementen aus R, wo nur endlich viele Elemente* $\neq 0$ *sind.*

Summe und Produkt von Polynomen ist durch

$$\begin{aligned} (a_n)_{n\in\mathbb{N}} + (b_n)_{n\in\mathbb{N}} &:= (a_n+b_n)_{n\in\mathbb{N}}, \\ (a_n)_{n\in\mathbb{N}} \cdot (b_n)_{n\in\mathbb{N}} &:= \left(\sum_{j=0}^{n} a_j b_{n-j}\right)_{n\in\mathbb{N}} \end{aligned}$$

definiert.

Es ist klar, wie ein Polynom $\sum\limits_{n\in\mathbb{N}} a_n x^n$ im obigen Sinn mit einem Polynom $(a_n)_{n\in\mathbb{N}}$ im jetzigen Sinn zu identifizieren ist. Insbesondere entsprechen die konstanten Polynome $f(x) = a_0$ den Folgen

$$(a_0, 0, 0, 0, \ldots).$$

Es ist daher durchaus sinnvoll, eine Folge der Form $(a, 0, 0, 0, \ldots)$ mit a zu identifizieren.

Damit ist aber auch klar, wie man die Unbestimmte x interpretieren kann. Sie entspricht der Folge $(0, 1, 0, 0, \ldots)$, wenn R ein Ring mit 1 ist.

Von nun an sei vorausgesetzt, daß R ein Ring mit 1 ist. Definiert man nun Polynome durch Folgen und setzt

$$x := (\delta_{n,1})_{n\in\mathbb{N}} = (0, 1, 0, 0, \ldots),$$

so ergibt sich induktiv

$$x^j = (\delta_{n,j})_{n\in\mathbb{N}} = (0, \ldots, 0, 1, 0, 0, \ldots)$$

und daher

$$\begin{aligned}(0, \ldots, 0, a_j, 0, 0, \ldots) &= (a_j, 0, 0, \ldots) \cdot (0, \ldots, 0, 1, 0, 0, \ldots) \\ &= a_j x^j\end{aligned}$$

Mit dieser Festsetzung von x läßt sich jedes Polynom $(a_n)_{n\in\mathbb{N}}$ in der Form

$$(a_n)_{n\in\mathbb{N}} = a_0 + a_1 x + \cdots + a_m x^m$$

(mit einem genügend großem $m \in \mathbb{N}$) darstellen.

8.1.3 Polynome und K-Algebren

Definition 8.7 *Ein Vektorraum* $\mathbf{A}$ *über einem Körper* K*, auf dem eine zusätzliche binäre Operation* $\cdot$ *definiert ist, so daß* $\langle \mathbf{A}, +, \cdot \rangle$ *ein Ring mit* 1 *ist und für alle* $y \in K$ *und* $\mathbf{a}, \mathbf{b} \in \mathbf{A}$

$$y(\mathbf{a} \cdot \mathbf{b}) = (y\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (y\mathbf{b})$$

gilt, heißt K-**Algebra**.

Beispiel 8.8 Der Polynomring $K[x]$ über einem Körper K ist eine K-Algebra.

Beispiel 8.9 Der Vektorraum $L(\mathbf{V}, \mathbf{V})$ der linearen Selbstabbildungen eines Vektoraums $\mathbf{V}$ (über einem Körper K) bildet mit der Hintereinanderausführung $\circ$ eine K-Algebra.

Beispiel 8.10 Die Menge $K^{n\times n}$ aller quadratischen Matrizen über einem Körper K bildet eine K-Algebra.

Bis jetzt wurde die Unbestimmte x rein formal behandelt, es ist aber naheliegend, *in ein Polynom einzusetzen*, d.h. man ersetzt die Unbestimmte x durch ein bestimmtes Element, etwa durch ein Element aus R.

Definition 8.11 *Sei R ein Ring und $c \in R$. Dann bezeichnet man durch*

$$\psi_c : R[x] \to R, \quad \sum_{n\in\mathbb{N}} a_n x^n \mapsto \sum_{n\in\mathbb{N}} a_n c^n$$

Einsetzungsabbildung *an der Stelle c. Das Bild $\psi_c(f(x))$ wird auch durch $f(c)$ abgekürzt.*

$c \in R$ heißt **Nullstelle** *eines Polynoms $f(x) \in R[x]$, wenn $\psi_c(f(x)) = f(c) = 0$.*

Satz 8.12 *Sei R ein kommutativer Ring. Dann ist die Einsetzungsabbildung $\psi_c : R[x] \to R$ für jedes $c \in R$ ein Ringhomomorphismus kommutativer Ringe.*

Die Einsetzungsabbildung ψ_c heißt in diesem Fall auch **Einsetzungshomomorphismus**.

Bei Polynomen über einem Körper K ist es auch möglich, Elemente einer K-Algebra einzusetzen.

Satz 8.13 *Sei K ein Körper, $\mathbf{A}$ eine K-Algebra und $\mathbf{c} \in \mathbf{A}$. Dann ist die Abbildung*

$$\psi_{\mathbf{c}} : K[x] \to \mathbf{A}, \quad \sum_{n\in\mathbb{N}} a_n x^n \mapsto \sum_{n\in\mathbb{N}} a_n \mathbf{c}^n$$

ein K-Algebrenhomomorphismus, d.h. $\psi_{\mathbf{c}}$ ist gleichzeitig Ringhomomorphismus und lineare Abbildung.

Jede Identität, die für Polynome aus $K[x]$ besteht, gilt damit automatisch auch dann, wenn man anstelle von x ein beliebiges Element einer K-Algebra (etwa eine Matrix oder eine lineare Selbstabbildung) einsetzt.

8.2 Divisionsalgorithmus und irreduzible Polynome

8.2.1 Teilbarkeit von Polynomen

Ähnlich wie in den ganzen Zahlen definiert man auf einem Polynomring Teilbarkeit.

Definition 8.14 *Sei K ein Körper und $p(x), q(x) \in K[x]$. $p(x)$ ist ein* **Teiler** *von $q(x)$, wenn es ein Polynom $r(x) \in K[x]$ mit*

$$q(x) = p(x)r(x)$$

gibt. Man sagt auch $p(x)$ **teilt** *$q(x)$ und schreibt dafür*

$$p(x)|q(x).$$

Definition 8.15 *Ein Polynom $p(x) \in K[x]$ vom $\mathrm{grad}(p(x)) > 0$ heißt* **irreduzibel**, *wenn es keine Polynome $a(x), b(x)$ mit*

$$p(x) = a(x)b(x)$$

und $\mathrm{grad}(a(x)) < \mathrm{grad}(p(x))$ und $\mathrm{grad}(b(x)) < \mathrm{grad}(p(x))$ gibt.

Irreduzible Polynome haben also nur *triviale* Teiler $c \in K^\times$ und $cp(x)$ (mit $c \in K^\times$). Beispielsweise sind alle Polynome $p(x) = a_0 + a_1 x$ von Grad 1 (mit $a_1 \neq 0$) irreduzible Polynome.

Lemma 8.16 *Sei K ein Körper. Dann gelten für Polynome $a(x), b(x), c(x) \in K[x]$ die folgenden Eigenschaften.*

1. $a(x)|b(x) \wedge b(x)|c(x) \Longrightarrow a(x)|c(x)$.
2. $a(x)|b(x) \wedge a(x)|c(x) \Longrightarrow a(x)|\left(b(x) + c(x)\right)$.
3. $a(x)|b(x) \wedge b(x)|a(x) \Longleftrightarrow \exists c \in K^\times : \; a(x) = cb(x)$.

Definition 8.17 *Sei K ein Körper. Ein Polynom $d(x) \in K[x]$ heißt* **größter gemeinsamer Teiler** *von zwei Polynomen $p(x), q(x) \in K[x]$, wenn*

1. $d(x)|p(x) \wedge d(x)|q(x)$ *und*
2. $(t(x)|p(x) \wedge t(x)|q(x)) \Longrightarrow t(x)|d(x)$ *gilt.*

Zwei Polynome $p(x), q(x) \in K[x]$ heißen **teilerfremd**, *wenn das Polynom 1 größter gemeinsamer Teiler ist.*

Man beachte, daß es bei Polynomen keinen eindeutigen größten gemeinsamen Teiler geben kann. Es gilt nur folgende Beziehung:

Lemma 8.18 *Sind $d_1(x), d_2(x)$ größte gemeinsame Teiler von zwei Polynomen $p(x), q(x) \in K[x]$, so gilt*

$$d_1(x)|d_2(x) \wedge d_2(x)|d_1(x),$$

bzw. es gibt ein $c \in K^\times$ mit

$$d_2(x) = cd_1(x).$$

Insbeondere sind bei teilerfremden Polynomen alle $c \in K^\times$ größte gemeinsame Teiler.

Definition 8.19 *Ein Polynom $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in K[x]$ vom Grad n heißt* **normiert**, *wenn $a_n = 1$ ist.*

Unter allen größten gemeinsamen Teilern zweier Polynome $p(x), q(x) \in K[x]$ gibt es genau ein normiertes. Dieses wird dann oft als *der* größte gemeinsame Teiler $d(x) = \mathrm{ggT}(p(x), q(x))$ bezeichnet.

8.2.2 Division mit Rest

Eine der grundlegenden Eigenschaften von Polynomringen über Körpern ist die **Division mit Rest**.

Satz 8.20 *Sei K ein Körper und $a(x), b(x) \in K[x]$ mit $b(x) \neq 0$. Dann gibt es Polynome $q(x), r(x) \in K[x]$ mit*

$$a(x) = b(x)q(x) + r(x),$$

wobei entweder $r(x) = 0$ oder $\operatorname{grad}(r(x)) < \operatorname{grad}(b(x))$ ist.

Die erste Anwendung der Division mit Rest ist der **Euklidische Algorithmus** zur Berechnung eines größten gemeinsamen Teilers.

Satz 8.21 *Sei K ein Körper und $a(x), b(x) \in K[x]$ mit $b(x) \neq 0$. Führt man die folgende Divisonskette durch:*

$$\begin{aligned} a(x) &= b(x)q_0(x) + r_0(x), & \operatorname{grad}(r_0(x)) &< \operatorname{grad}(b(x)) \\ b(x) &= r_0(x)q_1(x) + r_1(x) & \operatorname{grad}(r_1(x)) &< \operatorname{grad}(r_0(x)) \\ r_0(x) &= r_1(x)q_2(x) + r_2(x) & \operatorname{grad}(r_2(x)) &< \operatorname{grad}(r_1(x)) \\ &\vdots \\ r_{k-2}(x) &= r_{k-1}(x)q_k(x) + r_k(x) & \operatorname{grad}(r_k(x)) &< \operatorname{grad}(r_{k-1}(x)) \\ r_{k-1}(x) &= r_k(x)q_{k+1}(x) + 0 & , \end{aligned}$$

dann ist der letzte von 0 verschiedene Rest $r_k(x)$ größter gemeinsamer Teiler von $a(x)$ und $b(x)$.

Aus dieser Divisionskette gelingt es auch, den größten gemeinsamen Teiler durch eine Linearkombination darzustellen.

Satz 8.22 *Sei K ein Körper und $d(x) \in K[x]$ ein größter gemeinsamer Teiler von $a(x), b(x) \in K[x]$ Dann gibt es Polynome $e(x), f(x) \in K[x]$ vom Grad $\operatorname{grad}(e(x)) < \operatorname{grad}(b(x))$ und $\operatorname{grad}(f(x)) < \operatorname{grad}(a(x))$ mit*

$$d(x) = e(x)a(x) + f(x)b(x).$$

8.2.3 Eindeutige Zerlegung von Polynomen

Eine erste Folgerung aus Satz 8.22 ist die folgende Eigenschaft für irreduzible Polynome.

Lemma 8.23 *Sei $p(x) \in K[x]$ ein irreduzibles Polynom und $a(x), b(x) \in K[x]$. Dann gilt*

$$p(x)|a(x)b(x) \Longrightarrow p(x)|a(x) \vee p(x)|b(x).$$

Diese Eigenschaft ist grundlegend, um den folgenden Satz über die Eindeutigkeit der Zerlegung in irreduzible Polynome zu beweisen.

Satz 8.24 *Sei K ein Körper. Dann läßt sich jedes Polynom $a(x) \in K[x]$ als Produkt irreduzibler Polynome darstellen. Diese Darstellung ist im folgenden Sinn eindeutig. Sind*

$$a(x) = cp_1(x)p_2(x)\cdots p_n(x) = cq_1(x)q_2(x)\cdots q_m(x)$$

zwei solche Darstellungen mit normierten irreduziblen Polynomen $p_1(x), p_2(x), \ldots, p_n(x)$ bzw. $q_1(x), q_2(x), \ldots, q_m(x)$, so ist $n = m$ und die irreduziblen Polynome $p_1(x), p_2(x), \ldots, p_n(x)$ stimmen bis auf die Reihenfolge mit den irreduziblen Polynomen $q_1(x), q_2(x), \ldots, q_m(x)$ überein.

8.2.4 Nullstellen von Polynomen

Definition 8.25 *Sei K ein Körper und $p(x) \in K[x]$. Ein Körperelement $\alpha \in K$ heißt* **Nullstelle** *von $p(x)$, wenn $p(\alpha) = 0$ ist.*

Satz 8.26 *Sei K ein Körper und $p(x) \in K[x]$. Dann gilt für alle $\alpha \in K$*

$$p(\alpha) = 0 \Longleftrightarrow (x - \alpha) | p(x).$$

Definition 8.27 *Sei K ein Körper, $p(x) \in K[x]$ und $\alpha \in K$ Nullstelle von $p(x)$. Die* **Vielfachheit** *der Nullstelle α ist jene natürliche Zahl k mit*

$$(x - \alpha)^k | p(x) \ \wedge \ (x - \alpha)^{k+1} \nmid p(x),$$

Man kann auch sagen, daß $(x - \alpha)$ in der eindeutigen Zerlegung von $p(x)$ in irreduzible Faktoren genau k-mal vorkommt.

Eine wichtige Folgerung aus der eindeutigen Zerlegung in irreduzible Faktoren ist der folgende Satz.

Satz 8.28 *Sei K ein Körper und $p(x) \in K[x]$ ein Polynom von Grad $\mathrm{grad}(p(x)) = n$. Dann gibt es höchstens n verschiedene Nullstellen von $p(x)$.*

Dieser Satz kann noch verschärft werden. Bezeichnen $\alpha_1, \ldots, \alpha_m \in K$ die verschiedenen Nullstellen von $p(x)$ und $k_1, \ldots, k_m$ die entsprechenden Vielfachheiten, so gilt

$$k_1 + \cdots + k_m \leq n.$$

8.3 Fundamentalsatz der Algebra

8.3.1 Polynome über den komplexen Zahlen

Der **Fundamentalsatz der Algebra** lautet:

Satz 8.29 *Jedes Polynom $f(x) \in \mathbb{C}[x]$ hat eine komplexe Nullstelle $\alpha \in \mathbb{C}$.*

Daraus folgt unmittelbar

Satz 8.30 *Sei $f(x) \in \mathbb{C}[x]$ ein Polynom vom Grad $\mathrm{grad}(f(x)) = n$. Sind $\alpha_1, \ldots, \alpha_m \in \mathbb{C}$ die Nullstellen von $f(x)$ und $k_1, \ldots, k_m$ die Vielfachheiten dieser Nullstellen, so gilt*

$$f(x) = a(x - \alpha_1)^{k_1} \cdots (x - \alpha_m)^{k_m},$$

wobei $a \in \mathbb{C}^\times$ der Hauptkoeffizient von $f(x)$ ist, d.h. $f(x)$ zerfällt in Linearfaktoren.

Dies kann auch anders ausgedrückt werden.

Satz 8.31 *Die einzigen irreduziblen Polynome in $\mathbb{C}[x]$ sind die Polynome vom Grad 1.*

8.3.2 Polynome über den reellen Zahlen

Jedes Polynom $f(x) \in \mathbb{R}[x]$ kann selbstverständlich auch über $\mathbb{C}$ interpretiert werden, und hat dann natürlich, wenn es keine reelle Nullstelle haben sollte, jedenfalls eine Nullstelle $\alpha \in \mathbb{C}$.

Lemma 8.32 *Sei $f(x) \in \mathbb{R}[x]$ und $\alpha \in \mathbb{C}$ eine Nullstelle von $f(x)$. Dann ist auch die konjugiert komplexe Zahl $\overline{\alpha} \in \mathbb{C}$ eine Nullstelle von $f(x)$.*

Ist $\alpha = \sigma + i\tau$ mit $\tau \neq 0$ eine Nullstelle von $f(x) \in \mathbb{R}[x]$, so gewinnt man daher automatisch eine weitere Nullstelle $\overline{\alpha} = \sigma - i\tau$, also ein **konjugiert komplexes Nullstellenpaar** und somit einen Teiler von $f(x)$ der Form

$$(x-\alpha)(x-\overline{\alpha}) = x^2 - 2\sigma x + (\sigma^2 + \tau^2) \in \mathbb{R}[x].$$

Jedes reelle Polynom zerfällt daher in der folgenden Art und Weise.

Satz 8.33 *Sei $f(x) \in \mathbb{R}[x]$ ein Polynom vom Grad n. Sind $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ mit Vielfachheiten $k_1, \ldots, k_m$ die reellen Nullstellen von $f(x)$ und $\sigma_1 \pm i\tau_1, \ldots, \sigma_s \pm i\tau_s$ mit Vielfachheiten $l_1, \ldots, l_s$ die konjugiert komplexen Nullstellenpaare von $f(x)$, so gilt*

$$f(x) = a(x-\alpha_1)^{k_1} \cdots (x-\alpha_m)^{k_m}(x^2 - 2\sigma_1 x + (\sigma_1^2 + \tau_1^2))^{l_1} \cdots (x^2 - 2\sigma_s x + (\sigma_s^2 + \tau_s^2))^{l_s},$$

wobei $a \in \mathbb{R}^\times$ der Hauptkoeffizient von $f(x)$ ist, d.h. $f(x)$ zerfällt in Linearfaktoren und quadratische Faktoren.

Dies kann auch anders ausgedrückt werden.

Satz 8.34 *Die einzigen irreduziblen Polynome in $\mathbb{R}[x]$ sind die Polynome von Grad 1 und die quadratischen Polynome*

$$x^2 + px + q$$

mit negativer Diskriminante

$$p^2 - 4q < 0.$$

8.4 Algebraische Körpererweiterungen

8.4.1 Polynomielle Kongruenzen

Der Übergang von $\mathbb{R}$ nach $\mathbb{C}$ besteht im wesentlichen durch Hinzufügen einer Nullstelle des Polynoms $x^2 + 1$, die (üblicherweise) durch i bezeichnet wird. Es wird dabei nicht die Frage beantwortet, *was i ist*, sondern nur, welche *Eigenschaften i* hat. Die wesentliche Eigenschaft von i ist $i^2 = -1$. Damit kann man mit i rechnen und Ausdrücke der Form $a + bi$ mit $a, b \in \mathbb{R}$ addieren, subtrahieren, multiplizieren und, was besonders bemerkenswert ist, dividieren, falls $a \neq 0$ oder $b \neq 0$ ist.

Eine ähnliche Vorgangsweise gelingt auch bei den rationalen Zahlen $\mathbb{Q}$, wenn man die irrationale Zahl $\sqrt{2}$ dazufügt. Wie zuvor kann man Ausdrücke der Form $a + b\sqrt{2}$ mit $a, b \in \mathbb{Q}$ addieren, subtrahieren, multiplizieren und dividieren, falls $a \neq 0$ oder $b \neq 0$ ist.

Beiden Beispielen gemeinsam ist, daß zu einem Körper K ($\mathbb{R}$ bzw. $\mathbb{Q}$) ein Element α (i bzw. $\sqrt{2}$) hinzugefügt wird, so daß die Menge aller Ausdrücke $a + b\alpha$ mit $a, b \in K$ wieder einen Körper $L = K(\alpha)$ bilden, einen Erweiterungskörper von K. Das Element α ist über dem Körper L Nullstelle eines Polynoms $f(x) \in K[x]$ ($x^2 + 1 \in \mathbb{R}[x]$ bzw. $x^2 - 2 \in \mathbb{Q}[x]$), wobei $f(x)$ über K irreduzibel ist. (Man spricht hier auch von eine algebraischen Körpererweiterung, weil das Element, das zum Körper K dazugefügt wird, eine algebraische Gleichung erfüllt.)

Im folgenden sollen zwei Zugänge zu allgemeinen algebraischen Körpererweiterungen vorgestellt werden. Der erste basiert auf polynomiellen Kongruenzen.

Definition 8.35 *Sei K ein Körper. Zwei Polynome $a(x), b(x) \in K[x]$ heißen* **kongruent** *modulo einem Polynom $f(x) \in K[x]$, i.Z. $a(x) \equiv b(x) \bmod f(x)$, wenn*

$$f(x)|a(x) - b(x).$$

Offensichtlich bilden die Polynome

$$f(x)K[x] = \{a(x) \in K[x] \,|\, a(x) \equiv 0 \bmod f(x)\}$$

bezüglich der Addition eine Untergruppe.

Ganz analog zu den ganzen Zahlen gelten die folgenden Eigenschaften:

Lemma 8.36 *Sei K ein Körper und $a_1(x), a_2(x), b_1(x), b_2(x), f(x) \in K[x]$ und $a_1(x) \equiv b_1(x) \bmod f(x)$ und $a_2(x) \equiv b_2(x) \bmod f(x)$, dann gilt auch*

$$a_1(x) + a_2(x) \equiv b_1(x) + b_2(x) \bmod f(x)$$

und

$$a_1(x) \cdot a_2(x) \equiv b_1(x) \cdot b_2(x) \bmod f(x).$$

Man kann daher auf den Nebenklassen von $f(x)K[x]$ wieder in natürlicher Weise Addition und Multiplikation definieren.

Definition 8.37 *Sei K ein Körper und $f(x) \in K[x]$. Die Menge der Nebenklassen von $f(x)K[x]$ wird durch*

$$K[x]/f(x)K[x] = \{a(x) + f(x)K[x] \,|\, a(x) \in K[x]\}$$

bezeichnet. Die Summe zweier Nebenklassen wird durch

$$(a(x) + f(x)K[x]) + (b(x) + f(x)K[x]) := (a(x) + b(x)) + f(x)K[x]$$

und das Produkt durch

$$(a(x) + f(x)K[x]) \cdot (b(x) + f(x)K[x]) := (a(x) \cdot b(x)) + f(x)K[x]$$

definiert.

Aus den Nebenklassen $a(x) + f(x)K[x]$ kann ein kanonischer Repräsentant gefunden werden.

Lemma 8.38 *Sei K ein Körper und $f(x) \in K[x]$. Dann gibt es zu jeder Nebenklasse $a(x) + f(x)K[x] \in K[x]/f(x)K[x]$ ein Polynom $a_0(x)$ mit $\operatorname{grad}(a_0(x)) < \operatorname{grad}(f(x))$.*

Weiters sind zwei Polynome $a_0(x), b_0(x) \in K[x]$ mit $\operatorname{grad}(a_0(x)) < \operatorname{grad}(f(x))$ und $\operatorname{grad}(b_0(x)) < \operatorname{grad}(f(x))$ genau dann gleich, wenn sie kongruent modulo $f(x)$ sind.

$K[x]/f(x)K[x]$ kann daher auch durch

$$K[x]/f(x)K[x] = \{a(x) + f(x)K[x] \mid (a(x) \in K[x]) \wedge (\operatorname{grad}(a(x)) < \operatorname{grad}(f(x)))\}$$

beschrieben werden.

Satz 8.39 *Sei K ein Körper und $f(x) \in K[x]$. Die algebraische Struktur $\langle K[x]/f(x)K[x], +, \cdot\rangle$ ist ein kommutativer Ring mit 1. Er ist genau dann ein Körper, wenn $f(x)$ über K irreduzibel ist.*

Die Ringeigenschaften sind sofort nachzurechnen. Die Berechnung der inversen Nebenklasse erfolgt mit Hilfe des Euklidischen Algorithmus. Sei $f(x) \in K[x]$ irreduzibel über K und sei $a(x)+f(x)K[x] \neq f(x)K[x]$ eine Nebenklasse modulo $f(x)$. Dann sind $a(x)$ und $f(x)$ teilerfremd. Folglich können mit Hilfe des Euklidischen Algorithmus Polynome $b(x), q(x) \in K[x]$ mit

$$a(x)b(x) + q(x)f(x) = 1$$

bestimmt werden und es gilt

$$(a(x) + f(x)K[x]) \cdot (b(x) + f(x)K[x]) = (a(x) \cdot b(x)) + f(x)K[x] = 1 + f(x)K[x],$$

d.h. $b(x) + f(x)K[x]$ ist bezüglich der Multiplikation das zu $a(x) + f(x)K[x]$ inverse Element.

Im Fall eines irreduziblen Polynoms $f(x)$ kann $K[x]$ als **Erweiterung** von K angesehen werden, da die Nebenklassen der konstanten Polynome

$$a + f(x)K[x]$$

eine *Kopie* von K darstellen.

Weiteres hat die Nebenklasse

$$\alpha := x + f(x)K[x]$$

eine interessante Deutung. Es gilt nämlich

$$f(\alpha) = f(x + f(x)K[x]) = 0.$$

Im Erweiterungkörper $L = K[x]/f(x)K[x]$ hat daher das Polynom $f(x)$ eine Nullstelle. Weiters kann L durch

$$L = \{a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1} \mid a_0, a_1, \ldots, a_{n-1} \in K \wedge f(\alpha) = 0\}$$

beschrieben werden. Für $K = \mathbb{R}$ und $f(x) = x^2 + 1$ erhält man dadurch gerade $\mathbb{C} = \mathbb{R}(i)$ und für $K = \mathbb{Q}$ und $f(x) = x^2 - 2$ den Körper $\mathbb{Q}(\sqrt{2})$.

Es ist also zulässig, bei einem über K irreduziblen Polynom $f(x) \in K[x]$, das ja keine Nullstellen in K hat, die Existenz einer **algebraischen Nullstelle** α von $f(x)$ als gegeben zu nehmen und mit ihr in üblicher Weise zu rechnen. Der Erweiterungskörper L (der eigentlich in der obigen Weise konstruiert werden muß) wird durch $L = K(\alpha)$ bezeichnet.

8.4.2 Direkte Produkte

Im vorigen Abschnitt wurde gezeigt, daß alle Elemente aus $L = K[x]/f(x)K[x]$ durch

$$\begin{aligned} K[x]/f(x)K[x] &= \{a(x) + f(x)K[x] \mid a(x) \in K[x] \wedge \operatorname{grad}(a(x)) < \operatorname{grad}(f(x))\} \\ &= \{a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + f(x)K[x] \mid a_0, a_1, \ldots, a_{n-1} \in K\} \end{aligned}$$

bzw. durch

$$K[x]/f(x)K[x] = K(\alpha) = \{a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1} \mid a_0, a_1, \ldots, a_{n-1} \in K \wedge f(\alpha) = 0\}$$

beschrieben werden können. Jedenfalls kann man $K[x]/f(x)K[x]$ mit K^n identifizieren, d.h. der Restklasse

$$a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + f(x)K[x] \in K[x]/f(x)K[x]$$

ordnet man das n-Tupel

$$(a_0, a_1, \ldots, a_n) \in K^n$$

zu. Offensichtlich können alle Rechnungen auch in K^n durchgeführt werden. Beispielsweise entspricht der Addition zweier Restklassen die Addition der entsprechenden n-Tupel.

Definiert man beispielsweise auf den Paaren $(a, b) \in \mathbb{R}^2$ Addition und Multiplikation durch

$$(a, b) + (c, d) := (a + c, b + d) \quad \text{und} \quad (a, b) \cdot (c, d) := (ac - bd, ad + bc),$$

so entspricht dies der Konstruktion der komplexen Zahlen $\mathbb{C}$, wobei $a + bi$ bzw. die Restklasse $(a + bx) + (x^2 + 1)\mathbb{R}[x]$ durch das Paar (a, b) repräsentiert wird. Man könnte $\mathbb{C}$ auch in dieser Weise definieren und a-posteriori erkennen, daß die Elemente $(a, 0)$ mit den reellen Zahlen $a \in \mathbb{R}$ identifiert werden können, daß also eine Erweiterung des reellen Körpers $\mathbb{R}$ vorliegt, und das Element $(0, 1)$ als i bezeichnen und nachrechnen, daß

$$i^2 = (0, 1) \cdot (0, 1) = (-1, 0) = -1$$

gilt.

In ähnlicher Weise kann man auf den Paaren $(a, b) \in \mathbb{Q}^2$ rationaler Zahlen Addition und Multiplikation durch

$$(a, b) + (c, d) := (a + c, b + d) \quad \text{und} \quad (a, b) \cdot (c, d) := (ac + 2bd, ad + bc),$$

definieren. Wieder liegt eine Erweiterung vor. Die Elemente $(a, 0)$ können mit $a \in \mathbb{Q}$ identifiziert werden und das Element $(0, 1)$ übernimmt die Rolle von $\sqrt{2}$:

$$(0, 1) \cdot (0, 1) = (2, 0) = 2.$$

Die allgemeine Vorgangsweise ist daher die folgende. Sei K ein Körper und $f(x) \in K[x]$ ein irreduzibles Polynom vom Grad n. Man kann nun auf $L = K^n$ Addition und Multiplikation definieren, so daß die Elemente $(a, 0, 0, \ldots, 0) \in K^n$ mit $a \in K$ identifiziert werden können, d.h. K^n kann als Erweiterungskörper von K interpretiert werden, und daß das Element

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

Nullstelle von $f(x)$ ist: $f(\alpha) = 0$. (Dabei muß das Polynom $f(x) \in K[x]$ als Polynom in $L[x]$ interpretiert werden, indem K durch die gerade erwähnte Identifikation als Unterkörper von $L = K^n$ aufgefaßt wird.)

Die Addition auf $L = K^n$ wird einfach durch

$$(a_0, a_1, \ldots, a_{n-1}) + (b_0, b_1, \ldots, b_{n-1}) := (a_0 + b_0, a_1 + b_1, \ldots, a_{n-1} + b_{n-1})$$

definiert. Bevor man die Multiplikation definieren kann, muß man mit Hilfe des Divisionsalgorithmus zu Polynomen

$$a(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}, \quad b(x) = b_0 + b_1 x + \cdots + b_{n-1}x^{n-1}$$

ein Polynom

$$c(x) = c_0 + c_1 x + \cdots c_{n-1}x^{n-1}$$

mit

$$a(x)b(x) = q(x)f(x) + c(x)$$

bestimmen. Offensichtlich hängen die Koeffizienten c_j von $c(x)$ nur von den (festen) Koeffizienten von $f(x)$ und von den Koeffizienten a_k von $a(x)$ und b_l von $b(x)$ ab. Die Koeffizienten c_j können daher als Funktionen von a_k, b_l der folgenden Gestalt

$$\begin{aligned} c_j &= c_j(a_0, a_1, \ldots, a_{n-1}, b_0, b_1, \ldots, b_{n-1}) \\ &= \sum_{k,l=0}^{n-1} c_{j,kl} a_k b_l \end{aligned}$$

interpretiert werden. Führt man den Divisonsalgorithmus einmal mit allgemeinen Koeffzienten a_k, b_l durch, so können die Koeffizienten $c_{j,kl}$ auch explizit berechnet werden. (Es ist auch möglich, die Koeffizienten $c_{j,kl}$ durch Lösung eines entsprechenden linearen Gleichungssystems zu gewinnen.)

Das Produkt auf $L = K^n$ wird schließlich durch

$$(a_0, a_1, \ldots, a_{n-1}) \cdot (b_0, b_1, \ldots, b_{n-1}) := (c_0, c_1, \ldots, c_{n-1})$$

definiert.

8.5 Erweiterungen von Vektorräumen

8.5.1 Algebraische Erweiterungen von Vektorräumen

Bei der Bestimmung von Eigenwerten einer linearen Abbildung müssen nur die Nullstellen des charakteristischen Polynoms ermittelt werden. Die Kenntnis von Eigenwerten (und dazugehörigen Eigenräumen) dient zum besseren Verständnis der Struktur der linearen Abbildung. Hat das charakteristische Polynom keine Nullstellen, so kann man zunächst zu einem algebraischen Erweiterungsköper des Grundköpers übergehen, so daß das charakteristische Polynom im Erweiterungskörper Nullstellen hat. Um diese Nullstellen als Eigenwerte einer geeigneten linearen Abbildung zwischen geeignenten Vektorräumen interpretieren zu können, muß man die gegebenen Vektorräume und die lineare Abbildung erweitern.

Ist beispielsweise $\mathbf{V}$ ein Vektorraum über $\mathbb{R}$, so kann es sinnvoll sein, $\mathbf{V}$ zu einem Vektorraum $\mathbf{V}_\mathbb{C}$ über $\mathbb{C}$ zu erweitern. Um die allgemeine Vorgangsweise zu motivieren, gehen wir einmal zunächst formal vor. Wir nehmen an, die Vektoren $\mathbf{c} \in \mathbf{V}_\mathbb{C}$ werden (analog zur Konstruktion der komplexen Zahlen) in Real- und Imaginärteil

$$\mathbf{c} = \mathbf{a} + i\mathbf{b}$$

mit $\mathbf{a}, \mathbf{b} \in \mathbf{V}$ zerlegt. Offensichtlich bildet die Menge dieser Vektoren eine Erweiterung von $\mathbf{V}$. Die Addition dieser Vektoren wird einfach durch

$$(\mathbf{a}_1 + i\mathbf{b}_1) + (\mathbf{a}_2 + i\mathbf{b}_2) := (\mathbf{a}_1 + \mathbf{a}_2) + i(\mathbf{b}_1 + \mathbf{b}_2)$$

definiert. Die Skalarmultiplikation eines Skalars $z = x + iy$ mit einen Vektor $\mathbf{c} = \mathbf{a} + i\mathbf{b}$ wird in natürlicher Weise durch

$$z\mathbf{c} = (x + iy)(\mathbf{a} + i\mathbf{b}) := (x\mathbf{a} - y\mathbf{b}) + i(x\mathbf{b} + y\mathbf{a})$$

festgelegt. (Es wird einfach formal ausmultipliziert und die Rechenregel $i^2 = -1$ benützt.)

Ein exakter Weg, diese Vorgangsweise nachzuzeichnen, ist, eine Menge $\mathbf{V}_\mathbb{C} = \mathbf{V} \times \mathbf{V}$ aller Paare von Vektoren aus $\mathbf{V}$ zu betrachten, die Addition durch

$$(\mathbf{a}_1, \mathbf{b}_1) + (\mathbf{a}_2, \mathbf{b}_2) := (\mathbf{a}_1 + \mathbf{a}_2, \mathbf{b}_1 + \mathbf{b}_2)$$

und die Skalarmultiplikation durch

$$(x + iy)(\mathbf{a}, \mathbf{b}) := (x\mathbf{a} - y\mathbf{b}, x\mathbf{b} + y\mathbf{a})$$

zu definieren. Es ist leicht nachzurechnen, daß $\mathbf{V}_\mathbb{C}$ mit diesen Operation ein Vektorraum über $\mathbb{C}$ ist. Identifiziert man die Vektoren $(\mathbf{a}, \mathbf{0}) \in \mathbf{V}_\mathbb{C}$ mit $\mathbf{a} \in \mathbf{V}$, so erkennt man, daß $\mathbf{V}_\mathbb{C}$ eine Vektorraumerweiterung von $\mathbf{V}$ ist. Wegen

$$(\mathbf{a}, \mathbf{b}) = (\mathbf{a}, \mathbf{0}) + i(\mathbf{b}, \mathbf{0}) = \mathbf{a} + i\mathbf{b}$$

ist die obige Vorgangsweise auch gerechtfertigt.

In ähnlicher Weise kann man auch eine Vektorraum $\mathbf{V}$ über $\mathbb{Q}$ zu einem Vektorraum $\mathbf{V}_{\mathbb{Q}(\sqrt{2})}$ über $\mathbb{Q}(\sqrt{2})$ erweitern. Auf den Paaren $\mathbf{V}_{\mathbb{Q}(\sqrt{2})} = \mathbf{V} \times \mathbf{V}$ definiert man die Addition durch

$$(\mathbf{a}_1, \mathbf{b}_2) + (\mathbf{a}_2, \mathbf{b}_2) := (\mathbf{a}_1 + \mathbf{a}_2, \mathbf{b}_1 + \mathbf{b}_2)$$

und die Skalarmultiplikation durch

$$(x + \sqrt{2}y)(\mathbf{a}, \mathbf{b}) := (x\mathbf{a} + 2\mathbf{b}, x\mathbf{b} + y\mathbf{a}).$$

Wieder erkennt man durch Nachrechnen, daß $\mathbf{V}_{\mathbb{Q}(\sqrt{2})}$ einen Vektorraum über $\mathbb{Q}(\sqrt{2}) = \{x + \sqrt{2}y \,|\, x, y \in \mathbb{Q}\}$ bildet. Identifiziert man wieder die Vektoren $(\mathbf{a}, \mathbf{0}) \in \mathbf{V}_{\mathbb{Q}(\sqrt{2})}$ mit $\mathbf{a} \in \mathbf{V}$, so kann $\mathbf{V}_{\mathbb{Q}(\sqrt{2})}$ wieder als Vektorraumerweiterung von $\mathbf{V}$ interpretiert werden. Weiters führt diese Identifikation zu der einprägsameren Darstellung

$$(\mathbf{a}, \mathbf{b}) = (\mathbf{a}, \mathbf{0}) + \sqrt{2}(\mathbf{b}, \mathbf{0}) = \mathbf{a} + \sqrt{2}\mathbf{b}.$$

Die eben geschilderte Vorgangsweise kann auch allgemein durchgeführt werden.

Definition 8.40 *Sei* $\mathbf{V}$ *ein Vektorraum über dem Körper* K *und* $f(x) \in K[x]$ *ein über* K *irreduzibles Polynom vom Grad* $\text{grad}(f(x)) = n$. *Weiters bezeichne*

$$K(\alpha) = \{x_0 + x_1\alpha + \cdots + x_{n-1}\alpha^{n-1} \,|\, x_0, x_1, \ldots, x_{n-1} \in K\}$$

eine algebraische Körpererweiterung von K *um eine Nullstelle* α *von* $f(x)$.

Dann wird auf der Menge $\mathbf{V}_{K(\alpha)} := \mathbf{V}^n$ *durch*

$$(\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_{n-1}) + (\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_{n-1}) := (\mathbf{a}_0 + \mathbf{b}_0, \mathbf{a}_1 + \mathbf{b}_1, \ldots, \mathbf{a}_{n-1} + \mathbf{b}_{n-1})$$

eine Addition und durch

$$\begin{aligned}&(x_0 + x_1\alpha + \cdots + x_{n-1}\alpha^{n-1}) \cdot (\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_{n-1}) \\ &:= \left(\sum_{k,l=0}^{n-1} c_{0,kl} x_k \mathbf{a}_l, \sum_{k,l=0}^{n-1} c_{1,kl} x_k \mathbf{a}_l, \ldots, \sum_{k,l=0}^{n-1} c_{n-1,kl} x_k \mathbf{a}_l \right)\end{aligned}$$

eine Skalarmultiplikation definiert, wobei die Koeffizienten $c_{j,kl}$ *durch die Multiplikation*

$$\begin{aligned}&(x_0 + x_1\alpha + \cdots + x_{n-1}\alpha^{n-1}) \cdot (y_0 + y_1\alpha + \cdots + y_{n-1}\alpha^{n-1}) \\ &= \left(\sum_{k,l=0}^{n-1} c_{0,kl} x_k y_l, \sum_{k,l=0}^{n-1} c_{1,kl} x_k y_l, \ldots, \sum_{k,l=0}^{n-1} c_{n-1,kl} x_k y_l \right),\end{aligned}$$

in $K(\alpha)$ *gegeben sind.*

Satz 8.41 *Sei* $\mathbf{V}$ *ein Vektorraum über dem Körper* K, $f(x) \in K[x]$ *ein über* K *irreduzibles Polynom und* $K(\alpha)$ *ein Erweiterungskörper von* K *um eine Nullstelle* α *von* $f(x)$. *Dann bildet* $\mathbf{V}_{K(\alpha)}$ *mit der soeben definierten Addition und Sklarmultiplikation einen Vektorraum über* $K(\alpha)$, *der* $\mathbf{V}$ *mittels der Identifikation von* $(\mathbf{a}, \mathbf{0}, \ldots, \mathbf{0}) \in \mathbf{V}_{K(\alpha)}$ *mit* $\mathbf{a} \in \mathbf{V}$ *enthält.*

Jede Basis $\mathbf{B} = \{\mathbf{b}_i \,|\, i \in I\}$ *von* $\mathbf{V}$ *ist (mittels dieser Identifikation) auch eine Basis von* $\mathbf{V}_{K(\alpha)}$. *Insbesondere sind die Dimensionen von* $\mathbf{V}$ *(über* K*) und* $\mathbf{V}_{K(\alpha)}$ *(über* $K(\alpha)$*) gleich.*

Mit Hilfe der Identifikation von $(\mathbf{a}, \mathbf{0}, \ldots, \mathbf{0}) \in \mathbf{V}_{K(\alpha)}$ mit $\mathbf{a} \in \mathbf{V}$ können die Vektoren aus $\mathbf{V}_{K(\alpha)}$ auch durch

$$\begin{aligned}(\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_{n-1}) &= (\mathbf{a}_0, \mathbf{0}, \ldots, \mathbf{0})) + \alpha(\mathbf{a}_1, \mathbf{0}, \ldots, \mathbf{0})) + \cdots + \alpha^{n-1}(\mathbf{a}_{n-1}, \mathbf{0}, \ldots, \mathbf{0})) \\ &= \mathbf{a}_0 + \alpha\mathbf{a}_1 + \cdots + \alpha^{n-1}\mathbf{a}_{n-1}\end{aligned}$$

dargestellt werden.

8.5.2 Erweiterung von linearen Abbildungen

Definition 8.42 *Seien* $\mathbf{V}, \mathbf{W}$ *Vektorräume über dem Körper* K, $g(x) \in K[x]$ *ein über* K *irreduzibles Polynom vom Grad* $\text{grad}(g(x)) = n$ *und* $K(\alpha)$ *ein Erweiterungskörper von* K *um eine Nullstelle* α *von* $g(x)$. *Ist* $f : \mathbf{V} \to \mathbf{W}$ *eine Abbildung, so definiert man die* **Fortsetzung**

$$f_{K(\alpha)} : \mathbf{V}_{K(\alpha)} \to \mathbf{W}_{K(\alpha)}$$

durch

$$f_{K(\alpha)}((\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_{n-1})) := (f(\mathbf{a}_0), f(\mathbf{a}_1), \ldots, f(\mathbf{a}_{n-1})).$$

Satz 8.43 *Seien* $\mathbf{V}, \mathbf{W}$ *Vektorräume über dem Körper* K $g(x) \in K[x]$ *ein über* K *irreduzibles Polynom vom Grad* $\operatorname{grad}(g(x)) = n$ *und* $K(\alpha)$ *ein Erweiterungskörper von* K *um eine Nullstelle* α *von* $g(x)$. *Ist* $f \in L(\mathbf{V}, \mathbf{W})$ *eine lineare Abbildung, so ist auch die Fortsetzung* $f_{K(\alpha)} : \mathbf{V}_{K(\alpha)} \to \mathbf{V}_{K(\alpha)}$ *linear.*

Es gelten folgende Eigenschaften:

1. $f_{K(\alpha)}|_{\mathbf{V}} = f$.
2. $(f(\mathbf{V}))_{K(\alpha)} = f(\mathbf{V}_{K(\alpha)})$.
3. $(\mathrm{kern}(f))_{K(\alpha)} = \mathrm{kern}(f_{K(\alpha)})$.
4. *Sind* $\mathbf{V}$ *und* $\mathbf{W}$ *endlichdimensional und* $\mathbf{B}$ *bzw.* $\mathbf{C}$ *Basen von* $\mathbf{V}$ *bzw.* $\mathbf{W}$*, so gilt*
$$\Phi_{\mathbf{BC}}(f_{K(\alpha)}) = \Phi_{\mathbf{BC}}(f).$$

Insbesondere stimmen Rang und Defekt von f *und* $f_{K(\alpha)}$ *überein.*

Kapitel 9

Diagonalisierbarkeit und Jordansche Normalform

9.1 Eigenwerte und Eigenvektoren

9.1.1 Eigenwerte linearer Abbildungen

Definition 9.1 *Sei* $\mathbf{V}$ *ein Vektorraum über einem Körper* K *und* $f \in L(\mathbf{V}, \mathbf{V})$. *Ein Skalar* $t \in K$ *heißt* **Eigenwert** *von* f, *wenn es einen Vektor* $\mathbf{a} \in \mathbf{V} \setminus \{\mathbf{0}\}$ *mit*

$$f(\mathbf{a}) = t\mathbf{a}$$

gibt.

Die Menge aller Eigenwerte $t \in K$ *einer linearen Abbildung* $f \in L(\mathbf{V}, \mathbf{V})$ *heißt* **Spektrum** *von* f *und wird durch*

$$\mathrm{spec}(f)$$

bezeichnet.

Definition 9.2 *Ein Vektor* $\mathbf{a} \in \mathbf{V} \setminus \{\mathbf{0}\}$, *zu dem es ein* $t \in \mathrm{spec}(f)$ *mit* $f(\mathbf{a}) = t\mathbf{a}$ *gibt, heißt* **Eigenvektor** *von* f *(zum Eigenwert* t*).*

Die Menge aller Eigenvektoren zu einem Eigenwert $t \in \mathrm{spec}(f)$ *(ergänzt um* $\mathbf{0}$*)*

$$\mathbf{V}_t = \mathbf{V}_t(f) := \{\mathbf{a} \in \mathbf{V} \mid f(\mathbf{a}) = t\mathbf{a}\}$$

heißt der zu $t \in \mathrm{spec}(f)$ *gehörige* **Eigenraum** *von* f.

Beispiel 9.3 Sei $\mathbf{V}$ der Vektorraum aller unendlich oft differenzierbaren Funktionen $g : \mathbb{R} \to \mathbb{R}$ und $D : \mathbf{V} \to \mathbf{V}$, $g \to g'$ der Ableitungsoperator. Dann ist für jedes $t \in \mathbb{R}$ die Funktion $g_t(x) = e^{tx}$ **Eigenfunktion** von D zum Eigenwert t, d.h. $\mathrm{spec}(D) = \mathbb{R}$.

Definition 9.4 *Sei* $f \in L(\mathbf{V}, \mathbf{V})$. *Ein Unterraum* $\mathbf{U} \le \mathbf{V}$ *von* $\mathbf{V}$ *heißt* f–**invariant**, *wenn*

$$f(\mathbf{U}) \subseteq \mathbf{U}$$

gilt, d.h. $f|_{\mathbf{U}} \in L(\mathbf{U}, \mathbf{U})$.

Lemma 9.5 *Sei $f \in L(\mathbf{V}, \mathbf{V})$. Dann ist für jedes $t \in \operatorname{spec}(f)$ der Eigenraum $\mathbf{V}_t$ ein f–invarianter Unterraum von $\mathbf{V}$.*

Weiters gilt.

Satz 9.6 *Sei $f \in L(\mathbf{V}, \mathbf{V})$ und seien $t_1, t_2 \in \operatorname{spec}(f)$ zwei verschiedene Eigenwerte von f. Dann gilt*

$$\mathbf{V}_{t_1} \cap \mathbf{V}_{t_2} = \{\mathbf{0}\}.$$

Insbesondere sind zwei Eigenvektoren von f zu verschiedenen Eigenwerten linear unabhängig. Dies gilt sogar für ein System von Eigenvektoren.

Satz 9.7 *Sei $f \in L(\mathbf{V}, \mathbf{V})$. Dann ist jedes System $\{\mathbf{a}_i \,|\, i \in I\}$ von Eigenvektoren von f zu paarweise verschiedenen Eigenwerten linear unabhängig.*

Daraus ergibt sich unmittelbar die folgende Eigenschaft von Eigenräumen.

Satz 9.8 *Sei $f \in L(\mathbf{V}, \mathbf{V})$. Dann ist die Summe $\sum\limits_{t \in \operatorname{spec}(f)} \mathbf{V}_t$ direkt, d.h.*

$$\bigoplus_{t \in \operatorname{spec}(f)} \mathbf{V}_t$$

ist ein Unterraum von $\mathbf{V}$.

Korollar 9.9 *Sei $\mathbf{V}$ ein n-dimensionaler Vektorraum. Dann gibt es zu jeder Abbildung $f \in L(\mathbf{V}, \mathbf{V})$ höchstens n verschiedene Eigenwerte.*

Ein wichtiges Problem wird es sein, zu entscheiden ob

$$\bigoplus_{t \in \operatorname{spec}(f)} \mathbf{V}_t = \mathbf{V}$$

gilt. Dies ist genau jener Fall, wenn $\mathbf{V}$ eine Basis besitzt, die nur aus Eigenvektoren von f besteht.

Satz 9.10 *Sei $f \in L(\mathbf{V}, \mathbf{V})$. Dann gibt es genau dann eine Basis von $\mathbf{V}$, die nur aus Eigenvektoren von f besteht, wenn*

$$\bigoplus_{t \in \operatorname{spec}(f)} \mathbf{V}_t = \mathbf{V}$$

gilt.

Eine Abbildung $f \in L(\mathbf{V}, \mathbf{V})$ mit der Eigenschaft, daß es eine Basis von Eigenvektoren gibt, heißt *einfach strukturiert*.

9.1.2 Eigenwerte von Matrizen

Definition 9.11 *Sei K ein Körper und $A \in K^{n\times n}$ eine quadratische Matrix. Ein Skalar $t \in K$ heißt* **Eigenwert** *von A, wenn es einen Vektor $\mathbf{a} \in K^{n\times 1} \setminus \{\mathbf{0}\}$ mit*

$$A\mathbf{a} = t\mathbf{a}$$

gibt.

Die Menge aller Eigenwerte $t \in K$ von A heißt **Spektrum** *von A und wird durch*

$$\operatorname{spec}(A)$$

bezeichnet.

Offensichtlich gilt $\operatorname{spec}(A) = \operatorname{spec}(f_A)$.

Definition 9.12 *Sei K ein Körper und $A \in K^{n\times n}$ eine quadratische Matrix. Ein Vektor $\mathbf{a} \in K^{n\times 1} \setminus \{\mathbf{0}\}$, zu dem es ein $t \in \operatorname{spec}(A)$ mit $A\mathbf{a} = t\mathbf{a}$ gibt, heißt* **Eigenvektor** *von A (zum Eigenwert t).*

Die Menge aller Eigenvektoren zu einem Eigenwert $t \in \operatorname{spec}(A)$ (ergänzt um $\mathbf{0}$)

$$\mathbf{V}_t = \mathbf{V}_t(A) := \{\mathbf{a} \in \mathbf{V} \mid A\mathbf{a} = t\mathbf{a}\}$$

heißt der zu $t \in \operatorname{spec}(A)$ gehörige **Eigenraum** *von A.*

Offensichtlich besteht ein Zusammenhang zwischen Eigenwerten und Eigenräumen von linearen Abbildungen auf endlichdimensionalen Vektorräumen und Matrizen.

Satz 9.13 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum über einem Körper K und $\mathbf{B}$ eine Basis von $\mathbf{V}$. Dann gilt für $t \in \operatorname{spec}(f)$ und $\mathbf{a} \in \mathbf{V}_t$*

$$\Phi_{\mathbf{BB}}(f) \cdot \Phi_{\mathbf{B}}(\mathbf{a}) = t\,\Phi_{\mathbf{B}}(\mathbf{a}),$$

d.h.

$$\operatorname{spec}(f) = \operatorname{spec}(\Phi_{\mathbf{BB}}(f)) \qquad \textit{und} \qquad \Phi_{\mathbf{B}}(\mathbf{V}_t) = (K^{n\times 1})_t.$$

9.2 Charakteristisches Polynom

9.2.1 Matrizen über Polynomringen

Für viele Anwendungsfälle (wie etwa in der Eigenwerttheorie) ist es günstig, Matrizen zu verwenden, deren Eintragungen nicht nur Körperelemente sind, sondern Polynome. (Für unsere Zwecke wird es ausreichen, Polynome über einem Körper einzusetzen.)

Definition 9.14 *Sei K ein Körper. Die Menge aller $m \times n$-Matrizen mit Eintragungen aus $K[x]$ wird durch $K[x]^{m\times n}$ bezeichnet.*

Sind $A(x) = (a_{ij}(x))_{1\le i\le m, 1\le j\le n}$, $B(x) = (b_{ij}(x))_{1\le i\le m, 1\le j\le n}$ zwei Matrizen aus $K[x]^{m\times n}$ so bezeichnet

$$A(x) + B(x) := (a_{ij}(x) + b_{ij}(x))_{1\le i\le m, 1\le j\le n}$$

die **Summe** *von $A(x)$ und $B(x)$ und*

$$tA(x) := (ta_{ij}(x))_{1\le i\le m, 1\le j\le n}$$

die **Multiplikation** *mit einem* **Skalar** *$t \in K$.*

Lemma 9.15 *Für jeden Körper K ist $K[x]^{m\times n}$ ein Vektorraum über K.*

Ebenso kann man zwei Polynommatrizen miteinander multiplizieren.

Definition 9.16 *Sei $A(x) = (a_{ij}(x))_{1\le i\le m, 1\le j\le n} \in K[x]^{m\times n}$ und $B(x) = (b_{jl}(x))_{1\le j\le n, 1\le l\le p} \in K[x]^{n\times p}$. Dann bezeichnet die Matrix*

$$C(x) = (c_{il}(x))_{1\le i\le m, 1\le l\le p} \in K[x]^{m\times p}$$

mit den Eintragungen

$$c_{il}(x) = \sum_{j=1}^{n} a_{ij}(x) b_{jl}(x)$$

das Produkt

$$C(x) = A(x)B(x).$$

Es gelten dieselben Rechenregeln wie bei gewöhlichen Matrizen:

$$(A(x)B(x))C(x) = A(x)(B(x)C(x)),\ A(x)(B(x) + C(x)) = A(x)B(x) + A(x)C(x), \ldots$$

Insbesondere bilden quadratische Polynommatrizen eine K-Algebra.

Lemma 9.17 *Für jeden Körper bildet $K[x]^{n\times n}$ eine K-Algebra.*

Entsprechend ist das Einsetzen (von Körperelementen in x) ein K-Algebrenhomomorphismus.

Es können aber in x nicht nur Körperelemente eingesetzt werden, sondern auch Elemente einer K-Algebra, beispielsweise lineare Abbildungen $f \in L(\mathbf{V}, \mathbf{V})$, wobei $\mathbf{V}$ ein Vektorraum über K ist.

Definition 9.18 *Sei K ein Körper und $A(x) = (a_{ij}(x))_{1\le i,j\le n} \in K[x]^{n\times n}$ eine quadratische Polynommatrix. Dann bezeichnet*

$$\det A(x) := \sum_{\pi \in \mathbf{S}_n} \operatorname{sgn}(\pi) \cdot a_{\pi(1)1}(x) \cdots a_{\pi(n),n}(x)$$

die **Determinante** *von $A(x)$.*

Offensichtlich besagt $\det A(x) \neq 0$ nicht unbedingt, daß $A(x)$ invertierbar ist. Allerdings gilt der folgende Satz:

Satz 9.19 *Sei $A(x) = (a_{ij}(x))_{1\leq i,j\leq n} \in K[x]^{n\times n}$ eine quadratische Polynommatrix und bezeichne $A_{ij}(x)$ den* **Kofaktor** *von $A(x)$, also die Determinante jener Matrix, die dadurch hervorgeht, daß man die j-ten Spalte von $A(x)$ durch $\mathbf{e}_i$ ersetzt, so gilt*

$$A(x)\hat{A}(x)^T = \det A(x) \cdot E_n \qquad \text{mit } \hat{A}(x) = (A_{ij}(x))_{1\leq i,j\leq n}.$$

9.2.2 Das charakteristische Polynom einer Matrix

Definition 9.20 *Sei K ein Körper und $A = (a_{ij})_{1\leq i,j\leq n} \in K^{n\times n}$ eine quadratische Matrix. Das Polynom*

$$\chi_A(x) := \det(A - xE_n) = \det(a_{ij} - x\delta_{ij})_{1\leq i,j\leq n}$$

heißt **charakteristisches Polynom** *von A.*

Definition 9.21 *Die* **Spur** $\mathrm{spur}(A)$ *einer quadratischen Matrix $A = (a_{ij})_{1\leq i,j\leq n} \in K^{n\times n}$ist durch*

$$\mathrm{spur}(A) := \sum_{i=1}^{n} a_{ii}$$

definiert

Satz 9.22 *Das charakteristische Polynom $\chi_A(x)$ einer Matrix $A = (a_{ij})_{1\leq i,j\leq n} \in K^{n\times n}$ ist ein Polynom n-ten Grades der Form*

$$\chi_A(x) = (-1)^n x^n + (-1)^{n-1}\mathrm{spur}(A)x^{n-1} + \cdots + \det(A).$$

Die wichtigste Eigenschaft des charakteristischen Polynoms ist, daß es ermöglicht, die Eigenwerte von A zu berechnen.

Satz 9.23 *Sei $A \in K^{n\times n}$ eine quadratische Matrix und $\chi_A(x)$ das charakteristische Polynom von A. Dann gilt*

$$t \in \mathrm{spec}(A) \Longleftrightarrow \chi_A(t) = 0,$$

d.h. die Nullstellen von $\chi_A(x)$ sind die Eigenwerte von A.

Weiters sind die Eigenräume von A (für $t \in \mathrm{spec}(A)$) durch

$$\mathbf{V}_t = \{\mathbf{a} \in K^{n\times 1} \mid (A - tE_n)\mathbf{a} = \mathbf{0}\} = \mathrm{kern}(f_A - t\mathrm{id})$$

gegeben.

Kennt man also einen Eigenwert t von A, so ist das Bestimmen von $(K^{n\times 1})_t$ das Lösen eines homogenen linearen Gleichungssystems.

Korollar 9.24 *Zerfällt das charakteristische Polynom $\chi_A(x)$ einer Matrix $A \in K^{n \times n}$ in Linearfaktoren,*

$$\chi_A(x) = (-1)^n \prod_{i=1}^{n} (x - t_i),$$

d.h. $t_1, t_2, \dots, t_n$ sind die Eigenwerte von A, dann gilt

$$\sum_{i=1}^{n} t_i = \mathrm{spur}(A) \qquad \textit{und} \qquad \prod_{i=1}^{n} t_i = \det(A).$$

9.2.3 Das charakteristische Polynom einer Abbildung

Auf endlichdimensionalen Vektorräumen kann auch das charakteristische Polynom einer Abbildung $f \in L(\mathbf{V}, \mathbf{V})$ definiert werden:

Definition 9.25 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum. Das* **charakteristische Polynom** *einer linearen Abbildung $f \in L(\mathbf{V}, \mathbf{V})$ ist durch*

$$\chi_f(x) := \det(\Phi_{\mathbf{BB}}(f) - xE_n)$$

definiert, wobei $\mathbf{B}$ eine Basis von $\mathbf{V}$ und $n = \dim(\mathbf{V})$ bezeichnen.

Daß das charakteristische Polynom einer Abbildung nicht von der Wahl der gewählten Basis $\mathbf{B}$ abhängt, sichert die folgende Eigenschaft.

Lemma 9.26 *Sei $A \in K^{n \times n}$ eine quadratische Matrix und $B = T^{-1}AT$, wobei $T \in K^{n \times n}$ eine reguläre Matrix ist. Dann stimmen die charakteristischen Polynome von A und B überein:*

$$\chi_A(x) = \chi_B(x).$$

Definition 9.27 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum. Die* **Spur** $\mathrm{spur}(f)$ *einer linearen Abbildung $f \in L(\mathbf{V}, \mathbf{V})$ ist durch*

$$\mathrm{spur}(f) := \mathrm{spur}(\Phi_{\mathbf{BB}}(f))$$

definiert, wobei $\mathbf{B}$ eine Basis von $\mathbf{V}$ bezeichnet. Entsprechend ist die Determinante von f durch

$$\det(f) := \det(\Phi_{\mathbf{BB}}(f))$$

gegeben.

Offensichtlich hängen diese beiden Begriffe nicht von der Wahl der Basis $\mathbf{B}$ ab.

Satz 9.28 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum. Das charakteristische Polynom $\chi_f(x)$ einer linearen Abbildung $f \in L(\mathbf{V}, \mathbf{V})$ ist ein Polynom n-ten Grades der Form*

$$\chi_f(x) = (-1)^n x^n + (-1)^{n-1}\mathrm{spur}(f)x^{n-1} + \cdots + \det(f).$$

Satz 9.29 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum und* $f \in L(\mathbf{V}, \mathbf{V})$*. Dann gilt*

$$t \in \operatorname{spec}(f) \Longleftrightarrow \chi_f(t) = 0,$$

d.h. die Nullstellen von $\chi_f(x)$ *sind die Eigenwerte von* f*.*

Damit erhält man einen zweiten Beweis von Korollar 9.9, daßein Polynom n-ten Grades höchstens n verschiedene Nullstellen hat. Weiters gilt:

Korollar 9.30 *Sei* $\mathbf{V}$ *ein endlichdimensionlaler Vektorraum. Zerfällt das charakteristische Polynom* $\chi_f(x)$ *einer linearen Abbildung* $f \in L(\mathbf{V}, \mathbf{V})$ *in Linearfaktoren,*

$$\chi_f(x) = (-1)^n \prod_{i=1}^{n} (x - t_i),$$

d.h. $t_1, t_2, \ldots, t_n$ *sind die Eigenwerte von* f*, dann gilt*

$$\sum_{i=1}^{n} t_i = \operatorname{spur}(f) \qquad \textit{und} \qquad \prod_{i=1}^{n} t_i = \det(f).$$

9.2.4 Algebraische und geometrische Vielfachheit

Definition 9.31 *Die* **algebraische Vielfachheit** $\lambda_a(t)$ *eines Eigenwertes* $t \in \operatorname{spec}(A)$ *einer quadratischen Matrix* $A \in K^{n \times n}$ *ist die Vielfachheit von* t *als Nullstelle des charakteristischen Polynoms* $\chi_A(x)$*.*

Die **geometrische Vielfachheit** $\lambda_g(t)$ *eines Eigenwertes* $t \in \operatorname{spec}(A)$ *einer quadratischen Matrix* $A \in K^{n \times n}$ *ist die Dimension des Eigenraums* $\mathbf{V}_t \leq K^{n \times 1}$*:*

$$\lambda_g(t) := \dim \mathbf{V}_t.$$

$\lambda_g(t)$ ist daher auch die Dimension des Lösungsraums des linearen Gleichungssystems $(A - tE_n)\mathbf{x} = \mathbf{0}$.

Definition 9.32 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum. Die* **algebraische Vielfachheit** $\lambda_a(t)$ *eines Eigenwertes* $t \in \operatorname{spec}(f)$ *einer linearen Abbildung* $f \in L(\mathbf{V}, \mathbf{V})$ *ist die Vielfachheit von* t *als Nullstelle des charakteristischen Polynoms* $\chi_f(x)$*.*

Die **geometrische Vielfachheit** $\lambda_g(t)$ *eines Eigenwertes* $t \in \operatorname{spec}(f)$ *einer linearen Abbildung* $f \in L(\mathbf{V}, \mathbf{V})$ *ist die Dimension des Eigenraums* $\mathbf{V}_t \leq \mathbf{V}$*:*

$$\lambda_g(t) := \dim \mathbf{V}_t.$$

Wegen $\mathbf{V}_t = \operatorname{kern}(f - t\mathrm{id})$ gilt auch $\lambda_g(t) = \operatorname{def}(f - t\mathrm{id})$.

Satz 9.33 *Die geometrische Vielfachheit* $\lambda_g(t)$ *eines Eigenwerts* $t \in \operatorname{spec}(A)$ *resp.* $t \in \operatorname{spec}(f)$ *ist nie größer als die algebraische Vielfachheit:*

$$\lambda_g(t) \leq \lambda_a(t).$$

9.2.5 Summen von Unterräumen

Es soll nun untersucht werden, wie das charakteristische Polynom einer auf einen f-invarianten Unterraum eingeschränkten Abbildung mit dem ursprünglichen charakteristischen Polynom zusammenhängt.

Satz 9.34 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum und* $f \in L(\mathbf{V}, \mathbf{V})$. *Ist* $\mathbf{V} = \mathbf{U} \oplus \mathbf{W}$ *direkte Summe von zwei* f*-invarianten Unterräumen, so gilt*

$$\chi_f(x) = \chi_{f|_\mathbf{U}}(x) \cdot \chi_{f|_\mathbf{W}}(x).$$

9.3 Diagonalisierbarkeit

9.3.1 Einfach strukturierte Abbildungen

Definition 9.35 *Eine lineare Abbildung* $f \in L(\mathbf{V}, \mathbf{V})$ *heißt* **einfach strukturiert**, *wenn es eine Basis* $\mathbf{B}$ *von* $\mathbf{V}$ *gibt, die nur aus Eigenvektoren von* f *besteht.*

Satz 9.36 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum. Eine Abbildung* $f \in L(\mathbf{V}, \mathbf{V})$ *ist genau dann einfach strukturiert, wenn es eine Basis* $\mathbf{B}$ *von* $\mathbf{V}$ *gibt, so daß* $\Phi_{\mathbf{BB}}(f)$ *eine Diagonalmatrix ist:*

$$\Phi_{\mathbf{BB}}(f) = \mathrm{diag}(t_1, t_2, \ldots, t_n).$$

$t_1, t_2, \ldots, t_n$ *sind dabei die Eigenwerte von* f.

9.3.2 Ähnliche Matrizen

Definition 9.37 *Zwei Matrizen* $A, B \in K^{n\times n}$ *heißen* **ähnlich**, *wenn es eine reguläre Matrix* $T \in K^{n\times n}$ *mit*

$$B = T^{-1}AT$$

gibt.

Beispielsweise sind die Koordinatenmatrizen $\Phi_{\mathbf{B}_1,\mathbf{B}_1}(f)$, $\Phi_{\mathbf{B}_2,\mathbf{B}_2}(f)$ einer linearen Abbildung $f \in L(\mathbf{V}, \mathbf{V})$ bezüglich verschiedener Basen $\mathbf{B}_1$, $\mathbf{B}_2$ von $\mathbf{V}$ ähnlich. Insbesondere stimmen die charakteristischen Polynome ähnlicher Matrizen überein.

9.3.3 Diagonalisierbarkeit

Definition 9.38 *Eine Matrix* $A \in K^{n\times n}$ *heißt* **diagonalisierbar**, *wenn es eine reguläre Matrix* $T \in K^{n\times n}$ *gibt, so daß*

$$T^{-1}AT = \mathrm{diag}(t_1, t_2, \ldots, t_n)$$

eine Diagonalmatrix $\mathrm{diag}(t_1, t_2, \ldots, t_n)$ *ist.*

Die Werte $t_1, t_2, \ldots, t_n$ sind dabei die Eigenwerte von f.

Satz 9.39 *Eine Matrix $A \in K^{n\times n}$ ist genau dann diagonalisierbar, wenn die dazugehörige Abbildung $f_A : K^{n\times 1} \to K^{n\times 1}$, $\mathbf{x} \mapsto A\mathbf{x}$, einfach strukturiert ist.*

Satz 9.40 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum. Eine lineare Abbildung $f \in L(\mathbf{V},\mathbf{V})$ ist genau dann einfach strukturiert, wenn die Matrix $\Phi_{\mathbf{BB}}(f)$ für jede Wahl einer Basis $\mathbf{B}$ von $\mathbf{V}$ diagonalisierbar ist.*

Satz 9.41 *Sei $\mathbf{V}$ ein Vektorraum mit endlicher Dimension $\dim(\mathbf{V}) = n < \infty$ und $f \in L(\mathbf{V},\mathbf{V})$. Hat f n verschiedene Eigenwerte, so ist f einfach strukturiert.*

9.4 Minimalpolynom

9.4.1 Annullatorpolynom und Minimalpolynom

Definition 9.42 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum und $f \in L(\mathbf{V},\mathbf{V})$ eine lineare Abbildung. Ein Polynom $p(x) \in K[x] \setminus \{0\}$ heißt* **Annullatorpolynom** *von f, wenn*

$$p(f) = 0.$$

Lemma 9.43 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum und $f \in L(\mathbf{V},\mathbf{V})$ eine lineare Abbildung. Dann gibt es ein Annullatorpolynom von f.*

Definition 9.44 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum und $f \in L(\mathbf{V},\mathbf{V})$ eine lineare Abbildung. Ein normiertes Polynom $\mu(x) \in K[x]$ heißt* **Minimalpolynom** *von f, wenn $\mu(x)$ Annullatorpolynom von f vom kleinstmöglichen Grad ist.*

Satz 9.45 *Jede lineare Abbildung $f \in L(\mathbf{V},\mathbf{V})$ auf einem endlichdimensionalen Vektorraum besitzt ein eindeutig bestimmtes Minimalpolynom $\mu_f(x)$. Weiters ist ein Polynom $p(x)$ genau dann Annullatorpolynom von f, wenn das Minimalpolynom $\mu_f(x)$von f ein Teiler von $p(x)$ ist: $\mu_f(x)|p(x)$.*

Wegen der Eindeutigkeit spricht man auch von *dem* Minimalpolynom (und bezeichnet es mit $\mu_f(x)$).

9.4.2 Satz von Cayley-Hamilton

Überraschenderweise ist $\mu_f(x)$ ein Teiler des charakteristischen Polynoms. Dies ist die Aussage des Satzes von **Cayley-Hamilton**.

Satz 9.46 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum und $f \in L(\mathbf{V},\mathbf{V})$ eine lineare Abbildung. Dann ist das charakteristische Polynom $\chi_f(x)$ ein Annullatorpolynom von f, d.h.*

$$\mu_f(x)|\chi_f(x)$$

9.4.3 Zerlegung des Minimalpolynoms

Ähnlich wie beim charakteristischen Polynom gibt es auch beim Minimalpolynom eine *Produktformel.*

Satz 9.47 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum,* $f \in L(\mathbf{V}, \mathbf{V})$ *und* $\mathbf{U}, \mathbf{W}$ *zwei* f*-invariante Unterräume von* $\mathbf{V}$ *mit* $\mathbf{V} = \mathbf{U} \oplus \mathbf{W}$*, so daß die charakteristischen Polynome von* $\chi_{f|_{\mathbf{U}}}(x)$ *und* $\chi_{f|_{\mathbf{W}}}(x)$ *teilerfremd sind. Dann gilt*

$$\mu_f(x) = \mu_{f|_{\mathbf{U}}}(x) \cdot \mu_{f|_{\mathbf{W}}}(x).$$

Es wird sich herausstellen, daß man gemäß der Zerlegung des charakteristischen Polynoms

$$\chi_f(x) = (-1)^n \prod_{t \in \mathrm{spec}(f)} (x-t)^{\lambda_a(t)}$$

f-invariante Unterräume $\overline{\mathbf{V}}_t$ mit $\dim(\overline{\mathbf{V}}_t) = \lambda_a(t)$ und

$$\chi_{f|_{\overline{\mathbf{V}}_t}}(x) = (t-x)^{\lambda_a(t)}$$

finden kann. Daher wird es das erste Ziel sein, Abbildungen $f \in L(\mathbf{V}, \mathbf{V})$ zu untersuchen, derer charakteristisches Polynom von der Gestalt $\chi_f(x) = (t-x)^n$ ist. Der allgemeine Fall kann dann aus Spezialfällen dieser Art zurückgeführt werden. Insbesondere gilt für das Minimalpolynom:

Satz 9.48 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum und* $f \in L(\mathbf{V}, \mathbf{V})$ *mit* $\chi_f(x) = (t-x)^n$*. Dann gilt*

$$\mu_f(x) = (x-t)^m$$

mit

$$m = \min\{r \geq 1 \,|\, \mathrm{im}((f - t\mathrm{id})^r) = \{\mathbf{0}\}\}.$$

Wegen des Satzes von Cayley-Hamilton gilt sicherlich $m \leq n$, das Mimimum existiert in jedem Fall.

9.5 Jordansche Normalform

9.5.1 Primärzerlegung

Lemma 9.49 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum,* $f \in L(\mathbf{V}, \mathbf{V})$ *und* $t \in \mathrm{spec}(f)$*. Weiter sei* $r \geq 1$ *eine natürliche Zahl. Dann gelten die folgenden drei Eigenschaften:*

1. *Die Unterräume* $\mathrm{kern}((f - t\mathrm{id})^r)$ *und* $\mathrm{im}((f - t\mathrm{id})^r)$ *sind* f*-invariante Unterräume.*
2. $\mathrm{kern}((f - t\mathrm{id})^r) \oplus \mathrm{im}((f - t\mathrm{id})^r) = \mathbf{V}$ *für* $r \geq \mu(t)$*.*
3. *Ist* $s \in \mathrm{spec}(f) \setminus \{t\}$*, dann gilt* $\mathbf{V}_s \cap \mathrm{kern}((f - t\mathrm{id})^r) = \{\mathbf{0}\}$*.*

Definition 9.50 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum,* $f \in L(\mathbf{V},\mathbf{V})$ *und* $t \in \operatorname{spec}(f)$*. Dann bezeichnet*

$$\overline{\mathbf{V}}_t = \overline{\mathbf{V}}_t(f) := \operatorname{kern}((f - t\mathrm{id})^{\mu(t)})$$

den **Hauptraum** *von* f *bezüglich* $t \in \operatorname{spec}(f)$*, wobei* $\mu(t) \geq 1$ *durch*

$$\mu(t) := \min\{r \geq 1 \,|\, \operatorname{kern}((f - t\mathrm{id})^r) = \operatorname{kern}((f - t\mathrm{id})^{r+1})\}$$

gegeben ist.

$\overline{\mathbf{V}}_t$ ist also ein f-invarianter Unterraum von $\mathbf{V}$ und $\mathbf{V}_t$ ist ein f-invarianter Unterraum von $\overline{\mathbf{V}}_t$: $\mathbf{V}_t \leq \overline{\mathbf{V}}_t \leq \mathbf{V}$. Weiters gilt Gleicheit $\mathbf{V}_t = \overline{\mathbf{V}}_t$ genau dann, wenn $\mu(t) = 1$ ist.

Die Haupträume leisten genau das, was im vorigen Abschnitt angedeutet wurde.

Satz 9.51 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über dem Körper* K *und* $f \in L(\mathbf{V},\mathbf{V})$ *eine lineare Abbildung, deren charakteristisches Polynom über* K *in Linearfaktoren zerfällt. Dann gelten die folgenden drei Eigenschaften:*

1. $\mathbf{V} = \bigoplus_{t\in\operatorname{spec}(\mathbf{V})} \overline{\mathbf{V}}_t$.
2. $\dim(\overline{\mathbf{V}}_t) = \lambda_a(t), \quad (t \in \operatorname{spec}(f))$.
3. $\chi_{f|_{\overline{\mathbf{V}}_t}}(x) = (t - x)^{\lambda_a(t)}$.

Korollar 9.52 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum und* $f \in L(\mathbf{V},\mathbf{V})$ *eine lineare Abbildung, deren charakteristisches Polynom über* K *in Linearfaktoren zerfällt. Dann ist das Minimalpolynom von* f *durch*

$$\mu_f(x) = \prod_{t\in\operatorname{spec}(f)} (x - t)^{\mu(t)}$$

gegeben.

Jeder Vektor $\mathbf{d} \in \overline{\mathbf{V}}_t$ hat die Eigenschaft, daß $(f - t\mathrm{id})^{\mu(t)}(\mathbf{d}) = \mathbf{0}$ gilt. Das kleinste $m \geq 1$ mit dieser Eigenschaft $(f - t\mathrm{id})^m(\mathbf{d}) = \mathbf{0}$ wird auch als **Stufe** von $\mathbf{d}$ bezeichnet, d.h.

$$m = \min\{r \geq 1 \,|\, (f - t\mathrm{id})^r(\mathbf{d}) = \mathbf{0}\}.$$

Offensichtlich ist dann der Vektor

$$\mathbf{c} := (f - t\mathrm{id})^{m-1}(\mathbf{d}) \neq \mathbf{0}$$

und es gilt

$$f(\mathbf{c}) = t\mathbf{c},$$

d.h. $\mathbf{c}$ ist Eigenvektor von f zum Eigenwert t. Es zeigt sich, daß die (Eigen-)Vektoren

$$(f - t\mathrm{id})^{m-1}(\mathbf{d}),\ (f - t\mathrm{id})^{m-2}(\mathbf{d}), \ldots,\ (f - t\mathrm{id})^{m-m}(\mathbf{d}) = \mathbf{d}$$

linear unabhängig sind Es gilt sogar noch mehr.

Lemma 9.53 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum,* $f \in L(\mathbf{V}, \mathbf{V})$ *und* $t \in \operatorname{spec}(f)$*. Sind weiters* $\mathbf{d}_1, \ldots \mathbf{d}_k \in \overline{\mathbf{V}}_t$ *und* $m_1, \ldots, m_k$ *die entsprechenden Stufen, d.h.*

$$m_j = \min\{r \geq 1 \,|\, (f - t\mathrm{id})^r(\mathbf{d}_j) = \mathbf{0}\},$$

so daß die Vektoren

$$\mathbf{c}_1 = (f - t\mathrm{id})^{m_1 - 1}(\mathbf{d}_1), \ldots,\ \mathbf{c}_k = (f - t\mathrm{id})^{m_k - 1}(\mathbf{d}_k)$$

linear unabhängig sind, dann ist auch das System der Vektoren

$$\mathbf{d}_{ij} := (f - t\mathrm{id})^{m_j - i}(\mathbf{d}_j) \qquad (1 \leq j \leq k,\ 1 \leq i \leq m_j)$$

linear unabhängig.

9.5.2 Jordansche Normalform

Definition 9.54 *Eine* $m \times m$*-Matrix*

$$J_m(t) := \begin{pmatrix} t & 1 & 0 & 0 & \cdots & 0 \\ 0 & t & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \ldots & 0 & t & 1 \\ 0 & 0 & \ldots & 0 & 0 & t \end{pmatrix}$$

heißt **Jordan-Kästchen** *der Dimension* m *zum Eigenwert* $t \in K$*.*

Ist $\mathbf{d} \in \overline{\mathbf{V}}_t$ ein Vektor der Stufe m, so erfüllen die Vektoren

$$\mathbf{c}_i := (f - t\mathrm{id})^{m-i}(\mathbf{d}) \qquad (1 \leq i \leq m)$$

die Beziehungen

$$\begin{aligned} f(\mathbf{c}_1) &= t\mathbf{c_1}, \\ f(\mathbf{c}_2) &= (f - t\mathrm{id})(\mathbf{c}_2) + t\mathbf{c_2} = \mathbf{c}_1 + t\mathbf{c}_2, \\ &\vdots \\ f(\mathbf{c}_m) &= \mathbf{c}_{m-1} + t\mathbf{c}_m. \end{aligned}$$

Für den Unterraum $\mathbf{U} := [\{\mathbf{c}_1, \ldots, \mathbf{c_m}\}]$ mit der Basis $\mathbf{B} = \{\mathbf{c}_1, \ldots, \mathbf{c_m}\}$ ist die Koordinatenmatrix für $f|_{\mathbf{U}}$ durch

$$\Phi_{\mathbf{BB}}(f|_{\mathbf{U}}) = J_m(t)$$

gegeben. Gelingt es also, den Vektorraum $\mathbf{V}$ als direkte Summe solcher Unterräume darzustellen, so existiert eine Basis von $\mathbf{V}$, so daß die Koordinatenmatrix von $f \in L(\mathbf{V}, \mathbf{V})$ nur aus Jordan-Kästchen besteht, die an der Diagonale hintereinander aufgereiht sind. Dies kann tatsächlich erreicht werden.

Satz 9.55 *Sei* $\mathbf{V}$ *ein* n*-dimensionaler Vektorraum und* $f \in L(\mathbf{V}, \mathbf{V})$ *mit* $\chi_f(x) = (t - x)^n$*. Dann gibt es eine Basis* $\mathbf{B}$ *von* $\mathbf{V}$*, und natürliche Zahlen* $m_1 > m_2 > \cdots > m_p > 0$ *und* $k_1, k_2, \ldots, k_p > 0$*, so daß* $\Phi_{\mathbf{BB}}(f)$ *erweiterte Diagonalform*

$$\Phi_{\mathbf{BB}}(f) = \operatorname{diag}(J_{m_1}(t), \cdots, J_{m_1}(t), J_{m_2}(t), \ldots, J_{m_2}(t), \ldots, J_{m_p}(t), \ldots, J_{m_p}(t))$$

hat, wobei das Jordan-Kästchen $J_{m_1}(t)$ k_1*-mal vorkommt, das Jordan-Kästchen* $J_{m_2}(t)$ k_2*-mal etc. Die Zahlen* p*,* $m_1, m_2, \ldots, m_p$ *und* $k_1, k_2, \ldots, k_p$ *sind durch die Beziehungen*

$$\dim((f - t\mathrm{id})^j(\mathbf{V}) \cap \operatorname{kern}(f - t\mathrm{id})) = \begin{cases} k_1 + \cdots + k_p & \textit{für } j \in \{0, \ldots, m_p - 1\}, \\ k_1 + \cdots + k_{p-1} & \textit{für } j \in \{m_p, \ldots, m_{p-1} - 1\}, \\ \quad \vdots & \\ k_1 & \textit{für } j \in \{m_2, \ldots, m_1 - 1\}, \\ 0 & \textit{für } j \in \{m_1, m_1 + 1, \ldots\} \end{cases}$$

gegeben.

Man beachte, daß

$$p = \dim(\mathbf{V}_t)$$

ist und daß die Zahen $m_1, m_2, \ldots, m_p$ und $k_1, k_2, \ldots, k_p$ eindeutig bestimmt sind, d.h. jede andere Koordinatenmatrix von f, die nur aus Jordan-Kästchen aufgebaut ist, unterscheidet sich von der angegeben Form nur durch die Reihenfolge der Kästchen. In beiden Fällen spricht man von einer **Jordanschen Normalform**

Definition 9.56 *Eine (erweiterte) Diagonalmatrix, die nur aus Jordan-Kästchen besteht, heißt* **Jordansche Normalform***.*

Insgesamt erhält man folgende Charakterisierung.

Satz 9.57 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über einem Körper* K *und* $f \in L(\mathbf{V}, \mathbf{V})$*, so daß das charakteristische Polynom* $\chi_f(x)$ *über* K *in Linearfaktoren zerfällt. Dann gibt es eine Basis* $\mathbf{B}$ *von* $\mathbf{V}$*, so daß* $\Phi_{\mathbf{BB}}(f)$ *Jordansche Normalform hat, wobei die auftretenden Jordan-Kästchen (bis auf die Reihenfolge) durch die Abbildung* f *eindeutig bestimmt sind.*

9.5.3 Klassifikation von ähnlichen Matrizen

Satz 9.58 *Jede Matrix* $A \in K^{n \times n}$*, derer charakteristisches Polynom über* K *in Linearfaktoren zerfällt, ist zu einer Jordanschen Normalform ähnlich.*

Mit Hilfe der Jordanschen Normalform können ähnliche Matrizen klassifizert werden.

Satz 9.59 *Zwei Matrizen* $A, B \in K^{n \times n}$*, deren charakteristische Polynome über* K *in Linearfaktoren zerfallen, sind genau dann ähnlich, wenn sie zur selben Jordanschen Normalform ähnlich sind.*

Der einzige Nachteil dieses Satzes ist es, daß er nur für Matrizen gilt, deren charakteristisches Polynom in Linearfaktoren zerfällt. Dies ist insbesondere der Fall, wenn $K = \mathbb{C}$ gilt.

Satz 9.60 *Zwei Matrizen $A, B \in \mathbb{C}^{n\times n}$ sind genau dann ähnlich, wenn sie zur selben Jordanschen Normalform ähnlich sind.*

Im allgemeinen Fall müßte man den Körper K um die Nullstellen von $\chi_f(x)$ ergänzen. Dies ist bei den reellen Zahlen besonders einfach.

Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum über $\mathbb{R}$ und $f \in L(\mathbf{V}, \mathbf{V})$, deren charakteristisches Polynom $\chi_f(x)$ nicht in Linearfaktoren (über $\mathbb{R}$) zerfällt. In diesem Fall erweitert man f zu einer Abbildung $f_\mathbb{C} \in L(\mathbf{V}_\mathbb{C}, \mathbf{V}_\mathbb{C})$. Sei nun $t = \sigma + i\tau$ ein nichtreeller Eigenwert von $f_\mathbb{C}$ und $\mathbf{d} \in \overline{\mathbf{V}}_t$ ein Vektor der Stufe m, so erfüllen die Vektoren

$$\mathbf{c}_j := (f - t\mathrm{id})^{m-i}(\mathbf{d}) = \mathbf{a}_j + i\mathbf{b}_j \qquad (1 \le j \le m)$$

wie vorhin diskutiert die Beziehungen

$$\begin{aligned} f_\mathbf{C}(\mathbf{c}_1) &= t\mathbf{c_1}, \\ f_\mathbf{C}(\mathbf{c}_2) &= (f - t\mathrm{id})(\mathbf{c}_2) + t\mathbf{c_2} = \mathbf{c}_1 + t\mathbf{c}_2, \\ &\vdots \\ f_\mathbf{C}(\mathbf{c}_m) &= \mathbf{c}_{m-1} + t\mathbf{c}_m. \end{aligned}$$

Da $\chi_f(x)$ ein reelles Polynom ist, besitzt $f_\mathbb{C}$ auch die konjugiert komplexe Zahl $\bar{t} = \sigma - i\tau$ als Eigenwert, und wir bekommen entsprechende Beziehungen für $\bar{t}$ und $\overline{\mathbf{c}}_j = \mathbf{a}_j - \mathbf{b}_j$. Die reellen Vektoren $\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}_1, \ldots, \mathbf{b}_m$ spannen denselben Unterraum auf wie die Vektoren $\mathbf{c}_1, \ldots, \mathbf{c}_m, \overline{\mathbf{c}}_1, \ldots, \overline{\mathbf{c}}_m$. Außerdem kann man durch Vergleich von Real- und Imaginärteil ablesen, wie f auf $\mathbf{a}_j$ und $\mathbf{b}_j$ operiert:

$$\begin{aligned} f(\mathbf{a}_1) + if(\mathbf{b}_1) &= (\sigma\mathbf{a}_1 - \tau\mathbf{b}_1) + i(\sigma\mathbf{b}_1 + \tau\mathbf{a}_1), \\ f(\mathbf{a}_2) + if(\mathbf{b}_2) &= (\mathbf{a}_1 + \sigma\mathbf{a}_2 - \tau\mathbf{b}_2) + i(\mathbf{b}_1 + \sigma\mathbf{b}_2 + \tau\mathbf{a}_2), \\ &\vdots \\ f(\mathbf{a}_m) + if(\mathbf{b}_m) &= (\mathbf{a}_{m-1} + \sigma\mathbf{a}_m - \tau\mathbf{b}_m) + i(\mathbf{b}_{m-1} + \sigma\mathbf{b}_m + \tau\mathbf{a}_m). \end{aligned}$$

Dieser $2m$-dimensionale Unterraum $\mathbf{U}$ ist auch f-invariant und die Koordinatenmatrix für $f|_\mathbf{U}$ bezüglich der Basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}_1, \ldots, \mathbf{b}_m\}$ ist durch

$$\begin{pmatrix} J_m(\sigma) & \tau E_m \\ -\tau E_m & J_m(\sigma) \end{pmatrix}$$

gegeben. Man kann also immer noch eine Normalform angeben.

Definition 9.61 *Eine (erweiterte) Diagonalmatrix, die nur aus Jordan-Kästchen und aus Kästchen der Form*

$$\begin{pmatrix} J_m(\sigma) & \tau E_m \\ -\tau E_m & J_m(\sigma) \end{pmatrix}$$

besteht, heißt **reelle Jordansche Normalform***.*

Insgesamt erhält man folgende Charakterisierung.

Satz 9.62 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über einem Körper* $\mathbb{R}$ *und* $f \in L(\mathbf{V}, \mathbf{V})$. *Dann gibt es eine Basis* $\mathbf{B}$ *von* $\mathbf{V}$, *so daß* $\Phi_{\mathbf{BB}}(f)$ *reelle Jordansche Normalform hat, wobei die auftretenden Jordan-Kästchen und Kästchen der Form*

$$\begin{pmatrix} J_m(\sigma) & \tau E_m \\ -\tau E_m & J_m(\sigma) \end{pmatrix}$$

(bis auf die Reihenfolge) durch die Abbildung f *eindeutig bestimmt sind.*

Satz 9.63 *Zwei Matrizen* $A, B \in \mathbb{R}^{n \times n}$ *sind genau dann ähnlich, wenn sie zur selben reellen Jordanschen Normalform ähnlich sind.*

Kapitel 10

Bilinearformen

10.1 Bilinearformen und kongruente Matrizen

10.1.1 Bilinearformen und $L(\mathbf{V}, \mathbf{V}^*)$

Ist $A \in K^{n\times n}$ eine quadratische Matrix, so kann man Vektoren $\mathbf{x}, \mathbf{y} \in K^{n\times 1}$ von links und von rechts mit A multiplizieren:

$$\mathbf{x}^T A\, \mathbf{y}.$$

Das Resultat ist in K. Hält man $\mathbf{x}$ fest so ist die Zuordnung $\mathbf{y} \mapsto \mathbf{x}^T A\mathbf{y}$ linear und ebenso bei festgehaltenem $\mathbf{y}$ die Zuordung $\mathbf{x} \mapsto \mathbf{x}^T A\mathbf{y}$. Man erhält eine sogenannte Bilinearform.

Definition 10.1 *Sei $\mathbf{V}$ ein Vektorraum über dem Körper K. Eine Abbildung $\sigma : \mathbf{V} \times \mathbf{V} \to K$ heißt* **Bilinearform**, *wenn sie in beiden Komponenten linear ist, d.h. für $\mathbf{a}, \mathbf{a}_1, \mathbf{a}_2, \mathbf{b}, \mathbf{b}_1, \mathbf{b}_2 \in \mathbf{V}$ und $x, y \in K$ gilt*

$$\begin{aligned} \sigma(x\mathbf{a}_1 + y\mathbf{a}_2, \mathbf{b}) &= x\sigma(\mathbf{a}_1, \mathbf{b}) + y\sigma(\mathbf{a}_2, \mathbf{b}), \\ \sigma(\mathbf{a}, x\mathbf{b}_1 + y\mathbf{b}_2) &= x\sigma(\mathbf{a}, \mathbf{b}_1) + y\sigma(\mathbf{a}, \mathbf{b}_2). \end{aligned}$$

Die Menge aller Bilinearformen $\sigma : \mathbf{V} \times \mathbf{V} \to K$ wird mit $BL(\mathbf{V})$ bezeichnet.

$BL(\mathbf{V})$ kann ähnlich wie $\mathbf{V}^* = L(\mathbf{V}, K)$ als Vektorraum angesehen werden. Man definiert für $\sigma, \tau \in BL(\mathbf{V})$ und $x \in K$

$$\begin{aligned} (\sigma + \tau)(\mathbf{a}, \mathbf{b}) &:= \sigma(\mathbf{a}, \mathbf{b}) + \tau(\mathbf{a}, \mathbf{b}), \\ (x\sigma)(\mathbf{a}, \mathbf{b}) &:= x\,\sigma(\mathbf{a}, \mathbf{b}). \end{aligned}$$

Satz 10.2 *Sei $\mathbf{V}$ ein Vektorraum über dem Körper K. Dann ist $BL(\mathbf{V})$ als Vektorraum isomorph zu $L(\mathbf{V}, \mathbf{V}^*)$:*

$$BL(\mathbf{V}) \cong L(\mathbf{V}, \mathbf{V}^*).$$

Einer Bilinearform $\sigma \in BL(\mathbf{V})$ entspricht die lineare Abbildung $f_\sigma \in L(\mathbf{V}, \mathbf{V}^*)$ durch die Festlegung

$$\langle f_\sigma(\mathbf{b}), \mathbf{a}\rangle := \sigma(\mathbf{a}, \mathbf{b}).$$

Selbstverständlich hätte man die Zuordung $\sigma \mapsto g_\sigma$ mit

$$\langle g_\sigma(\mathbf{a}), \mathbf{b}\rangle := \sigma(\mathbf{a}, \mathbf{b})$$

betrachten können. Der Zusammenhang bei Koordinatenmatrizen ist aber einfacher, wenn man f_σ statt g_σ betrachtet.

Definition 10.3 *Sei* $\mathbf{V}$ *ein* n*-dimensionaler Vektorraum und* $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ *eine Basis von* $\mathbf{V}$*. Die* **Koordinatenmatrix** $\Phi_{\mathbf{B}}(\sigma)$ *einer Bilinearform* $\sigma \in BL(\mathbf{V})$ *ist durch*

$$\Phi_{\mathbf{B}}(\sigma) := (\sigma(\mathbf{b}_i, \mathbf{b}_j))_{1\le i,j\le n}$$

gegeben.

Lemma 10.4 *Sei* $\mathbf{V}$ *ein* n*-dimensionaler Vektorraum und* $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ *eine Basis von* $\mathbf{V}$*. Dann gilt für* $\sigma \in BL(\mathbf{V})$ *und* $\mathbf{a}, \mathbf{b} \in \mathbf{V}$

$$\sigma(\mathbf{a}, \mathbf{b}) = \Phi_{\mathbf{B}}(\mathbf{a})^T \Phi_{\mathbf{B}}(\sigma)\Phi_{\mathbf{B}}(\mathbf{b}).$$

Satz 10.5 *Sei* $\mathbf{V}$ *ein* n*-dimensionaler Vektorraum,* $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ *eine Basis von* $\mathbf{V}$ *und* $\mathbf{B}^* \subseteq \mathbf{V}^*$ *die duale Basis von* $\mathbf{B}$*. Dann gilt für* $\sigma \in BL(\mathbf{V})$ *und der entsprechenden Abbildung* $f_\sigma \in L(\mathbf{V}, \mathbf{V}^*)$

$$\Phi_{\mathbf{B}}(\sigma) = \Phi_{\mathbf{BB}^*}(f_\sigma).$$

Für g_σ gilt übrigens

$$\Phi_{\mathbf{BB}^*}(g_\sigma) = \Phi_{\mathbf{B}}(\sigma)^T = \Phi_{\mathbf{BB}^*}(f_\sigma)^T.$$

Definition 10.6 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum und* $\sigma \in BL(\mathbf{V})$*. Dann werden Rang* $\operatorname{rg}(\sigma)$ *und Defekt* $\operatorname{def}(\sigma)$ *von* σ *durch*

$$\begin{aligned} \operatorname{rg}(\sigma) &:= \operatorname{rg}(f_\sigma), \\ \operatorname{def}(\sigma) &:= \operatorname{def}(f_\sigma) \end{aligned}$$

definiert.

Man beachte, daß dann für jede Basis $\mathbf{B}$ von $\mathbf{V}$

$$\operatorname{rg}(\sigma) = \operatorname{rg}(f_\sigma) = \operatorname{rg}(\Phi_{\mathbf{BB}^*}(f_\sigma)) = \operatorname{rg}(\Phi_{\mathbf{B}}(\sigma)) = \operatorname{rg}(\Phi_{\mathbf{BB}^*}(g_\sigma)) = \operatorname{rg}(g_\sigma)$$

gilt.

10.1.2 Kongruente Matrizen

Lemma 10.7 *Sei* $\mathbf{V}$ *ein Vektorraum und* $\sigma \in BL(\mathbf{V})$ *eine Bilinearform. Dann ist für jede lineare Abbildung* $f \in L(\mathbf{V}, \mathbf{V})$ *die Abbildung*

$$\tau = \sigma \circ (f, f) : \mathbf{V} \times \mathbf{V}, \quad (\mathbf{a}, \mathbf{b}) \mapsto \sigma(f(\mathbf{a}), f(\mathbf{b}))$$

wieder eine Bilinearform (in $BL(\mathbf{V})$*).*

Definition 10.8 *Sei* $\mathbf{V}$ *ein Vektorraum. Zwei Bilinearformen* $\sigma, \tau \in BL(\mathbf{V})$ *heißen* **kongruent**, *wenn es eine bijektive lineare Abbildung* $f \in L(\mathbf{V}, \mathbf{V})$ *mit*

$$\tau(\mathbf{a}, \mathbf{b}) = \sigma(f(\mathbf{a}), f(\mathbf{b})) \quad \textit{resp.} \quad \tau = \sigma \circ (f, f)$$

gibt.

Offensichtlich ist die Relation σ *ist kongruent zu* τ eine Äquivalenzrelation auf $BL(\mathbf{V})$.

Satz 10.9 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum,* $\mathbf{B}$ *eine Basis von* $\mathbf{V}$*,* $\sigma \in BL(\mathbf{V})$ *und* $f \in L(\mathbf{V}, \mathbf{V})$*. Dann gilt für* $\tau = \sigma \circ (f, f) \in BL(\mathbf{V})$

$$\Phi_{\mathbf{B}}(\tau) = \Phi_{\mathbf{BB}}(f)^T \Phi_{\mathbf{B}}(\sigma) \Phi_{\mathbf{BB}}(f).$$

Satz 10.10 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum und* $\mathbf{B}, \tilde{\mathbf{B}}$ *zwei Basen von* $\mathbf{V}$*. Dann gilt für jede Bilinearform* $\sigma \in BL(\mathbf{V})$

$$\Phi_{\tilde{\mathbf{B}}}(\sigma) = T_{\tilde{\mathbf{B}}\mathbf{B}}^T \Phi_{\mathbf{B}}(\sigma) T_{\tilde{\mathbf{B}}\mathbf{B}}.$$

Diese beiden Sätze motivieren die folgende Definition.

Definition 10.11 *Zwei quadratische Matrizen* $A, B \in K^{n \times n}$ *heißen* **kongruent**, *wenn es eine reguläre Matrix* $P \in K^{n \times n}$ *mit*

$$B = P^T A P$$

gibt.

Offensichtlich gelten die folgenden Eigenschaften.

Satz 10.12 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum und* $\mathbf{B}$ *eine Basis von* $\mathbf{V}$*. Zwei Bilinearformen* $\sigma, \tau \in BL(\mathbf{V})$ *sind genau dann kongruent, wenn die Koordinatenmatrizen* $\Phi_{\mathbf{B}}(\sigma), \Phi_{\mathbf{B}}(\tau)$ *kongruent sind.*

Weiters sind die Koordinatenmatrizen $\Phi_{\mathbf{B}}(\sigma), \Phi_{\tilde{\mathbf{B}}}(\sigma)$ von σ bezüglich zweier Basen $\mathbf{B}, \tilde{\mathbf{B}}$ von $\mathbf{V}$ kongruent.

10.2 Orthosymmetrische Bilinearformen

10.2.1 Symmetrische und alternierende Bilinearformen

Mit Hilfe von Bilinearformen kann ein Orthogonalitätsbegriff auf einem Vektorraum eingeführt werden.

Definition 10.13 *Sei* $\mathbf{V}$ *ein Vektorraum und* $\sigma \in BL(\mathbf{V})$ *eine Bilinearform. Zwei Vektoren* $\mathbf{a}, \mathbf{b} \in \mathbf{V}$ *heißen* **orthogonal** *bezüglich* σ*, wenn* $\sigma(\mathbf{a}, \mathbf{b}) = 0$*. Man schreibt dafür auch*

$$\mathbf{a} \perp_\sigma \mathbf{b}.$$

Ist die Relation $\perp_\sigma$ *symmetrisch, d.h.*

$$\sigma(\mathbf{a}, \mathbf{b}) = 0 \iff \sigma(\mathbf{b}, \mathbf{a}) = 0,$$

dann heißt σ **orthosymmetrisch***.*

Offensichtlich ist σ orthosymmetrisch, wenn $\sigma(\mathbf{a}, \mathbf{b}) = \sigma(\mathbf{b}, \mathbf{a})$ für alle $\mathbf{a}, \mathbf{b} \in \mathbf{V}$ ist. Dieser und ähnliche Spezialfälle sollen nun genauer untersucht werden.

Definition 10.14 *Sei* $\mathbf{V}$ *ein Vektorraum und* $\sigma \in BL(\mathbf{V})$*.*

1. σ *heißt* **symmetrisch***, wenn*

$$\forall \mathbf{a}, \mathbf{b} \in \mathbf{V} : \ \sigma(\mathbf{a}, \mathbf{b}) = \sigma(\mathbf{b}, \mathbf{a}).$$

2. σ *heißt* **schiefsymmetrisch***, wenn*

$$\forall \mathbf{a}, \mathbf{b} \in \mathbf{V} : \ \sigma(\mathbf{a}, \mathbf{b}) = -\sigma(\mathbf{b}, \mathbf{a}).$$

3. σ *heißt* **alternierend***, wenn*

$$\forall \mathbf{a} \in \mathbf{V} : \ \sigma(\mathbf{a}, \mathbf{a}) = 0.$$

Satz 10.15 *Sei* $\mathbf{V}$ *ein Vektorraum über dem Körper* K*. Dann gelten für eine Bilinearform* $\sigma \in BL(\mathbf{V})$ *folgende Implikationen:*

1. σ *symmetrisch* $\Longrightarrow$ σ *orthosymmetrisch.*
2. σ *schiefsymmetrisch* $\Longrightarrow$ σ *orthosymmetrisch.*
3. σ *alternierend* $\Longrightarrow$ σ *schiefsymmetrisch.*
4. σ *schiefsymmetrisch* $\wedge \mathrm{char}(K) \neq 2 \Longrightarrow \sigma$ *alternierend.*

Man beachte, daß nur im Fall $\mathrm{char}(K) = 2$ die Begriffe *schiefsymmetrisch* und *alternierend* verschieden sind. Im Fall $\mathrm{char}(K) = 2$ stimmen allerdings die Begriffe *symmetrisch* und *schiefsymmetrisch* überein.

Für Matrizen können analoge Begriffe eingeführt werden.

Definition 10.16 *Sei* $A = (a_{ij})_{1 \le i,j \le n} \in K^{n \times n}$ *eine quadratische Matrix.*

1. A *heißt* **symmetrisch***, wenn* $A^T = A$*.*
2. A *heißt* **schiefsymmetrisch***, wenn* $A^T = -A$*.*

3. *A heißt* **alternierend**, *wenn $A^T = -A$ und alle Diagonalelemente $a_{ii} = 0$, $1 \leq i \leq n$, sind.*

Satz 10.17 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum und $\mathbf{B}$ eine Basis von $\mathbf{V}$. Dann gelten für $\sigma \in BL(\mathbf{V})$ folgende Eigenschaften:*

1. *σ symmetrisch $\iff \Phi_{\mathbf{B}}(\sigma)$ symmetrisch.*
2. *σ schiefsymmetrisch $\iff \Phi_{\mathbf{B}}(\sigma)$ schiefsymmetrisch.*
3. *σ alternierend $\iff \Phi_{\mathbf{B}}(\sigma)$ alternierend.*

10.2.2 Klassifikation orthosymmetrischer Bilinearformen

Im vorigen Abschnitt wurde erwähnt, daß jede Bilinearform, die symmetrisch oder schiefsymmetrisch ist, (offensichtlich) auch orthosymmetrisch ist. Interessanterweise hat man damit bereits alle orthosymmetrischen Bilinearformen beschrieben.

Satz 10.18 *Sei $\mathbf{V}$ ein Vektorraum. Dann ist eine Bilinearform $\sigma \in BL(\mathbf{V})$ genau dann orthosymmetrisch, wenn sie symmetrisch oder schiefsymmetrisch ist.*

Für den Nachweis dieses Satzes benötigt man einige Vorbereitungen.

Lemma 10.19 *Sei $\mathbf{V}$ ein Vektorraum und $\mathbf{a}^* \in \mathbf{V}^*$ eine Linearform. Dann gilt*

$$(\mathrm{kern}(\mathbf{a}^*))^\circ = [\mathbf{a}^*].$$

Satz 10.20 *Sei $\mathbf{V}$ ein Vektorraum. Stimmen für zwei Bilinearformen $\sigma, \tau \in BL(\mathbf{V})$ die Orthogonalitätsrelationen $\perp_\sigma, \perp_\tau$ überein, d.h.*

$$\sigma(\mathbf{a}, \mathbf{b}) = 0 \iff \tau(\mathbf{a}, \mathbf{b}) = 0,$$

dann gibt es $c \in K^\times$ mit

$$\sigma = c\tau.$$

Satz 10.18 folgt nun direkt aus Satz 10.20.

10.2.3 Orthogonalräume

Definition 10.21 *Sei $\mathbf{V}$ ein Vektorraum und $\sigma \in BL(\mathbf{V})$. Für eine nichtleere Teilmenge $\mathbf{M} \subseteq \mathbf{V}$ bezeichnen*

$$\begin{aligned} \mathbf{M}^\perp &:= \{\mathbf{b} \in \mathbf{V} \mid \forall \mathbf{a} \in \mathbf{M}: \ \sigma(\mathbf{a}, \mathbf{b}) = 0\}, \\ {}^\perp\mathbf{M} &:= \{\mathbf{a} \in \mathbf{V} \mid \forall \mathbf{b} \in \mathbf{M}: \ \sigma(\mathbf{a}, \mathbf{b}) = 0\}. \end{aligned}$$

*den Rechts- bzw. Links-***Orthogonalraum** *von $\mathbf{M}$. Die Mengen $\mathbf{V}^\perp$ und ${}^\perp\mathbf{V}$ heißen Rechts- bzw. Links-***Radikal** *von $\mathbf{V}$.*

Ist σ orthosymmetrisch, so stimmen $\mathbf{M}^{\perp}$ und ${}^{\perp}\mathbf{M}$ für alle $\mathbf{M} \subseteq \mathbf{V}$ überein.

Definition 10.22 *Eine Bilinearform $\sigma \in BL(\mathbf{V})$ auf einem Vektorraum heißt* **nichtausgeartet**, *wenn*

$$\mathbf{V}^{\perp} = {}^{\perp}\mathbf{V} = \{\mathbf{0}\}.$$

Ein Unterraum $\mathbf{U} \leq \mathbf{V}$ heißt **nichtausgeartet**, *wenn $\sigma|_{\mathbf{U}}$ auf $\mathbf{U}$ nichtausgeartet ist.*

$\mathbf{U}$ ist daher genau dann nichtausgeartet, wenn

$$\mathbf{U} \cap \mathbf{U}^{\perp} = \mathbf{U} \cap {}^{\perp}\mathbf{U} = \{\mathbf{0}\}.$$

Einige Eigenschaften von Orthogonalräumen sind im folgenden Satz zusammengefaßt:

Satz 10.23 *Sei $\mathbf{V}$ ein Vektorraum, $\sigma \in BL(\mathbf{V})$, $\mathbf{M}, \mathbf{N}$ nichtleere Teilmengen von $\mathbf{V}$ und $\mathbf{U}_i$ $(i \in I)$ ein System von Unterräumen von $\mathbf{V}$.*

1. $\mathbf{M} \subseteq \mathbf{N} \Longrightarrow \mathbf{N}^{\perp} \subseteq \mathbf{M}^{\perp} \;\wedge\; {}^{\perp}\mathbf{N} \subseteq {}^{\perp}\mathbf{M}$.
2. $[\mathbf{M}]^{\perp} = \mathbf{M}^{\perp}$, ${}^{\perp}[\mathbf{M}] = {}^{\perp}\mathbf{M}$.
3. $\mathbf{M}^{\perp} \leq \mathbf{V}$, ${}^{\perp}\mathbf{M} \leq \mathbf{V}$.
4. ${}^{\perp}\mathbf{M} = g_{\sigma}^{-1}(\mathbf{M}^{\circ})$, $\mathbf{M}^{\perp} = f_{\sigma}^{-1}(\mathbf{M}^{\circ})$.
5. $\mathbf{V}^{\perp} = \mathrm{kern}(f_{\sigma})$, ${}^{\perp}\mathbf{V} = \mathrm{kern}(g_{\sigma})$.
6. $\mathbf{M} \subseteq {}^{\perp}(\mathbf{M}^{\perp})$, $\mathbf{M} \subseteq ({}^{\perp}\mathbf{M})^{\perp}$.
7. $\left(\sum\limits_{i\in I} \mathbf{U}_i\right)^{\perp} = \bigcap\limits_{i\in I} \mathbf{U}_i^{\perp}$, ${}^{\perp}\left(\sum_{i\in I} \mathbf{U}_i\right) = \bigcap_{i\in I} {}^{\perp}\mathbf{U}_i$.
8. $\sum\limits_{i\in I} \mathbf{U}_i^{\perp} \subseteq \left(\bigcap\limits_{i\in I} \mathbf{U}_i\right)^{\perp}$, $\sum_{i\in I} {}^{\perp}\mathbf{U}_i \subseteq {}^{\perp}\left(\bigcap_{i\in I} \mathbf{U}_i\right)$,
   wobei Gleichheit gilt, wenn $\dim(\mathbf{V})$ endlich ist und σ nichtausgeartet ist.

Ist $\mathbf{U}$ ein Unterraum eines endlichdimensionalen Vektorraums $\mathbf{V}$, so können genaue Aussagen über die Dimensionen der auftretenden Vektorräume gemacht werden.

Satz 10.24 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum, $\sigma \in BL(\mathbf{V})$ und $\mathbf{U} \leq \mathbf{V}$. Dann gilt:*

1. $\dim(\mathbf{U}) + \dim(\mathbf{U}^{\perp}) = \dim(\mathbf{V}) + \dim(\mathbf{U} \cap {}^{\perp}\mathbf{V})$.
2. $\dim(\mathbf{U}) + \dim({}^{\perp}\mathbf{U}) = \dim(\mathbf{V}) + \dim(\mathbf{U} \cap \mathbf{V}^{\perp})$.

Korollar 10.25 *Ist $\mathbf{V}$ ein endlichdimensionaler Vektorraum und $\sigma \in BL(\mathbf{V})$ nichtausgeartet, so gilt für alle Unterräume $\mathbf{U} \leq \mathbf{V}$*

$$\mathbf{U} = {}^{\perp}(\mathbf{U}^{\perp}) = ({}^{\perp}\mathbf{U})^{\perp}.$$

Von besonderer Bedeutung für die Struktursätze alternierender und symmetrischer Bilinearformen ist die folgende Eigenschaft:

Satz 10.26 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum und* $\sigma \in BL(\mathbf{V})$. *Dann gilt für jeden nichtausgearteten Unterraum* $\mathbf{U} \leq \mathbf{V}$

$$\mathbf{U} \oplus \mathbf{U}^{\perp} = \mathbf{U} \oplus {}^{\perp}\mathbf{U} = \mathbf{V}.$$

10.3 Stuktursätze von Bilinearformen

10.3.1 Alternierende Bilinearformen

Satz 10.27 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum und* $\sigma \in BL(\mathbf{V})$ *eine alternierende Bilinearform vom Rang* r. *Dann ist* r *gerade, und es gibt* $s = \frac{r}{2}$ *paarweise orthogonale, nichtausgeartete, zweidimensionale Unterräume* $\mathbf{L}_1, \ldots, \mathbf{L}_s$ *von* $\mathbf{V}$ *mit*

$$\mathbf{V} = \mathbf{L}_1 \oplus \cdots \oplus \mathbf{L}_s \oplus \mathbf{V}^{\perp}.$$

Es gibt daher eine Basis $\mathbf{B}$ von $\mathbf{V}$, so daß die Koordinatenmatrix $\Phi_{\mathbf{B}}(\sigma)$ die Gestalt

$$\Phi_{\mathbf{B}}(\sigma) = \begin{pmatrix} 0 & 1 & & & & & & & & & \\ -1 & 0 & & & & & & & & & \\ & & 0 & 1 & & & & & & & \\ & & -1 & 0 & & & & & & & \\ & & & & \ddots & & & & & & \\ & & & & & 0 & 1 & & & & \\ & & & & & -1 & 0 & & & & \\ & & & & & & & 0 & & & \\ & & & & & & & & 0 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 0 \end{pmatrix} \tag{10.1}$$

hat, wobei die restlichen Elemente durch 0 gefüllt werden müssen. Man sagt auch, $\Phi_{\mathbf{B}}(\sigma)$ hat erweiterte Diagonalgestalt und schreibt dafür

$$\Phi_{\mathbf{B}}(\sigma) = \operatorname{diag}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, 0, 0, \ldots, 0\right).$$

Satz 10.27 löst auch das Kongruenzproblem von alternierenden Bilinearformen.

Satz 10.28 *Zwei alternierende Bilinearformen* $\sigma, \tau \in BL(\mathbf{V})$ *auf einem endlichdimensionalen Vektorraum sind genau dann kongruent wenn ihre Ränge gleich sind.*

Entsprechend gilt, daß eine alternierende Matrix immer zu einer Matrix der Form (10.1) kongruent ist, und daraus folgt unmittelbar:

Satz 10.29 *Zwei alternierende Matrizen* $A, B \in K^{n \times n}$ *sind genau dann kongruent, wenn ihre Ränge gleich sind.*

10.3.2 Symmetrische Bilinearformen

Satz 10.30 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum und* $\sigma \in BL(\mathbf{V})$ *eine symmetrische aber nicht alternierende Bilinearform vom Rang* r*. Dann gibt es* r *paarweise orthogonale, nichtausgeartete, eindimensionale Unterräume* $\mathbf{A}_1, \ldots, \mathbf{A}_r$ *von* $\mathbf{V}$ *mit*

$$\mathbf{V} = \mathbf{A}_1 \oplus \cdots \oplus \mathbf{A}_r \oplus \mathbf{V}^\perp.$$

Es gibt daher eine Basis $\mathbf{B}$ von $\mathbf{V}$, so daß die Koordinatenmatrix $\Phi_\mathbf{B}(\sigma)$ Diagonalgestalt

$$\Phi_\mathbf{B}(\sigma) = \mathrm{diag}(g_1, g_2, \ldots, g_r, 0, 0, \ldots, 0)$$

mit $g_1, g_2, \ldots, g_r \in K^\times$ hat. Entsprechend ist eine symmetrische Matrix immer zu einer Diagonalmatrix kongruent.

Damit ist das allgemeine Kongruenzproblem für symmetrische Bilinearformen bzw. Matrizen **nicht gelöst**, da die Diagonalelemente $g_1, \ldots, g_r$ (außer in Spezialfällen) nicht genau spezifiziert werden können. Allgemein kann nur gesagt werden, daß es auf quadratische Faktoren nicht ankommt. (Ersetzt man nämlich den Basisvektor $\mathbf{b}_j$ durch $x\mathbf{b}_j$, so wird das Diagonalelement $g_j = \sigma(\mathbf{b}_j, \mathbf{b}_j)$ durch $x^2 g_j$ ersetzt.)

10.3.3 Symmetrische Bilinearformen über $\mathbb{C}$

Satz 10.31 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über den komplexen Zahlen* $\mathbb{C}$ *und* $\sigma \in BL(\mathbf{V})$ *eine symmetrische Bilinearform vom Rang* r*. Dann gibt es eine Basis* $\mathbf{B}$ *von* $\mathbf{V}$ *mit*

$$\Phi_\mathbf{B}(\sigma) = \begin{pmatrix} E_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \mathrm{diag}(1, 1, \ldots, 1, 0, 0, \ldots, 0).$$

Dieser Satz löst das Kongruenzproblem für symmetrische Bilinearformen über $\mathbb{C}$.

Satz 10.32 *Zwei symmetrische Bilinearformen* $\sigma, \tau \in BL(\mathbf{V})$ *auf einem endichdimensionalen Vektorraum* $\mathbf{V}$ *über* $\mathbb{C}$ *sind genau dann kongruent, wenn ihre Ränge gleich sind.*

Entsprechend gilt, daß jede Matrix $A \in \mathbb{C}^{n\times n}$ zu einer Matrix der Form $\mathrm{diag}(1, 1, \ldots, 1, 0, 0, \ldots, 0)$ kongruent ist, und daraus folgt:

Satz 10.33 *Zwei symmetrische Matrizen* $A, B \in \mathbb{C}^{n\times n}$ *sind genau dann kongruent, wenn ihre Ränge gleich sind.*

10.3.4 Trägheitssatz von Sylvester

Die Situation bei reellen endlichdimensionalen Vektorräumen ist nicht ganz so einfach wie im komplexen Fall. Trotzdem gelingt es hier, das Kongruenzproblem zu lösen. Dazu führt man folgende Begriffe ein.

Definition 10.34 *Sei* $\mathbf{V}$ *ein Vektorraum über* $\mathbb{R}$*,* $\sigma \in BL(\mathbf{V})$ *eine symmetrische Bilinearform auf* $\mathbf{V}$ *und* $\mathbf{U} \leq \mathbf{V}$ *ein Unterraum von* $\mathbf{V}$*.*

1. $\mathbf{U}$ *heißt* **positiv definit***, wenn für alle* $\mathbf{a} \in \mathbf{U} \setminus \{\mathbf{0}\}$ $\sigma(\mathbf{a},\mathbf{a}) > 0$ *ist.*
2. $\mathbf{U}$ *heißt* **positiv semidefinit***, wenn für alle* $\mathbf{a} \in \mathbf{U}$ $\sigma(\mathbf{a},\mathbf{a}) \geq 0$ *ist.*
3. $\mathbf{U}$ *heißt* **negativ definit***, wenn für alle* $\mathbf{a} \in \mathbf{U} \setminus \{\mathbf{0}\}$ $\sigma(\mathbf{a},\mathbf{a}) < 0$ *ist.*
4. $\mathbf{U}$ *heißt* **negativ semidefinit***, wenn für alle* $\mathbf{a} \in \mathbf{U}$ $\sigma(\mathbf{a},\mathbf{a}) \leq 0$ *ist.*
5. $\mathbf{U}$ *heißt* **indefinit***, wenn es* $\mathbf{a},\mathbf{b} \in \mathbf{U}$ *mit* $\sigma(\mathbf{a},\mathbf{a}) > 0$ *und* $\sigma(\mathbf{b},\mathbf{b}) < 0$ *gibt.*

Definition 10.35 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über* $\mathbb{R}$ *und* $\sigma \in BL(\mathbf{V})$ *eine symmetrische Bilinearform. Definiert man*

$$p := \max\{\dim(\mathbf{U}) \,|\, \mathbf{U} \leq \mathbf{V} \ \wedge \ \mathbf{U} \text{ ist positiv definit}\}$$

und

$$q := \max\{\dim(\mathbf{U}) \,|\, \mathbf{U} \leq \mathbf{V} \ \wedge \ \mathbf{U} \text{ ist negativ definit}\},$$

dann heißt die Differenz

$$p - q$$

Signatur *von* σ*.*

Entsprechend kann man die Signatur einer symmetrischen Matrix $A \in \mathbb{R}^{n\times n}$ definieren.

Satz 10.36 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über* $\mathbb{R}$ *und* $\sigma \in BL(\mathbf{V})$ *eine symmetrische Bilinearform. Dann gibt es eine Basis* $\mathbf{B}$ *von* $\mathbf{V}$ *mit*

$$\Phi_{\mathbf{B}}(\sigma) = \begin{pmatrix} E_p & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -E_q & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{10.2}$$

wobei $p+q$ *der Rang von* σ *und* $p-q$ *die Signatur von* σ *ist.*

Der **Trägheitssatz von Sylvester** löst nun das Kongruenzproblem für reelle symmetrische Bilinearformen.

Satz 10.37 *Zwei symmetrische Bilinearformen* $\sigma,\tau \in BL(\mathbf{V})$ *auf einem endlichdimensionalen Vektorraum* $\mathbf{V}$ *über den reellen Zahlen sind genau dann kongruent, wenn ihre Ränge und Signaturen gleich sind.*

Entsprechend gilt, daß jede symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ zu einer Matrix der Form (10.2) kongruent ist, wobei p und q eindeutig bestimmt sind. Es gilt auch:

Satz 10.38 *Zwei symmetrische Matrizen* $A, B \in \mathbb{R}^{n\times n}$ *sind genau dann kongruent, wenn ihre Ränge und ihre Signaturen gleich sind.*

10.4 Quadratische Formen

10.4.1 Bilinearformen und Quadratische Formen

Betrachtet man zu einer (symmetrischen) Matrix $A \in K^{n\times n}$ die Abbildung $q(\mathbf{a}) = \mathbf{a}^T A\mathbf{a}$ (bzw. setzt man in einer Bilinearform $\sigma \in BL(\mathbf{V})$ in beide Stellen denselben Vektor ein: $q(\mathbf{a}) = \sigma(\mathbf{a},\mathbf{a})$), so erhält man eine Abbildung $q : \mathbf{V} \to K$ mit der Eigenschaft $q(x\mathbf{a}) = x^2 q(\mathbf{a})$. Genauer definiert man:

Definition 10.39 *Sei* $\mathbf{V}$ *ein Vektorraum über dem Körper* K*. Eine Abbildung* $q : \mathbf{V} \to K$ *mit den Eigenschaften*

$$q(x\mathbf{a}) = x^2 q(\mathbf{a}) \qquad (x \in K, \mathbf{a} \in \mathbf{V})$$

und

$$\sigma(\mathbf{a},\mathbf{b}) := q(\mathbf{a}+\mathbf{b}) - q(\mathbf{a}) - q(\mathbf{b}) \ \text{ ist eine Bilinearform}$$

heißt **quadratische Form** *auf* $\mathbf{V}$.

Man beachte, daß die einer quadratischen Form q zugeordnete Bilinearform σ immer symmetrisch ist. Umgekehrt kann man wegen

$$\sigma(\mathbf{a},\mathbf{a}) = 2^2 q(\mathbf{a}) - 2q(\mathbf{a}) = 2q(\mathbf{a})$$

im Fall $\mathrm{char}(K) \neq 2$ sofort die quadratische Form zurückgewinnen.

Satz 10.40 *Sei* $\mathbf{V}$ *ein Vektorraum über einem Körper* K *mit* $\mathrm{char}(K) \neq 2$*. Dann entsprechen die quadratischen Formen und die symmetrischen Bilinearformen einander eineindeutig.*

Ist $\mathbf{V}$ überdies endlichdimensional (über einem Körper K mit $\mathrm{char}(K) \neq 2$), dann kann man q auch mit Hilfe einer Matrix beschreiben. Dazu lege man eine Basis $\mathbf{B}$ von $\mathbf{V}$ fest und setze

$$G := \frac{1}{2}\Phi_{\mathbf{B}}(\sigma).$$

Dann gilt

$$q(\mathbf{a}) = \Phi_{\mathbf{B}}(\mathbf{a})^T G\, \Phi_{\mathbf{B}}(\mathbf{a}).$$

10.4.2 Definite quadratische Formen

Definition 10.41 *Sei* $\mathbf{V}$ *ein Vektorraum über den reellen Zahlen* $\mathbb{R}$ *und* q *eine quadratische Form auf* $\mathbf{V}$.

1. q *heißt* **positiv definit**, *wenn* $q(\mathbf{a}) > 0$ *für alle* $\mathbf{a} \in \mathbf{V} \setminus \{\mathbf{0}\}$.
2. q *heißt* **positiv semidefinit**, *wenn* $q(\mathbf{a}) \geq 0$ *für alle* $\mathbf{a} \in \mathbf{V}$.
3. q *heißt* **negativ definit**, *wenn* $q(\mathbf{a}) < 0$ *für alle* $\mathbf{a} \in \mathbf{V} \setminus \{\mathbf{0}\}$.
4. q *heißt* **negativ semidefinit**, *wenn* $q(\mathbf{a}) \leq 0$ *für alle* $\mathbf{a} \in \mathbf{V}$.

5. *q heißt* **indefinit**, *wenn es Vektoren $\mathbf{a}, \mathbf{b} \in \mathbf{V}$ mit $q(\mathbf{a}) > 0$ und $q(\mathbf{b}) < 0$ gibt.*

Entsprechend definiert man für symmetrische Matrizen:

Definition 10.42 *Sei $G \in \mathbb{R}^{n \times n}$ eine reelle symmetrische Matrix.*

1. *G heißt* **positiv definit**, *wenn $\mathbf{a}^T G \mathbf{a} > 0$ für alle $\mathbf{a} \in \mathbb{R}^{n \times 1} \setminus \{\mathbf{0}\}$.*
2. *G heißt* **positiv semidefinit**, *wenn $\mathbf{a}^T G \mathbf{a} \geq 0$ für alle $\mathbf{a} \in \mathbb{R}^{n \times 1}$.*
3. *G heißt* **negativ definit**, *wenn $\mathbf{a}^T G \mathbf{a} < 0$ für alle $\mathbf{a} \in \mathbb{R}^{n \times 1} \setminus \{\mathbf{0}\}$.*
4. *G heißt* **negative semidefinit**, *wenn $\mathbf{a}^T G \mathbf{a} \leq 0$ für alle $\mathbf{a} \in \mathbb{R}^{n \times 1}$.*
5. *G heißt* **indefinit**, *wenn es Vektoren $\mathbf{a}, \mathbf{b} \in \mathbb{R}^{n \times 1}$ mit $\mathbf{a}^T G \mathbf{a} > 0$ und $\mathbf{b}^T G \mathbf{b} < 0$ gibt.*

Sei $G = (g_{ij})_{1 \leq i,j \leq n} \in \mathbb{R}^{n \times n}$ eine Matrix. Die Determinanten

$$\det G_{(1,2,\ldots k)}$$

der $k \times k$-Untermatrizen $G_{(1,2,\ldots k)} := (g_{ij})_{1 \leq i,j \leq k}$, die aus den ersten k Zeilen resp. Spalten von G gebildet werden, heißen **Hauptminoren** der Matrix G. Mit Hilfe der Hauptminoren kann entschieden werden, ob eine quadratische Form positiv bzw. negativ definit ist (**Hauptminorenkriterium**).

Satz 10.43 *Sei q eine quadratische Form auf einem endlichdimensionalen Vektorraum $\mathbf{V}$ über $\mathbb{R}$ und $\sigma \in BL(\mathbf{V})$ jene symmetrische Bilinearform auf $\mathbf{V}$ mit $q(\mathbf{a}) = \sigma(\mathbf{a}, \mathbf{a})$.*

Dann ist q genau dann positiv definit, wenn alle Hauptminoren der Matrix $G = \Phi_{\mathbf{B}}(\sigma)$ (für eine beliebige Basis $\mathbf{B}$ von $\mathbf{V}$) positiv sind:

$$\det G_{(1,2,\ldots,k)} > 0 \qquad (1 \leq k \leq n).$$

q ist genau dann negativ definit, wenn die Hauptminoren abwechselnd negativ und positiv sind:

$$\det G_{(1,2,\ldots,2k-1)} < 0 \quad \textit{und} \quad \det G_{(1,2,\ldots,2k)} > 0 \quad (1 \leq k \leq n/2).$$

Satz 10.44 *Sei $G \in \mathbb{R}^{n \times n}$ eine reelle symmetrische Matrix.*

Dann ist G genau dann positiv definit, wenn alle Hauptminoren der Matrix G positiv sind:

$$\det G_{(1,2,\ldots,k)} > 0 \qquad (1 \leq k \leq n).$$

G ist genau dann negativ definit, wenn die Hauptminoren abwechselnd negativ und positiv sind:

$$\det G_{(1,2,\ldots,2k-1)} < 0 \quad \textit{und} \quad \det G_{(1,2,\ldots,2k)} > 0 \quad (1 \leq k \leq n/2).$$

Man beachte, daß man positive bzw. negativ semidefinite quadratische Formen (bzw. Matrizen) nicht mit einem Satz dieser Form charakterisieren kann.

10.5 Sesquilinearformen und Hermitesche Formen

10.5.1 Sesquilinearformen

Definition 10.45 *Sei* $\mathbf{V}$ *ein Vektorraum über dem Körper* K *und* $\zeta : K \to K$ *ein Körperautomorphismus. Eine Abbildung* $\sigma : \mathbf{V} \times \mathbf{V} \to K$ *heißt* ζ-**Sesquilinearform**, *wenn* σ *in der ersten Komponente linear und in der zweiten Komponente* ζ*-semilinear ist, d.h. für* $\mathbf{a}, \mathbf{a}_1, \mathbf{a}_2, \mathbf{b}, \mathbf{b}_1, \mathbf{b}_2 \in \mathbf{V}$ *und* $x, y \in K$ *gilt*

$$\begin{aligned}\sigma(x\mathbf{a}_1 + y\mathbf{a}_2, \mathbf{b}) &= x\sigma(\mathbf{a}_1, \mathbf{b}) + y\sigma(\mathbf{a}_2, \mathbf{b}),\\ \sigma(\mathbf{a}, x\mathbf{b}_1 + y\mathbf{b}_2) &= \zeta(x)\sigma(\mathbf{a}, \mathbf{b}_1) + \zeta(y)\sigma(\mathbf{a}, \mathbf{b}_2).\end{aligned}$$

Im Spezialfall $\zeta = \mathrm{id}_K$ ist eine ζ-Sesquilinearform eine Bilinearform. Die Menge aller ζ-Sesquilinearformen bildet natürlich wieder einen Vektorraum.

Ist $A \in K^{n\times n}$ eine Matrix, so ist z.B. die Abbildung

$$\sigma(\mathbf{a}, \mathbf{b}) := \mathbf{a}^T A\zeta(\mathbf{b})$$

eine ζ-Sesquilinearform auf $\mathbf{V} = K^{n\times n}$ (wobei $\zeta(\mathbf{b})$ natürlich komponentenweise zu verstehen ist).

Definition 10.46 *Eine* ζ*-Sesquilinearform* σ *heißt* ζ-**symmetrisch**, *wenn für alle* $\mathbf{a}, \mathbf{b} \in \mathbf{V}$

$$\sigma(\mathbf{a}, \mathbf{b}) = \zeta(\sigma(\mathbf{b}, \mathbf{a}))$$

gilt.

Eine ζ*-Sesquilinearform* σ *heißt* ζ-**schiefsymmetrisch**, *wenn für alle* $\mathbf{a}, \mathbf{b} \in \mathbf{V}$

$$\sigma(\mathbf{a}, \mathbf{b}) = -\zeta(\sigma(\mathbf{b}, \mathbf{a}))$$

gilt.

Eine ζ*-Sesquilinearform* σ *heißt* **alternierend**, *wenn für alle* $\mathbf{a} \in \mathbf{V}$

$$\sigma(\mathbf{a}, \mathbf{a}) = 0$$

gilt.

Lemma 10.47 *Eine alternierende Sesquilinearform ist eine Bilinearform. (Insbesondere ist* $\zeta = \mathrm{id}_K$*.)*

Der praktisch wichtigste Fall einer Sesquilinearform ist eine Hermitesche Form.

Definition 10.48 *Sei* $\mathbf{V}$ *ein Vektorraum über den komplexen Zahlen* $\mathbb{C}$ *und bezeichne* $\zeta : \mathbb{C} \to \mathbb{C}$ *die komplexe Konjugation* $\zeta(z) = \overline{z}$*. Eine (in diesem Sinn)* ζ*-symmetrische* ζ*-Sesquilinearform* σ *heißt* **Hermitesche Form** *oder* **Hermitesche Sesquilinearform***:*

$$\begin{aligned}\sigma(x\mathbf{a}_1 + y\mathbf{a}_2, \mathbf{b}) &= x\sigma(\mathbf{a}_1, \mathbf{b}) + y\sigma(\mathbf{a}_2, \mathbf{b}),\\ \sigma(\mathbf{a}, x\mathbf{b}_1 + y\mathbf{b}_2) &= \overline{x}\sigma(\mathbf{a}, \mathbf{b}_1) + \overline{y}\sigma(\mathbf{a}, \mathbf{b}_2),\\ \sigma(\mathbf{a}, \mathbf{b}) &= \overline{\sigma(\mathbf{b}, \mathbf{a})}.\end{aligned}$$

Ist $A \in \mathbb{C}^{n\times n}$ eine **Hermitesche Matrix**, d.h. $A^T = \overline{A}$, dann ist die Abbildung

$$\sigma(\mathbf{a}, \mathbf{b}) := \mathbf{a}^T A \overline{\mathbf{b}}$$

eine Hermitesche Form, denn

$$\overline{\sigma(\mathbf{b}, \mathbf{a})} = \overline{\mathbf{b}^T A \overline{\mathbf{a}}} = \overline{\mathbf{b}}^T \overline{A} \mathbf{a} = \overline{\mathbf{b}}^T A^T \mathbf{a} = \mathbf{a}^T A \overline{\mathbf{b}} = \sigma(\mathbf{a}, \mathbf{b}).$$

Ähnlich wie bei einer Bilinearform kann man einer ζ-Sesquilinearform σ eine ζ-semilineare Funktion $f_\sigma : \mathbf{V} \to \mathbf{V}^*$ durch

$$\sigma(\mathbf{a}, \mathbf{b}) = \langle f_\sigma(\mathbf{b}), \mathbf{a}\rangle$$

zuordnen.

Satz 10.49 *Der Vektorraum aller ζ-Sesquilinearformen auf $\mathbf{V}$ ist zum Vektorraum aller ζ-semilinearen Abbildungen $\mathbf{V} \to \mathbf{V}^*$ isomorph.*

Genauso wie bei den Bilinearformen kann man Sesquilinearformen bezüglich einer Basis eine Matrix zuordnen.

Definition 10.50 *Sei $\mathbf{V}$ ein n-dimensionaler Vektorraum über dem Körper K, $\zeta : K \to K$ ein Körperautomorphismus und $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ eine Basis von $\mathbf{V}$. Die* **Koordinatenmatrix** *$\Phi_{\mathbf{B}}(\sigma)$ einer ζ-Sesquilinearform σ auf $\mathbf{V}$ ist durch*

$$\Phi_{\mathbf{B}}(\sigma) := (\sigma(\mathbf{b}_i, \mathbf{b}_j))_{1\le i,j\le n}$$

gegeben.

Lemma 10.51 *Sei $\mathbf{V}$ ein n-dimensionaler Vektorraum und $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ eine Basis von $\mathbf{V}$. Dann gilt für eine ζ-Sesquilinearform σ auf $\mathbf{V}$ und Vektoren $\mathbf{a}, \mathbf{b} \in \mathbf{V}$*

$$\sigma(\mathbf{a}, \mathbf{b}) = \Phi_{\mathbf{B}}(\mathbf{a})^T \Phi_{\mathbf{B}}(\sigma) \zeta(\Phi_{\mathbf{B}}(\mathbf{b})).$$

D.h., man kann eine Sesquilinearform auf der Ebene der Koordinaten immer durch eine Matrix beschreiben.

Satz 10.52 *Sei $\mathbf{V}$ ein n-dimensionaler Vektorraum, $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ eine Basis von $\mathbf{V}$ und $\mathbf{B}^* \subseteq \mathbf{V}^*$ die duale Basis von $\mathbf{B}$. Dann gilt für eine ζ-Sesquilinearform σ auf $\mathbf{V}$ und der entsprechenden ζ-semilinearen Abbildung $f_\sigma : \mathbf{V} \to \mathbf{V}^*$*

$$\Phi_{\mathbf{B}}(\sigma) = \zeta(\Phi_{\mathbf{B}\mathbf{B}^*}(f_\sigma)).$$

Definition 10.53 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum über dem Körper K, $\zeta : K \to K$ ein Körperautomorphismus und σ eine ζ-Sesquilinearform auf $\mathbf{V}$. Dann werden Rang* $\mathrm{rg}(\sigma)$ *und Defekt* $\mathrm{def}(\sigma)$ *von σ durch*

$$\begin{aligned} \mathrm{rg}(\sigma) &:= \mathrm{rg}(f_\sigma), \\ \mathrm{def}(\sigma) &:= \mathrm{def}(f_\sigma) \end{aligned}$$

definiert.

Man beachte, daß dann für jede Basis $\mathbf{B}$ von $\mathbf{V}$

$$\mathrm{rg}(\sigma) = \mathrm{rg}(f_\sigma) = \mathrm{rg}(\Phi_{\mathbf{BB}^*}(f_\sigma)) = \mathrm{rg}(\Phi_{\mathbf{B}}(\sigma))$$

gilt.

Abschließend beschäftigen wir uns noch kurz mit ζ-symmetrischen bzw. ζ-schiefsymmetrischen Matrizen und der Zusammenhang zu ζ-symmetrischen bzw. ζ-schiefsymmetrischen Sesquilinearformen.

Definition 10.54 *Sei $\zeta : K \to K$ ein Körperautomorphismus und $A = (a_{ij})_{1\leq i,j\leq n} \in K^{n\times n}$ eine quadratische Matrix.*

1. *A heißt ζ-**symmetrisch**, wenn $A^T = \zeta(A)$.*
2. *A heißt ζ-**schiefsymmetrisch**, wenn $A^T = -\zeta(A)$.*

Durch direktes Nachrechnen erhält man, daß die entsprechenden Abbildungen

$$\sigma(\mathbf{a}, \mathbf{b}) := \mathbf{a}^T A \zeta(\mathbf{b})$$

ζ-symmetrisch bzw. ζ-schiefsymmetrisch sind.

Die Eigenschaften ζ-symmetrische und ζ-schiefsymmetrisch spiegeln sich auch bei den Koordinatenmatrizen wider.

Satz 10.55 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum über dem Körper K, $\zeta : K \to K$ ein Körperautomorphismus und $\mathbf{B}$ eine Basis von $\mathbf{V}$. Dann gelten für eine ζ-Sesquilinearform σ auf $\mathbf{V}$ folgende Eigenschaften:*

1. *σ ζ-symmetrisch $\iff \Phi_{\mathbf{B}}(\sigma)$ ζ-symmetrisch.*
2. *σ ζ-schiefsymmetrisch $\iff \Phi_{\mathbf{B}}(\sigma)$ ζ-schiefsymmetrisch.*

Insbesondere erfüllen die Koordinatenmatrizen $\Phi_{\mathbf{B}}(\sigma)$ von Hermiteschen Formen

$$\Phi_{\mathbf{B}}(\sigma)^T = \overline{\Phi_{\mathbf{B}}(\sigma)}$$

bzw.

$$\Phi_{\mathbf{B}}(\sigma)^H = \Phi_{\mathbf{B}}(\sigma),$$

wenn man für

$$A^H := (\overline{A})^T$$

setzt. A^H heißt auch die **Hermiteschtransponierte Matrix**.

Eine Matrix $A \in \mathbb{C}^{n\times n}$, die $A^H = (\overline{A})^T = A$ erfüllt, heißt auch Hermitesch. Hermitesche Formen entsprechen also Hermiteschen Matrizen.

10.5.2 ζ-kongruente Matrizen

Wieder kann ganz analog zu Bilinearformen vorgegangen werden:

Lemma 10.56 *Sei* $\mathbf{V}$ *ein Vektorraum über dem Körper* K*,* $\zeta : K \to K$ *ein Körperautomorphismus und* σ *eine* ζ*-Sesquilinearform auf* $\mathbf{V}$*. Dann ist für jede lineare Abbildung* $f \in L(\mathbf{V}, \mathbf{V})$ *die Abbildung*

$$\tau = \sigma \circ (f, f) : \mathbf{V} \times \mathbf{V} \to K, \quad (\mathbf{a}, \mathbf{b}) \mapsto \sigma(f(\mathbf{a}), f(\mathbf{b}))$$

wieder eine ζ*-Sesquilinearform auf* $\mathbf{V}$*.*

Definition 10.57 *Sei* $\mathbf{V}$ *ein Vektorraum über dem Körper* K *und* $\zeta : K \to K$ *ein Körperautomorphismus. Zwei* ζ*-Sesquilinearformen* σ, τ *auf* $\mathbf{V}$ *heißen* **kongruent***, wenn es eine bijektive lineare Abbildung* $f \in L(\mathbf{V}, \mathbf{V})$ *mit*

$$\tau(\mathbf{a}, \mathbf{b}) = \sigma(f(\mathbf{a}), f(\mathbf{b})) \quad \textit{resp.} \quad \tau = \sigma \circ (f, f)$$

gibt.

Offensichtlich ist die Relation σ *ist kongruent zu* τ eine Äquivalenzrelation auf den ζ-Sesquilinearformen.

Satz 10.58 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über dem Körper* K*,* $\zeta : K \to K$ *ein Körperautomorphismus,* $\mathbf{B}$ *eine Basis von* $\mathbf{V}$*,* σ *eine* ζ*-Sesquilinearform auf* $\mathbf{V}$ *und* $f \in L(\mathbf{V}, \mathbf{V})$*. Dann gilt für* $\tau = \sigma \circ (f, f) \in BL(\mathbf{V})$

$$\Phi_{\mathbf{B}}(\tau) = \Phi_{\mathbf{BB}}(f)^T \Phi_{\mathbf{B}}(\sigma)\, \zeta(\Phi_{\mathbf{BB}}(f)).$$

Satz 10.59 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über dem Körper* K*,* $\zeta : K \to K$ *ein Körperautomorphismus und* $\mathbf{B}, \tilde{\mathbf{B}}$ *zwei Basen von* $\mathbf{V}$*. Dann gilt für jede* ζ*-Sesquilinearform* σ *auf* $\mathbf{V}$

$$\Phi_{\tilde{\mathbf{B}}}(\sigma) = T_{\tilde{\mathbf{B}}\mathbf{B}}^T \Phi_{\mathbf{B}}(\sigma)\, \zeta(T_{\tilde{\mathbf{B}}\mathbf{B}}).$$

Diese beiden Sätze motivieren die folgende Definition.

Definition 10.60 *Sei* $\zeta : K \to K$ *ein Körperautomorphismus. Zwei quadratische Matrizen* $A, B \in K^{n \times n}$ *heißen* ζ-**kongruent***, wenn es eine reguläre Matrix* $P \in K^{n \times n}$ *mit*

$$B = P^T A\, \zeta(P)$$

gibt.

Offensichtlich gelten die folgenden Eigenschaften.

Satz 10.61 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über dem Körper* K*,* $\zeta : K \to K$ *ein Körperautomorphismus und* $\mathbf{B}$ *eine Basis von* $\mathbf{V}$*. Zwei* ζ*-Sesquilinearformen* σ, τ *auf* $\mathbf{V}$ *sind genau dann kongruent, wenn die Koordinatenmatrizen* $\Phi_{\mathbf{B}}(\sigma), \Phi_{\mathbf{B}}(\tau)$ ζ*-kongruent sind.*

Weiters sind die Koordinatenmatrizen $\Phi_{\mathbf{B}}(\sigma), \Phi_{\tilde{\mathbf{B}}}(\sigma)$ von zwei Basen $\mathbf{B}, \tilde{\mathbf{B}}$ von $\mathbf{V}$ ζ-kongruent.

10.5.3 Orthosymmetrische Sesquilinearformen

Mit Hilfe von Sesquilinearformen wird wie bei Bilinearformen ein Orthogonalitätsbegriff eingeführt.

Definition 10.62 *Sei $\mathbf{V}$ ein Vektorraum über dem Körper K, $\zeta : K \to K$ ein Körperautomorphismus und σ eine Sesquilinearform auf $\mathbf{V}$. Zwei Vektoren $\mathbf{a}, \mathbf{b} \in \mathbf{V}$ heißen* **orthogonal** *bezüglich σ, wenn $\sigma(\mathbf{a}, \mathbf{b}) = 0$. Man schreibt dafür auch*

$$\mathbf{a} \perp_\sigma \mathbf{b}.$$

Ist die Relation $\perp_\sigma$ symmetrisch, d.h.

$$\sigma(\mathbf{a}, \mathbf{b}) = 0 \Longleftrightarrow \sigma(\mathbf{b}, \mathbf{a}) = 0,$$

dann heißt σ **orthosymmetrisch**.

Die Strukturtheorie ist bei allgemeinen Sesquilinearformen ein wenig komplizierter als bei Bilinearformen.

Satz 10.63 *Sei $\mathbf{V}$ ein Vektorraum über dem Körper K und $\zeta : K \to K$ ein Körperautomorphismus. Stimmen für zwei ζ-Sesquilinearformen σ, τ auf $\mathbf{V}$ die Orthogonalitätsrelationen $\perp_\sigma, \perp_\tau$ überein, d.h.*

$$\sigma(\mathbf{a}, \mathbf{b}) = 0 \Longleftrightarrow \tau(\mathbf{a}, \mathbf{b}) = 0,$$

wobei $\mathrm{rg}(\sigma) \geq 2$ ist, dann gibt es $c \in K^\times$ mit

$$\sigma = c\tau.$$

Daraus ergibt sich die folgende Charakterisierung orthosymmetrischer Sesquilinearformen.

Satz 10.64 *Sei $\mathbf{V}$ ein Vektorraum über dem Körper K und $\zeta : K \to K$ ein Körperautomorphismus. Dann ist eine ζ-Sesquilinearform σ auf $\mathbf{V}$ genau dann orthosymmetrisch, wenn sie alternierend ist (d.h. σ ist eine alternierende Bilinearform und $\zeta = \mathrm{id}_K$) oder wenn $\zeta^2 = \mathrm{id}_K$ ist und es eine Konstante $c \in K^\times$ gibt, so daß $c\,\sigma$ ζ-symmetrisch ist.*

Ebenso kann man auch bezüglich Sesquilinearformen Orthogonalräume betrachten. Einfachheitshalber werden in diesem Zusammenhang nur orthosymmetrische Sesquilinearformen betrachtet.

Definition 10.65 *Sei $\mathbf{V}$ ein Vektorraum über dem Körper K, $\zeta : K \to K$ ein Körperautomorphismus und σ eine orthosymmetrische ζ-Sesquilinearform auf $\mathbf{V}$. Zu einer nichtleeren Teilmenge $\mathbf{M} \subseteq \mathbf{V}$ bezeichnet*

$$\mathbf{M}^\perp := \{\mathbf{b} \in \mathbf{V} \mid \forall \mathbf{a} \in \mathbf{M} : \ \sigma(\mathbf{a}, \mathbf{b}) = 0\}$$

den **Orthogonalraum** *von $\mathbf{M}$. Die Menge $\mathbf{V}^\perp$ heißt* **Radikal** *von $\mathbf{V}$.*

Definition 10.66 *Sei* $\mathbf{V}$ *ein Vektorraum über dem Körper* K *und* $\zeta : K \to K$ *ein Körperautomorphismus. Eine* ζ*-Sesquilinearform* σ *auf* $\mathbf{V}$ *heißt* **nichtausgeartet***, wenn*

$$\mathbf{V}^{\perp} = \{\mathbf{0}\}.$$

Ein Unterraum $\mathbf{U} \leq \mathbf{V}$ *heißt* **nichtausgeartet***, wenn* $\sigma|_{\mathbf{U}}$ *auf* $\mathbf{U}$ *nichtausgeartet ist.*

$\mathbf{U}$ ist daher genau dann nichtausgeartet, wenn

$$\mathbf{U} \cap \mathbf{U}^{\perp} = \{\mathbf{0}\}.$$

Einige Eigenschaften von Orthogonalräumen sind im folgenden Satz zusammengefaßt:

Satz 10.67 *Sei* $\mathbf{V}$ *ein Vektorraum über dem Körper* K*,* $\zeta : K \to K$ *ein Körperautomorphismus,* σ *eine orthosymmetrische* ζ*-Sesquilinearform auf* $\mathbf{V}$*,* $\mathbf{M}, \mathbf{N}$ *nichtleere Teilmengen von* $\mathbf{V}$ *und* $\mathbf{U}_i$ $(i \in I)$ *ein System von Unterräumen von* $\mathbf{V}$*.*

1. $\mathbf{M} \subseteq \mathbf{N} \Longrightarrow \mathbf{N}^{\perp} \subseteq \mathbf{M}^{\perp}$.
2. $[\mathbf{M}]^{\perp} = \mathbf{M}^{\perp}$.
3. $\mathbf{M}^{\perp} \leq \mathbf{V}$.
4. $\mathbf{M}^{\perp} = f_{\sigma}^{-1}(\mathbf{M}^{\circ})$.
5. $\mathbf{V}^{\perp} = \mathrm{kern}(f_{\sigma})$.
6. $\mathbf{M} \subseteq (\mathbf{M}^{\perp})^{\perp}$.
7. $\left(\sum_{i\in I} \mathbf{U}_i\right)^{\perp} = \bigcap_{i\in I} \mathbf{U}_i^{\perp}$.
8. $\sum_{i\in I} \mathbf{U}_i^{\perp} \subseteq \left(\bigcap_{i\in I} \mathbf{U}_i\right)^{\perp}$,
   wobei Gleichheit gilt, wennn $\dim(\mathbf{V})$ *endlich ist und* σ *nichtausgeartet ist.*

Ist $\mathbf{U}$ ein Unterraum eines endlichdimensionalen Vektorraums $\mathbf{V}$, so können genaue Aussagen über die Dimensionen der auftretenden Vektorräume gemacht werden.

Satz 10.68 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über dem Körper* K*,* $\zeta : K \to K$ *ein Körperautomorphismus und* σ *eine orthosymmetrische* ζ*-Sesquilinearform* σ *auf* $\mathbf{V}$*. Dann gilt:*

$$\dim(\mathbf{U}) + \dim(\mathbf{U}^{\perp}) = \dim(\mathbf{V}) + \dim(\mathbf{U} \cap \mathbf{V}^{\perp}).$$

Korollar 10.69 *Sei* $\mathbf{V}$ *ein Vektorraum über dem Körper* K*,* $\zeta : K \to K$ *ein Körperautomorphismus,* σ *eine nichtausgeartete orthosymmetrische* ζ*-Sesquilinearform auf* $\mathbf{V}$*, so gilt für alle Unterräume* $\mathbf{U} \leq \mathbf{V}$

$$\mathbf{U} = (\mathbf{U}^{\perp})^{\perp}.$$

Der folgende Satz spielt bei den Struktursätzen für Hermitesche Formen eine große Rolle.

Satz 10.70 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über dem Körper* K*,* $\zeta : K \to K$ *ein Körperautomorphismus und eine* ζ*-Sesquilinearform* σ *auf* $\mathbf{V}$*. Dann gilt für jeden nichtausgearteten Unterraum* $\mathbf{U} \leq \mathbf{V}$

$$\mathbf{U} \oplus \mathbf{U}^{\perp} = \mathbf{V}.$$

Damit erhält man wie bei symmetrischen Bilinearformen die Möglichkeit zu diagonalisieren.

Satz 10.71 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über dem Körper* K*,* $\zeta : K \to K$ *ein Körperautomorphismus und* σ *eine* ζ*-symmetrische Sesquilinearform vom Rang* r*. Dann gibt es* r *paarweise orthogonale, nichtausgeartete, eindimensionale Unterräume* $\mathbf{A}_1, \ldots, \mathbf{A}_r$ *von* $\mathbf{V}$ *mit*

$$\mathbf{V} = \mathbf{A}_1 \oplus \cdots \oplus \mathbf{A}_r \oplus \mathbf{V}^{\perp}.$$

Es gibt daher eine Basis $\mathbf{B}$ von $\mathbf{V}$, so daß die Koordinatenmatrix $\Phi_{\mathbf{B}}(\sigma)$ Diagonalgestalt

$$\Phi_{\mathbf{B}}(\sigma) = \operatorname{diag}(g_1, g_2, \ldots, g_r, 0, 0, \ldots, 0)$$

mit $g_1, g_2, \ldots, g_r \in K^{\times}$ hat.

Entsprechend gilt:

Satz 10.72 *Sei* $\zeta : K \to K$ *ein Körperautomorphismus mit* $\zeta^2 = \mathrm{id}_K$*. Dann ist jede* ζ*-symmetrische Matrix* $A \in K^{n \times n}$ *zu einer Diagonalmatrix* ζ*-kongruent.*

10.5.4 Hermitesche Formen

Eine Hermitesche Sesquilinearform (kurz auch Hermitesche Form, was aber zu Verwechslungen führen kann, wie wir gleich sehen werden) σ ist eine ζ-symmetrische ζ-Sesquilinearform auf einem Vektorraum $\mathbf{V}$ über $\mathbb{C}$, wobei ζ die komplexe Konjugation $\zeta(z) = \overline{z}$ bezeichnet. Setzt man in σ an beiden Stellen denselben Vektor ein, so erhält man ein Analogon zu quadratischen Formen, nämlich sogenannte *Hermitesche Formen*.

Definition 10.73 *Sei* $\mathbf{V}$ *ein Vektorraum über* $\mathbb{C}$*. Dann heißt eine Abbildung* $h : \mathbf{V} \to \mathbb{R}$ **Hermitesche Form***, wenn folgende drei Eigenschaften erfüllt sind:*

1. $h(x\mathbf{a}) = x\overline{x}\, h(\mathbf{a})$.
2. $h(\mathbf{a} + \mathbf{b}) + h(\mathbf{a} - \mathbf{b}) = 2(h(\mathbf{a}) + h(\mathbf{b}))$.
3. $\sigma(\mathbf{a}, \mathbf{b}) := \frac{1}{2}(h(\mathbf{a}+\mathbf{b}) + ih(\mathbf{a}-\mathbf{b}) - (1+i)(h(\mathbf{a}) + h(\mathbf{b}))$ *ist eine Hermitesche Sesquilinearform.*

Man beachte, daß Hermitesche Formen immer reellwertig sind.

Ist $A \in \mathbb{C}^{n \times n}$ eine Hermitesche Matrix, d.h. $A^H = A$, dann ist

$$h(\mathbf{a}) := \mathbf{a}^T A \overline{\mathbf{a}}$$

eine Hermitesche Form.

Hermitesche Formen treten auch als Fortsetzung reeller symmetrischer Bilinearformen in natürlicher Weise auf.

Lemma 10.74 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über* $\mathbb{R}$ *und bezeichne* $\mathbf{V}_{\mathbb{C}}$ *die komplexe Erweiterung von* $\mathbf{V}$. *Weiters sei* σ *eine symmetrische Bilinearform auf* $\mathbf{V}$. *Dann gibt es genau zwei Möglichkeiten,* σ *auf* $\mathbf{V}_{\mathbb{C}}$ *fortzusetzen, einerseits zu einer symmetrischen Bilinearform auf* $\mathbf{V}_{\mathbb{C}}$ *und andererseits zu einer Hermiteschen Form.*

Satz 10.75 *Sei* $\mathbf{V}$ *ein Vektorraum über* $\mathbb{C}$. *Jede Hermitesche Sesquilinearform* σ *bedingt durch die Festlegung*

$$h(\mathbf{a}) := \sigma(\mathbf{a}, \mathbf{a})$$

eine Hermitesche Form. Andererseits kann σ *aus der Kenntnis von* h *ermittelt werden. D.h. Hermitesche Sesquilinearformen und Hermitesche Formen entsprechen einander eineindeutig.*

Da Hermitesche Formen reellwertig sind, kann man in Analogie zu den quadratischen Formen folgende Begriffe einführen:

Definition 10.76 *Sei* $\mathbf{V}$ *ein Vektorraum über* $\mathbb{C}$, σ *eine Hermitesche Form auf* $\mathbf{V}$ *und* $\mathbf{U} \leq \mathbf{V}$ *ein Unterraum von* $\mathbf{V}$.

1. $\mathbf{U}$ *heißt* **positiv definit**, *wenn für alle* $\mathbf{a} \in \mathbf{U} \setminus \{\mathbf{0}\}$ $\sigma(\mathbf{a}, \mathbf{a}) > 0$ *ist.*
2. $\mathbf{U}$ *heißt* **positiv semidefinit**, *wenn für alle* $\mathbf{a} \in \mathbf{U}$ $\sigma(\mathbf{a}, \mathbf{a}) \geq 0$ *ist.*
3. $\mathbf{U}$ *heißt* **negativ definit**, *wenn für alle* $\mathbf{a} \in \mathbf{U} \setminus \{\mathbf{0}\}$ $\sigma(\mathbf{a}, \mathbf{a}) < 0$ *ist.*
4. $\mathbf{U}$ *heißt* **negativ semidefinit**, *wenn für alle* $\mathbf{a} \in \mathbf{U}$ $\sigma(\mathbf{a}, \mathbf{a}) \leq 0$ *ist.*
5. $\mathbf{U}$ *heißt* **indefinit**, *wenn es* $\mathbf{a}, \mathbf{b} \in \mathbf{U}$ *mit* $\sigma(\mathbf{a}, \mathbf{a}) > 0$ *und* $\sigma(\mathbf{b}, \mathbf{b}) < 0$ *gibt.*

In derselben Weise kann man entsprechende Begriffe für Hermitesche Matrizen $A \in \mathbb{C}^{n \times n}$ einführen, indem man $\sigma(\mathbf{a}, \mathbf{a}) = h(\mathbf{a}) = \mathbf{a}^T A \overline{\mathbf{a}}$ setzt.

Definition 10.77 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über* $\mathbb{C}$ *und* σ *eine Hermitesche Form. Definiert man*

$$p := \max\{\dim(\mathbf{U}) \,|\, \mathbf{U} \leq \mathbf{V} \ \wedge \ \mathbf{U} \textit{ ist positiv definit}\}$$

und

$$q := \max\{\dim(\mathbf{U}) \,|\, \mathbf{U} \leq \mathbf{V} \ \wedge \ \mathbf{U} \textit{ ist negativ definit}\},$$

dann heißt die Differenz

$$p - q$$

Signatur *von* σ.

Entsprechend wird die Signatur einer Hermiteschen Matrix $A \in \mathbb{C}^{n \times n}$ (d.h. $A^H = A$) definiert.

Satz 10.78 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über* $\mathbb{C}$ *und* σ *eine Hermitesche Form. Dann gibt es eine Basis* $\mathbf{B}$ *von* $\mathbf{V}$ *mit*

$$\Phi_{\mathbf{B}}(\sigma) = \begin{pmatrix} E_p & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -E_q & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{10.3}$$

wobei $p + q$ *der Rang von* σ *und* $p - q$ *die Signatur von* σ *ist.*

Diese Version des **Trägheitssatzes von Sylvester** löst nun das Kongruenzproblem Hermitesche Formen.

Satz 10.79 *Zwei Hermitesche Formen σ, τ auf einem endlichdimensionalen Vektorraum $\mathbf{V}$ über $\mathbb{C}$ sind genau dann kongruent, wenn ihre Ränge und Signaturen gleich sind.*

Entsprechend ist jede Hermitesche Matrix $A \in \mathbb{C}^{n\times n}$ zu einer Matrix der Form (10.3) Hermitesch-kongruent, d.h. es gibt eine reguläre Matrix $P \in \mathbb{C}^{n\times n}$ mit

$$PAP^H = \begin{pmatrix} E_p & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -E_q & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix},$$

wobei p und q eindeutig bestimmt sind. Außerdem gilt:

Satz 10.80 *Zwei Hermitesche Matrizen $A, B \in \mathbb{C}^{n\times n}$ (d.h. $A^H = A$, $B^H = B$) sind genau dann Hermitesch-kongruent, wenn ihre Ränge und ihre Signaturen gleich sind.*

Man definiert nun in entsprechender Weise definite Hermitesche Formen bzw. definite Hermitesche Matrizen.

Definition 10.81 *Sei $\mathbf{V}$ ein Vektorraum über den komplexen Zahlen $\mathbb{C}$ und h eine Hermitesche Form auf $\mathbf{V}$.*

1. *h heißt* **positiv definit**, *wenn $h(\mathbf{a}) > 0$ für alle $\mathbf{a} \in \mathbf{V} \setminus \{\mathbf{0}\}$.*
2. *h heißt* **positiv semidefinit**, *wenn $h(\mathbf{a}) \geq 0$ für alle $\mathbf{a} \in \mathbf{V}$.*
3. *h heißt* **negativ definit**, *wenn $h(\mathbf{a}) < 0$ für alle $\mathbf{a} \in \mathbf{V} \setminus \{\mathbf{0}\}$.*
4. *h heißt* **negativ semidefinit**, *wenn $h(\mathbf{a}) \leq 0$ für alle $\mathbf{a} \in \mathbf{V}$.*
5. *h heißt* **indefinit**, *wenn es Vektoren $\mathbf{a}, \mathbf{b} \in \mathbf{V}$ mit $h(\mathbf{a}) > 0$ und $h(\mathbf{b}) < 0$ gibt.*

Entsprechend definiert man für symmetrische Matrizen:

Definition 10.82 *Sei $G \in \mathbb{C}^{n\times n}$ eine Hermitesche Matrix, d.h. $G^H = G$.*

1. *G heißt* **positiv definit**, *wenn $\mathbf{a}^T G\overline{\mathbf{a}} > 0$ für alle $\mathbf{a} \in \mathbb{C}^{n\times 1} \setminus \{\mathbf{0}\}$.*
2. *G heißt* **positiv semidefinit**, *wenn $\mathbf{a}^T G\overline{\mathbf{a}} \geq 0$ für alle $\mathbf{a} \in \mathbb{C}^{n\times 1}$.*
3. *G heißt* **negativ definit**, *wenn $\mathbf{a}^T G\overline{\mathbf{a}} < 0$ für alle $\mathbf{a} \in \mathbb{C}^{n\times 1} \setminus \{\mathbf{0}\}$.*
4. *G heißt* **negative semidefinit**, *wenn $\mathbf{a}^T G\overline{\mathbf{a}} \leq 0$ für alle $\mathbf{a} \in \mathbb{C}^{n\times 1}$.*
5. *G heißt* **indefinit**, *wenn es Vektoren $\mathbf{a}, \mathbf{b} \in \mathbb{C}^{n\times 1}$ mit $\mathbf{a}^T G\overline{\mathbf{a}} > 0$ und $\mathbf{a}^T G\overline{\mathbf{a}} < 0$ gibt.*

Wieder kann mit Hilfe der Hauptminoren entschieden werden, ob eine hermitesche Form positiv bzw. negativ definit ist (**Hauptminorenkriterium**).

Satz 10.83 *Sei h eine Hermitesche Form auf einem endlichdimensionalen Vektorraum $\mathbf{V}$ über $\mathbb{C}$ und σ jene Hermitesche Sesquilinearform auf $\mathbf{V}$ mit $h(\mathbf{a}) = \sigma(\mathbf{a}, \mathbf{a})$.*

Dann ist h genau dann positiv definit, wenn alle Hauptminoren der Matrix $G = \Phi_{\mathbf{B}}(\sigma)$ (für eine beliebige Basis $\mathbf{B}$ von $\mathbf{V}$) positiv sind:

$$\det G_{(1,2,\ldots,k)} > 0 \qquad (1 \le k \le n).$$

h ist genau dann negativ definit, wenn die Hauptminoren abwechselnd negativ und positiv sind:

$$\det G_{(1,2,\ldots,2k-1)} < 0 \quad \textit{und} \quad \det G_{(1,2,\ldots,2k)} > 0 \quad (1 \le k \le n/2).$$

Satz 10.84 *Sei $G \in \mathbb{C}^{n\times n}$ eine Hermitesche Matrix, d.h. $G^H = G$.*

Dann ist G genau dann positiv definit, wenn alle Hauptminoren der Matrix G positiv sind:

$$\det G_{(1,2,\ldots,k)} > 0 \qquad (1 \le k \le n).$$

G ist genau dann negativ definit, wenn die Hauptminoren abwechselnd negativ und positiv sind:

$$\det G_{(1,2,\ldots,2k-1)} < 0 \quad \textit{und} \quad \det G_{(1,2,\ldots,2k)} > 0 \quad (1 \le k \le n/2).$$

10.6 Quadriken

Ziel dieses Abschnittes ist es im wesentlichen, die Nullstellenmenge eines quadratischen Polynoms in mehreren Veränderlichen zu untersuchen. Sei etwa

$$\lambda(x_1, x_2) = 2x_1^2 - 4x_1x_2 + 5x_2^2 - 3x_1 + x_2 - 7$$

ein quadratischen Polynom vom Grad 2. Es besteht aus einem rein quadratischen Teil $(2x_1^2 - 4x_1x_2 + 5x_2^2)$, einem linearen Teil $(-3x_1 + x_2)$ und einer Konstanten (-7). Es stellt sich heraus, daß die Nullstellenmenge $\{(x_1, x_2) \in K^2 \mid \lambda(x_1, x_2) = 0\}$ leichter untersucht werden kann, wenn man zu *homogenen Koordinaten* übergeht. Man betrachtete anstelle von λ das rein quadratische Polynom

$$q(x_0, x_1, x_2) = 2x_1^2 - 4x_1x_2 + 5x_2^2 - 3x_0x_1 + x_0x_2 - 7x_0^2 = x_0^2\lambda(x_1/x_0, x_2/x_0).$$

Offensichtlich ist q eine quadratische Form und es gilt $\lambda(x_1, x_2) = q(1, x_1, x_2)$. Da q quadratische Form ist, ist mit jeder Lösung (x_0, x_1, x_2) von $q(x_0, x_1, x_2) = 0$ auch jedes Vielfache $t(x_0, x_1, x_2) = (tx_0, tx_1, tx_2)$ Lösung: $q(tx_0, tx_1, tx_2) = t^2q(x_0, x_1, x_2) = 0$. Insbesondere erhält man aus einer Lösung (x_0, x_1, x_2) von $q(x_0, x_1, x_2) = 0$ (mit $x_0 \ne 0$) auch die Lösung der Form $(1, x_1/x_0, x_2/x_0)$ und damit eine Lösung von $\lambda(x_1', x_2') = 0$ mit $x_1' = x_1/x_0$ und $x_2' = x_2/x_0$.

Das heißt, es ist (im wesentlichen) gleichwertig, die Nullstellenmenge von λ oder die Nullstellenmenge von q zu betrachten. Der Vorteil von q ist, daß Methoden der linearen Algebra direkt auf q anwendbar sind, da q mit einer symmetrischen Bilinearform σ identifiert werden kann. Weiters ist q ein homogenes Polynom und kann daher auch über einem projektiven Raum betrachtet werden, was eine weitere Vereinfachung mit sich bringt. Das ist auch der Ausgangspunkt des folgenden Unterabschnitts.

10.6.1 Projektive Quadriken

Es wird jetzt immer vorausgesetzt, daß die Charakteristik des Skalarkörpers K ungleich 2 ist, damit symmetrische Bilinearformen und quadratische Formen einander eineindeutig entsprechen.

Definition 10.85 *Sei* $\mathbf{V}$ *ein Vektorraum über einem Körper* K *mit* $\mathrm{char}(K) \neq 2$*. Unter einer* **projektiven Quadrik** Q *auf dem projektiven Raum* $\mathcal{P}(\mathbf{V})$ *versteht man die Menge der projektiven Punkte*

$$Q = Q(\sigma) := \{A = K\mathbf{a} \in \mathcal{P}(\mathbf{V}) \,|\, \sigma(\mathbf{a}, \mathbf{a}) = 0\},$$ [1]

wobei $\sigma \in BL(\mathbf{V})$ *eine symmetrische Bilinearform ist.*

Man hätte anstelle der symmetrischen Bilinearform σ auch die durch $q(\mathbf{a}) = \sigma(\mathbf{a}, \mathbf{a})$ gegebene quadratische Form q verwenden können.

Ist $\dim(\mathbf{V}) = n + 1 < \infty$ und $\mathbf{B} = \{\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_n\}$ eine Basis von $\mathbf{V}$ und $G = (g_{ij}) = \Phi_{\mathbf{B}}(\sigma)$ die Koordinatenmatrix von σ, dann ist $Q = Q(\sigma)$ die Menge der Punkte $A = K(x_0\mathbf{b}_0 + x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n)$ mit

$$\sum_{i,j=0}^{n} g_{ij} x_i x_j = 0.$$

Es geht also um die *Nullstellenmenge* eines (homogenen) quadratischen Polynoms.

Da σ symmetrisch ist, gibt es auch eine Basis $\mathbf{B} = \{\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_n\}$ von $\mathbf{B}$, so daß $G = \Phi_{\mathbf{B}}(\sigma)$ eine Diagonalmatrix ist, d.h. man muß nur quadratische Gleichungen der Form

$$g_0 x_0^2 + g_1 x_1^2 + \cdots + g_n x_n^2 = 0$$

untersuchen.

Definition 10.86 *Sei* $\mathbf{V}$ *ein Vektorraum über einem Körper* K *mit* $\mathrm{char}(K) \neq 2$ *und* Q *eine projektive Quadrik im projektiven Raum* $\mathcal{P}(\mathbf{V})$*.*

Eine projektive Gerade $g \in \mathcal{P}(\mathbf{V})$ *heißt* **Tangente***, wenn entweder* g *und* Q *nur einen Punkt gemeinsam haben oder wenn* g *in* Q *enthalten ist.*

Ein Punkt $S \in Q$ *heißt* **singulär***, wenn jede projektive Gerade* g*, die* S *enthält Tangente von* Q *ist.*

Ein Punkt $S \in Q$ *heißt* **regulär***, wenn er nicht singulär ist.*

Q *heißt* **singulär***, wenn* Q *einen singulären Punkt enthält.*

Q *heißt* **regulär***, wenn alle Punkte von* Q *regulär sind.*

Die Menge aller singulären Punkte einer Quadrik wird als **Spitzenraum** *bezeichnet.*

Satz 10.87 *Sei* $\mathbf{V}$ *ein Vektorraum über einem Körper* K *mit* $\mathrm{char}(K) \neq 2$ *und* Q *eine projektive Quadrik im projektiven Raum* $\mathcal{P}(\mathbf{V})$*. Dann ist eine projektive Gerade genau dann Tangente von* Q *im Punkt* P*, d.h.* $P \in Q \cap g$*, wenn*

$$g \subseteq P^{\perp}.$$

[1] Die projektive Geometrie $\mathcal{P}(\mathbf{V})$ über einem Vektorraum $\mathbf{V}$ Besteht aus allen Unterräumen von $\mathbf{V}$. Insbesondere ist ein *projektiver Punkt* ein eindimensionaler Unterraum, der durch die Vielfachen $K\mathbf{a} = \{x\mathbf{a} \,|\, x \in K\}$ eines Vektors $\mathbf{a} \neq \mathbf{0}$ beschrieben werden kann.

wobei $\perp = \perp_\sigma$ *die von* σ *induzierte Orthogonalitätsrelation ist.*

Weiters gilt für jede projektive Gerade g*, die nicht in* Q *enthalten ist, daß sie höchstens* 2 *Punkte mit* Q *gemeinsam hat.*

Satz 10.88 *Sei* $\mathbf{V}$ *ein Vektorraum über einem Körper* K *mit* $\mathrm{char}(K) \neq 2$ *und* Q *eine projektive Quadrik. Dann besteht der Spitzenraum genau aus den Punkten von* $\mathcal{P}(\mathbf{V}^\perp)$*.*

Korollar 10.89 *Eine projektive Quadrik* $Q = Q(\sigma)$ *ist genau dann regulär, wenn die Bilinearform* σ *nichtausgeartet ist.*

Der nächste Satz zeigt, daß (im wesentlichen) die Punktmenge Q die zugrundeliegende Bilinearform σ rekonstruiert.

Satz 10.90 *Sei* $\mathbf{V}$ *ein Vektorraum über einem Körper* K *mit* $\mathrm{char}(K) \neq 2$ *und* Q *eine projektive Quadrik in* $\mathcal{P}(\mathbf{V})$*, die wenigstens einen regulären Punkt enthält, so daß es zwei Bilinearformen* $\sigma_1, \sigma_2 \in BL(\mathbf{V})$ *mit*

$$Q = Q(\sigma_1) = Q(\sigma_2)$$

gibt. Dann existiert $c \in K^\times$ *mit*

$$\sigma_1 = c\,\sigma_2.$$

10.6.2 Affine Quadriken

Definition 10.91 *Sei* $\mathbf{V}$ *ein Vektorraum über einem Körper* K *mit* $\mathrm{char}(K) \neq 2$ *und* $\mathbf{N} = \mathbf{c} + \mathbf{U}$ *ein Nebenraum eines Unterraums* $\mathbf{U} \leq \mathbf{V}$ *von* $\mathbf{V}$*. Eine Abbildung der Form*

$$\lambda(\mathbf{a}) = q(\mathbf{a} - \mathbf{c}) + l(\mathbf{a} - \mathbf{c}) + g,$$

wobei q *eine quadratische Form auf* $\mathbf{U}$*,* l *eine Linearform auf* $\mathbf{U}$ *und* $g \in K$ *ist, heißt* **quadratische Funktion** *auf dem affinen Raum* $\mathcal{A}(\mathbf{N})$*.*

Für die Definition einer quadratischen Funktion benötigt man also einen Vektor (Repräsentaten) $\mathbf{c} \in \mathbf{N}$, eine quadratische Form q, eine Linearform l und eine Konstante g. Verändert man den Repräsentanten $\mathbf{c}$, so verändern sich auch l und g, aber nicht q.

Betrachtet man eine affine Basis $\{\mathbf{c}, \mathbf{c} + \mathbf{b}_1, \ldots, \mathbf{c} + \mathbf{b}_n\}$ von $\mathbf{N}$ (mit Ursprung $\mathbf{c}$ und Einheitspunkten $\mathbf{c} + \mathbf{b}_i$), dann gilt für $\mathbf{a} = \mathbf{c} + \sum_{i=1}^n x_i \mathbf{b}_i$

$$\lambda(\mathbf{a}) = \sum_{i,j=1}^n g_{ij} x_i x_j + \sum_{i=0}^n l_i x_i + g.$$

Auf der Ebene der Koordinaten ist eine quadratische Funktion daher ein allgemeines Polynom von Grad 2.

Quadratische Funktionen hängen eng mit quadratischen Formen bzw. mit symmetrischen Bilinearformen zusammen.

Satz 10.92 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über einem Körper* K *mit* $\operatorname{char}(K) \neq 2$, $\mathbf{H} \leq \mathbf{V}$ *eine Hyperebene von* $\mathbf{V}$ *und* $\mathbf{c} \notin \mathbf{H}$. *Ist* λ *eine quadratische Funktion auf* $\mathbf{N} = \mathbf{c} + \mathbf{H}$, *dann gibt es eine eindeutig bestimmte quadratische Form* q *auf ganz* $\mathbf{V}$, *so daß für alle* $\mathbf{a} \in \mathbf{N}$

$$\lambda(\mathbf{a}) = q(\mathbf{a})$$

ist. Die quadratischen Funktionen auf $\mathbf{N}$ *und die quadratischen Formen auf* $\mathbf{V}$ *entsprechen daher eineindeutig.*

Definition 10.93 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über einem Körper* K *mit* $\operatorname{char}(K) \neq 2$ *und* $\mathbf{N} = \mathbf{c} + \mathbf{U}$ *ein Nebenraum von* $\mathbf{U} \leq \mathbf{V}$ *eines Unterraums von* $\mathbf{V}$. *Eine* **affinen Quadrik** $Q_{aff} = Q_{aff}(\lambda)$ *auf dem affinen Raum* $\mathcal{A}(\mathbf{N})$ *ist durch*

$$Q_{aff} = Q_{aff}(\lambda) := \{a = \{\mathbf{a}\} \in \mathcal{A}(\mathbf{N}) \,|\, \lambda(\mathbf{a}) = 0\}$$

definiert, wobei λ *eine quadratische Funktion auf* $\mathbf{N}$ *bezeichnet.*

Wegen Satz 10.92 kann jede affine Quadrik als *Schnitt* einer projektiven Quadrik mit dem zugrundeliegenden Nebenraum angesehen werden. Daher können viele Eigenschaften affiner Quadriken aus entsprechenden Eigenschaften für projektive Quadriken abgeleitet werden.

Wieder läßt sich die definierende quadratische Funktion einer affinen Quadrik $Q_{aff}(\lambda)$ (im wesentlichen) aus den Menge der Punkte aus $Q_{aff}(\lambda)$ rekonstruieren.

Satz 10.94 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über einem Körper* K *mit* $\operatorname{char}(K) \neq 2$ *und* $\mathbf{N} = \mathbf{c} + \mathbf{U}$ *ein Nebenraum von* $\mathbf{U} \leq \mathbf{V}$ *eines Unterraums von* $\mathbf{V}$. *Weiters sei* $Q_{aff} \subseteq \mathbf{N}$ *eine affine Quadrik auf* $\mathbf{N}$, *so daß die affine Hülle der Punkte aus* Q_{aff} *ganz* $\mathbf{N}$ *ist. Sind nun* $\lambda_1(\mathbf{a}), \lambda_2(\mathbf{a})$ *zwei quadratische Funktionen auf* $\mathbf{N}$ *mit* $Q_{aff} = Q_{aff}(\lambda_1) = Q_{aff}(\lambda_2)$. *Dann gibt es ein* $c \in K^\times$ *mit*

$$\lambda_1 = c\lambda_2.$$

10.6.3 Struktursätze für Quadriken

Da projektive Quadriken auf symmetrische Bilinearformen zurückgeführt werden können, gilt der folgende Satz

Satz 10.95 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über einem Körper* K *mit* $\operatorname{char}(K) \neq 2$ *und* Q *eine projektive Quadrik im projektiven Raum* $\mathcal{P}(\mathbf{V})$. *Dann gibt es eine Basis* $\mathbf{B}$ *von* $\mathbf{V}$, *so daß*

$$\sigma(\mathbf{b}_i, \mathbf{b}_j) = g_i \delta_{ij},$$

gilt.

Bezüglich dieser Basis wird die Quadrik durch die Gleichung

$$g_0 x_0^2 + g_1 x_1^2 + \cdots + g_n x_n^2 = 0$$

beschrieben. Über den reellen Zahlen $\mathbb{R}$ kann durch Skalierung $|g_i| = 1$ erreicht werden, das Vorzeichen aber nicht verändert werden. Man erhält daher die Fälle

$$x_0^2 + x_1^2 + \cdots + x_{p-1}^2 - x_p^2 - \cdots - x_{p+q-1}^2 = 0$$

mit $p \geq q$, $p + q \leq n + 1$,

Der affine Fall ist etwas komplizierter, da man auf die Hyperebene $\mathbf{H}$ Rücksicht nehmen muß. Hier gilt:

Satz 10.96 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum über einem Körper* K *mit* $\mathrm{char}(K) \neq 2$ *und* $\mathbf{N} = \mathbf{c} + \mathbf{U}$ *ein Nebenraum von* $\mathbf{U} \leq \mathbf{V}$ *eines Unterraums von* $\mathbf{V}$*. Weiters sei* $Q_{aff} \subseteq \mathbf{N}$ *eine affine Quadrik auf* $\mathbf{N}$*. Dann gibt es eine Basis* $\mathbf{B}$ *von* $\mathbf{V}$*, so daß* Q_{aff} *bezüglich dieser Basis entweder durch*

$$g_0 + g_1 x_1^2 + \cdots + g_n x_n^2 = 0$$

oder durch

$$g_1 x_1 + g_2 x_2^2 + \cdots + g_n x_n^2 = 0$$

beschrieben werden kann.

Über $\mathbb{R}$ ist die Situation wieder etwas einfacher. Hier gibt es zu jeder affinen Quadrik Q_{aff} ein Koordinatensystem, so daß die Punkte $a = \{\mathbf{a}\} \in Q_{aff}$ mit den Koordinaten $(x_1, \ldots, x_n)$ durch eine Gleichung der Form

$$x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2 = 0$$

mit $p \geq q$, $p + q \leq n$, durch

$$x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2 = \pm 1$$

mit $p \geq q$, $p + q \leq n$ oder durch

$$x_2^2 + \cdots + x_{p+1}^2 - x_{p+2}^2 - \cdots - x_{p+q-1}^2 = \pm x_1$$

beschrieben werden kann.

Kapitel 11

Skalarprodukte und Euklidische Vektorräume

11.1 Skalarprodukte und reziproke Basen

11.1.1 Vektorräume mit Skalarprodukt

Definition 11.1 *Sei* $\mathbf{V}$ *ein Vektorraum über dem Körper* K *und* σ

- *eine nicht ausgeartete symmetrische Bilinearform auf* $\mathbf{V} \times \mathbf{V}$*,*
- *eine nicht ausgeartete alternierende Bilinearform auf* $\mathbf{V} \times \mathbf{V}$ *oder*
- *eine nicht ausgeartete* ζ*-symmetrische* ζ*-Sesquilinearform auf* $\mathbf{V} \times \mathbf{V}$*,*

wobei $\zeta : K \to K$ *einen Körperautomorphismus bezeichnet.* σ *heißt dann* **Skalarprodukt** *und das Paar* $(\mathbf{V}, \sigma)$ **Vektorraum mit Skalarprodukt***.*

In einem Vektorraum mit Skalarprodukt schreibt man anstelle von $\sigma(\mathbf{a}, \mathbf{b})$ *auch*

$$\langle \mathbf{a}, \mathbf{b} \rangle \qquad \textit{oder} \qquad \langle \mathbf{a}, \mathbf{b} \rangle_\sigma .$$

Ein Vektorraum $\mathbf{V}$ *mit einer symmetrischen Bilinearform (bzw. mit einer* ζ*-symmetrischen* ζ*-Sesquilinearform) als Skalarprodukt wird auch als (*ζ*–)***symmetrischer Raum** *bezeichnet.*

Ein Vektorraum $\mathbf{V}$ *mit einer alternierenden Bilinearform als Skalarprodukt wird als* **symplektischer Raum** *bezeichnet.*

Symmetrische Bilinearformen sind natürlich auch ζ-symmetrische ζ-Sesquilinearformen, wenn $\zeta = \mathrm{id}_K$ ist. Wegen ihrer Bedeutung wurden sie hier trotzdem extra genannt. Weiters beachte man, daß jedes Skalarprodukt orthosymmetrisch ist. Es muß also, falls $\dim(\mathbf{V}) > 1$ ist, immer $\zeta^2 = \mathrm{id}_K$ gelten.

Beispiel 11.2 Im Vektorraum $\mathbf{V} = K^{n \times 1}$ ist

$$\langle (x_1, \ldots, x_n)^T, (y_1, \ldots, y_n)^T \rangle := x_1 y_1 + \cdots + x_n y_n$$

das *gewöhnliche* Skalarprodukt.

Beispiel 11.3 Im Vektorraum $\mathbb{C}^{n\times 1}$ ist

$$\langle (x_1,\dots,x_n)^T,(y_1,\dots,y_n)^T\rangle := x_1\overline{y}_1+\dots+x_n\overline{y}_n$$

ein von einer Hermiteschen Form stammendes Skalarprodukt.

Beispiel 11.4 Sei $\mathbf{V}$ der Vektorraum aller stetigen Funktionen $f:[0,1]\to\mathbb{C}$. Dann ist

$$\langle f,g\rangle := \int_0^1 f(x)\overline{g}(x)\,dx$$

ein (von einer Hermiteschen Form stammendes) Skalarprodukt.

Beispiel 11.5 Sei $\mathbf{V}=\mathbb{R}^{4\times 1}$ und ein Skalarprodukt durch

$$\langle (x_1,x_2,x_3,x_4)^T,(y_1,y_2,y_3,y_4)^T\rangle := x_1y_1+x_2y_2+x_3y_3-x_4y_4$$

definiert. Man beachte, daß dieses Skalarprodukt (definitionsgemäß) nicht ausgeartet ist, es aber Unterräume, wie z.B.

$$\mathbf{U} := [(1,0,0,1)^T]$$

gibt, auf denen das Skalarprodukt ausgeartet ist.

Definition 11.6 *Ein Vektorraum mit Skalarprodukt* $(\mathbf{V},\sigma)$ *heißt* **anisotrop**, *wenn* σ *auf jedem Unterraum* $\mathbf{U}\le\mathbf{V}$ *nicht ausgeartet ist.*

Lemma 11.7 *Ein Vektorraum mit Skalarprodukt* $(\mathbf{V},\sigma)$ *ist genau dann anisotrop, wenn für alle Vektoren* $\mathbf{a}\in\mathbf{V}\setminus\{\mathbf{0}\}$

$$\langle \mathbf{a},\mathbf{a}\rangle \neq 0$$

gilt.

Man bezeichnet auch einen Vektor $\mathbf{a}$ eines Vektorraums mit Skalarprodukt $(\mathbf{V},\sigma)$ als anisotrop, wenn $\langle \mathbf{a},\mathbf{a}\rangle\neq 0$.

Im endlichdimensionalen Fall kann bei fest gewählter Basis $\mathbf{B}$ von $\mathbf{V}$ einem Skalarprodukt σ eine Matrix $\Phi_{\mathbf{B}}(\sigma)$ zugeornet werden. Da $\Phi_{\mathbf{B}}(\sigma)$ genau dann regulär ist, wenn σ nicht ausgeartet ist, ist im Fall eines Skalarprodukts $\Phi_{\mathbf{B}}(\sigma)$ immer regulär.

Satz 11.8 *Sei* $(\mathbf{V},\sigma)$ *ein* (ζ*–)symmetrischer Vektorraum mit endlicher Dimension* $\dim(\mathbf{V})=n$. *Dann gibt es eine Basis* $\mathbf{B}$ *von* $\mathbf{V}$ *und Skalare* $c_1,c_2,\dots,c_n\in K^\times$ *mit*

$$\langle \mathbf{a},\mathbf{b}\rangle = c_1x_1\zeta(y_1)+c_2x_2\zeta(y_2)+\dots+c_nx_n\zeta(y_n),$$

wobei $\Phi_{\mathbf{B}}(\mathbf{a})=(x_1,\dots,x_n)^T$ *und* $\Phi_{\mathbf{B}}(\mathbf{b})=(y_1,\dots,y_n)^T$.

Satz 11.9 *Sei* $(\mathbf{V},\sigma)$ *ein symplektischer Vektorraum mit endlicher Dimension* $\dim(\mathbf{V})=n$. *Dann ist* n *gerade und es gibt eine Basis* $\mathbf{B}$ *von* $\mathbf{V}$ *mit*

$$\langle \mathbf{a},\mathbf{b}\rangle = x_1y_2-x_2y_1+x_3y_4-x_4y_3+\dots+x_{n-1}y_n-x_ny_{n-1}$$

wobei $\Phi_{\mathbf{B}}(\mathbf{a})=(x_1,\dots,x_n)^T$ *und* $\Phi_{\mathbf{B}}(\mathbf{b})=(y_1,\dots,y_n)^T$.

11.1.2 Orthogonalsysteme

Definition 11.10 *Sei* $(\mathbf{V}, \sigma)$ *ein Vektorraum mit Skalarprodukt und* $\mathbf{C}$ *eine Teilmenge von* $\mathbf{V}$*.* $\mathbf{C}$ *heißt* **Orthogonalsystem** *(OS), wenn*

1. $\langle \mathbf{c}, \mathbf{c} \rangle \neq 0$ *für alle* $\mathbf{c} \in \mathbf{C}$ *und*
2. $\langle \mathbf{c}_1, \mathbf{c}_2 \rangle = 0$ *für alle* $\mathbf{c}_1, \mathbf{c}_2 \in \mathbf{C}$ *mit* $\mathbf{c}_1 \neq \mathbf{c}_2$.

$\mathbf{C}$ *heißt* **Orthonormalsystem** *(ONS), wenn*

1. $\langle \mathbf{c}, \mathbf{c} \rangle = 1$ *für alle* $\mathbf{c} \in \mathbf{C}$ *und*
2. $\langle \mathbf{c}_1, \mathbf{c}_2 \rangle = 0$ *für alle* $\mathbf{c}_1, \mathbf{c}_2 \in \mathbf{C}$ *mit* $\mathbf{c}_1 \neq \mathbf{c}_2$.

Eine Basis $\mathbf{B}$ *von* $\mathbf{V}$ *heißt* **Orthogonalbasis** *(OB), wenn* $\mathbf{B}$ *ein Orthogonalsystem bildet.*
Eine Basis $\mathbf{B}$ *von* $\mathbf{V}$ *heißt* **Orthonormalbasis** *(ONB), wenn* $\mathbf{B}$ *ein Orthonormalsystem bildet.*

Satz 11.11 *Jedes Orthogonalsystem ist linear unabhängig.*

Satz 11.8 kann auch so interpretiert werden, daß es in einem endlichdimensonalen symmetrischen Raum immer eine Orthogonalbasis gibt. Diese Eigenschaft kann noch etwas verfeinert werden und führt zum **Orthogonalisierungverfahren von Gram-Schmidt**:

Satz 11.12 *Sei* $(\mathbf{V}, \sigma)$ *ein anisotroper* $(\zeta-)$*symmetrischer Vektorraum und* $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2 \ldots\}$ *ein endliches oder abzählbares System von linear unabhängigen Vektoren. Bildet man nun induktiv*

$$\begin{aligned} \mathbf{b}_1 &:= \mathbf{c}_1, \\ \mathbf{b}_{n+1} &:= \mathbf{c}_{n+1} - \sum_{k=1}^{n} \frac{\langle \mathbf{c}_{n+1}, \mathbf{b}_k \rangle}{\langle \mathbf{b}_k, \mathbf{b}_k \rangle} \mathbf{b}_k, \end{aligned}$$

so ist $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots\}$ *ein Orthogonalsystem mit*

$$[\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_k\}] = [\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k\}] =$$

für alle $k \geq 1$*. Insbesondere gilt*

$$[\mathbf{C}] = [\mathbf{B}].$$

Korollar 11.13 *Jeder anisotrope höchstens abzählbar dimensionale Vektorraum mit Skalarprodukt besitzt eine Orthogonalbasis.*

Korollar 11.14 *Jedes endliche Orthogonalsystem eines anisotropen höchstens abzählbar dimensionalen Vektorraums mit Skalarprodukt läßt sich zu einer Orthogonalbasis fortsetzen.*

11.1.3 Reziproke Basen

Definition 11.15 *Sei* $(\mathbf{V}, \sigma)$ *ein Vektorraum mit Skalarprodukt und* $\mathbf{B} = \{\mathbf{b}_i \mid i \in I\}$ *eine Basis von* $\mathbf{V}$. *Eine Basis* $\mathbf{C} = \{\mathbf{c}_i \mid i \in I\}$ *heißt zu* $\mathbf{B}$ **reziprok** *oder* **reziproke Basis**, *wenn für alle* $i \in I$

$$\langle \mathbf{b}_i, \mathbf{c}_j \rangle = \delta_{ij}$$

gilt.

Daß man (im Falle ihrer Existenz) von *der* reziproken Basis sprechen kann, sichert die folgende Eigenschaft.

Lemma 11.16 *Sei* $(\mathbf{V}, \sigma)$ *ein Vektorraum mit Skalarprodukt. Besitzt eine Basis* $\mathbf{B}$ *von* $\mathbf{V}$ *eine reziproke Basis* $\mathbf{C}$, *so ist diese eindeutig bestimmt. Weiters ist* $\mathbf{B}$ *dann die zu* $\mathbf{C}$ *reziproke Basis.*

Man schreibt für die reziproke Basis von $\mathbf{B} = \{\mathbf{b}_i \mid i \in I\}$ manchmal auch $\hat{\mathbf{B}} = \{\hat{\mathbf{b}}_i \mid i \in I\}$.

Mit Hilfe einer reziproken Basis kann man die Koeffizienten bezüglich einer Basis leicht angeben.

Satz 11.17 *Sei* $(\mathbf{V}, \sigma)$ *ein Vektorraum mit Skalarprodukt und* $\mathbf{B} = \{\mathbf{b}_i \mid i \in I\}$ *eine Basis von* $\mathbf{V}$ *und* $\hat{\mathbf{B}} = \{\hat{\mathbf{b}}_i \mid i \in I\}$ *die zu* $\mathbf{B}$ *reziproke Basis. Dann gilt für jeden Vektor* $\mathbf{a} \in \mathbf{V}$

$$\mathbf{a} = \sum_{i \in I} \langle \mathbf{a}, \hat{\mathbf{b}}_i \rangle \mathbf{b}_i.$$

Dieser Satz erinnert an eine ähnliche Eigenschaft der dualen Basis $\mathbf{B}^*$, wo anstelle von $\langle \mathbf{a}, \hat{\mathbf{b}}_i \rangle$ der Koeffizient $\mathbf{b}_i^*(\mathbf{a}) = \langle \mathbf{b}_i^*, \mathbf{a} \rangle$ zu setzen ist. Tatsächlich *realisiert* $\hat{\mathbf{b}}_i$ mit Hilfe des Skalarprodukts σ die Linearform $\mathbf{b}_i^*$. Dadurch erhält man auch ein Existenzkriterium für reziproke Basen.

Satz 11.18 *Sei* $(\mathbf{V}, \sigma)$ *ein Vektorraum mit Skalarprodukt. Weiters bezeichne* $f_\sigma : \mathbf{V} \to \mathbf{V}^*$ *jene injektive (semi–)lineare Abbildung, die durch*

$$\langle f_\sigma(\mathbf{b}), \mathbf{a} \rangle = \sigma(\mathbf{a}, \mathbf{b})$$

definiert ist. Dann existiert zu einer Basis $\mathbf{B} = \{\mathbf{b}_i \mid i \in I\}$ *genau dann eine reziproke Basis, wenn* $\mathbf{B}^* \subseteq f_\sigma(\mathbf{V})$ *gilt. In diesem Fall ist*

$$\hat{\mathbf{B}} = \{\hat{\mathbf{b}}_i = f_\sigma^{-1}(\mathbf{b}_i^*) \mid i \in I\}$$

die reziproke Basis.

Ist $\mathbf{V}$ unendlichdimensional, so ist f_σ nicht surjektiv. Es ist daher im allgemeinen nicht gesichert, daß es eine reziproke Basis gibt, und tatsächlich ist das nicht immer der Fall.

Trotzdem gibt es auch im Unendlichdimensionalen Fälle, wo es immer reziproke Basen gibt.

Satz 11.19 *Sei* $(\mathbf{V}, \sigma)$ *ein Vektorraum mit Skalarprodukt und* $\mathbf{B} = \{\mathbf{b}_i \mid i \in I\}$ *eine Orthogonalbasis. Dann ist*

$$\hat{\mathbf{B}} = \left\{ \hat{\mathbf{b}}_i = \frac{1}{\langle \mathbf{b}_i, \mathbf{b}_i \rangle} \mathbf{b}_i \mid i \in I \right\}$$

die zu $\mathbf{B}$ *reziproke Basis.*

Korollar 11.20 *Sei* $(\mathbf{V}, \sigma)$ *ein Vektorraum mit Skalarprodukt. Dann ist eine Basis* $\mathbf{B}$ *genau dann Orthonormalbasis, wenn sie selbstreziprok ist.*

11.1.4 Orthogonalprojektion

In einem Vektorraum mit Skalarprodukt $(\mathbf{V}, \sigma)$ bezeichnet $\perp = \perp_\sigma$ die zu σ gehörige Orthogonalitätsrelation.

Definition 11.21 *Sei $(\mathbf{V}, \sigma)$ ein Vektorraum mit Skalarprodukt. Gilt für einen Unterraum $\mathbf{U} \leq \mathbf{V}$*

$$\mathbf{U} \oplus \mathbf{U}^\perp = \mathbf{V},$$

so heißt $\mathbf{U}^\perp$ **orthogonales Komplement** *von $\mathbf{U}$.*

Jeder Vektor $\mathbf{a} \in \mathbf{V}$ kann dann eindeutig durch $\mathbf{a} = \mathbf{u} + (\mathbf{a} - \mathbf{u})$ dargestellt werden, wobei $\mathbf{u} \in \mathbf{U}$ und $\mathbf{u} \perp (\mathbf{a} - \mathbf{u})$.

Definition 11.22 *Besitzt ein Unterraum $\mathbf{U} \leq \mathbf{V}$ ein orthogonales Komplement $\mathbf{U}^\perp$, so heißt die Projektion*

$$p_\mathbf{U} : \mathbf{V} \to \mathbf{U}, \quad \mathbf{a} \mapsto \mathbf{u} \quad \textit{mit} \quad \mathbf{u} \perp (\mathbf{a} - \mathbf{u})$$

Orthogonalprojektion *auf $\mathbf{U}$.*

Die Existenz eines orthogonalen Komplements ist im Unendlichdimensionalen nicht immer gesichert. Eine Ausnahme bilden endlichdimensionale Unterräume $\mathbf{U}$.

Satz 11.23 *Sei $(\mathbf{V}, \sigma)$ ein Vektorraum mit Skalarprodukt und $\mathbf{U} \leq \mathbf{V}$ ein endlichdimensionaler, anisotroper Unterraum von $\mathbf{V}$, d.h. $\sigma|_\mathbf{U}$ ist nicht ausgeartet. Dann besitzt $\mathbf{U}$ ein orthogonales Komplement.*

Bezeichnet $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m\}$ eine Basis von $\mathbf{U}$ und $\hat{\mathbf{B}} = \{\hat{\mathbf{b}}_1, \hat{\mathbf{b}}_2, \ldots, \hat{\mathbf{b}}_m\}$ die reziproke Basis von $\mathbf{B}$, dann ist die Orthogonalprojektion $p_\mathbf{U} : \mathbf{V} \to \mathbf{U}$ durch

$$p_\mathbf{U}(\mathbf{a}) = \sum_{j=1}^{m} \langle \mathbf{a}, \hat{\mathbf{b}}_j \rangle \, \mathbf{b}_j$$

gegeben.

Man beachte, daß es in jedem endlichdimensionalen Vektorraum mit Skalarprodukt zu jeder Basis eine reziproke Basis gibt. Insbesondere kann man eine Orthogonalbasis $\mathbf{B}$ verwenden, so daß sich die reziproke Basis $\hat{\mathbf{B}}$ durch $\hat{\mathbf{b}}_j = (\langle \mathbf{b}_j, \mathbf{b}_j \rangle)^{-1} \mathbf{b}_j$ berechnen läßt. Die Orthogonalprojektion $p_\mathbf{U}$ ist dann durch

$$p_\mathbf{U}(\mathbf{a}) = \sum_{j=1}^{m} \frac{\langle \mathbf{a}, \mathbf{b}_j \rangle}{\langle \mathbf{b}_j, \mathbf{b}_j \rangle} \, \mathbf{b}_j$$

gegeben.

Leider kann man diese Darstellung von $p_\mathbf{U}$ nicht allgemein im unendlichdimensionalen Fall verwenden, auch wenn man eine Orthogonalbasis $\mathbf{B} = \{\mathbf{b}_j \mid j \in J\}$ von $\mathbf{U}$ kennt. Es kann nämlich passieren, daß $\langle \mathbf{a}, \mathbf{b}_j \rangle$ für unendlich viele $j \in J$ ungleich Null ist. Allerdings wird der formale Ausdruck

$$\sum_{j \in J} \frac{\langle \mathbf{a}, \mathbf{b}_j \rangle}{\langle \mathbf{b}_j, \mathbf{b}_j \rangle} \mathbf{b}_j$$

als **Fourierreihe** bezeichnet. Nur in ganz bestimmten Fällen, etwa in Hilberträumen, kann dieser formalen Summe wieder ein Vektor in $\mathbf{V}$ zugeordnet werden.

11.2 Euklidische und unitäre Vektorräume

11.2.1 Positive definite Skalarprodukte

Definition 11.24 *Ein Vektorraum* $\mathbf{V}$ *über den reellen Zahlen* $\mathbb{R}$ *mit einer positiv definiten symmetrischen Bilinearform* σ *heißt* **euklidischer Vektorraum**.

Ein Vektorraum $\mathbf{V}$ *über den komplexen Zahlen* $\mathbb{C}$ *mit einer positiv definiten Hermiteschen Form* σ *heißt* **unitärer Vektorraum**.

Lemma 11.25 *Jeder euklidische bzw. unitäre Vektorraum ist anisotrop.*

Ist $\mathbf{V}$ ein endlichdimensionaler euklidischer bzw. unitärer Vektorraum, so gibt es jedenfalls eine Basis $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ von $\mathbf{V}$ mit

$$\Phi_{\mathbf{B}}(\sigma) = E_n,$$

d.h. für $\Phi_{\mathbf{B}}(\mathbf{a}) = (x_1, \ldots, x_n)^T$ und $\Phi_{\mathbf{B}}(\mathbf{b}) = (y_1, \ldots, y_n)^T$ gilt

$$\langle \mathbf{a}, \mathbf{b} \rangle = x_1 y_1 + \cdots + x_n y_n$$

bzw.

$$\langle \mathbf{a}, \mathbf{b} \rangle = x_1 \overline{y}_1 + \cdots + x_n \overline{y}_n.$$

Außerdem ist $\mathbf{B}$ eine Orthonormalbasis.

In euklidischen bzw. unitären Vektorräumen kann man übrigens immer von einer Orthogonalbasis zu einer Orthonormalbasis übergehen. Dazu benötigt man den Begriff der Länge eines Vektors.

Definition 11.26 *Sei* $(\mathbf{V}, \sigma)$ *ein euklidischer bzw. unitärer Vektorraum. Dann bezeichnet*

$$\|\mathbf{a}\| := \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle}$$

die **Länge** *bzw. die* **Norm** *des Vektors* $\mathbf{a} \in \mathbf{V}$.

Ein Vektor $\mathbf{a} \in \mathbf{V}$ *heißt* **normiert**, *wenn*

$$\|\mathbf{a}\| = 1.$$

Satz 11.27 *Sei* $(\mathbf{V}, \sigma)$ *ein euklidischer bzw. unitärer Vektorraum und* $\mathbf{B} = \{\mathbf{b}_i \mid i \in I\}$ *eine Orthogonalbasis von* $\mathbf{V}$. *Dann ist*

$$\mathbf{B}' := \left\{ \mathbf{b}'_i := \frac{1}{\|\mathbf{b}_i\|} \mathbf{b}_i \mid i \in I \right\}$$

eine Orthonormalbasis von $\mathbf{V}$.

11.2.2 Normierte Räume

Grundlegend für viele Betrachtungen in euklidischen bzw. unitären Vektorräumen ist die **Ungleichung von Cauchy–Schwarz**.

Satz 11.28 *Für je zwei Vektoren* $\mathbf{a}, \mathbf{b}$ *eines euklidischen bzw. unitären Vektorraums* $(\mathbf{V}, \sigma)$ *gilt*

$$|\langle \mathbf{a}, \mathbf{b} \rangle| \leq \|\mathbf{a}\| \, \|\mathbf{b}\|.$$

Weiters gilt Gleichheit genau dann, wenn $\mathbf{a}, \mathbf{b}$ *linear abhängig sind.*

Beispielsweise gilt für komplexe Zahlen $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n \in \mathbb{C}$

$$\left| \sum_{i=1}^{n} x_i \overline{y}_i \right| \leq \left(\sum_{i=1}^{n} |x_i|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} |y_i|^2 \right)^{\frac{1}{2}}$$

und für (stetige) Funktionen $f, g : [0, 1] \to \mathbb{C}$

$$\left| \int_0^1 f(x) \overline{g}(x) \, dx \right| \leq \sqrt{\int_0^1 |f(x)|^2 dx} \sqrt{\int_0^1 |g(x)|^2 dx}.$$

Die Cauchy–Schwarzsche Ungleichung erlaubt es, in euklidischen bzw. unitären Vektorräumen den Winkel zwischen zwei Vektoren zu definieren.

Definition 11.29 *Sei* $(\mathbf{V}, \sigma)$ *ein euklidischer bzw. ein unitärer Vektorraum. Der Winkel* $\alpha \in [0, \pi]$ *zweier Vektoren* $\mathbf{a}, \mathbf{b} \in \mathbf{V} \setminus \{\mathbf{0}\}$ *ist durch*

$$\cos \alpha = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|}$$

definiert.

Zwei orthogonale Vektoren haben also Winkel $\alpha = \frac{\pi}{2}$.

Eine unmittelbare Folgerung der Ungleichung von Cauchy–Schwarz ist die **Dreiecksungleichung**.

Satz 11.30 *Für je zwei Vektoren* $\mathbf{a}, \mathbf{b}$ *eines euklidischen bzw. unitären Vektorraums* $(\mathbf{V}, \sigma)$ *gilt*

$$\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|.$$

Weiters gilt Gleichheit genau dann, wenn $\mathbf{a} = \mathbf{0}$ *oder* $\mathbf{b} = \mathbf{0}$ *oder wenn es eine positive reelle Zahl* t *mit* $\mathbf{a} = t\mathbf{b}$ *gibt.*

Die Dreiecksungleichung ist die wichtigste Eigenschaft eines *normierten Raums*, der folgendermaßen definiert ist.

Definition 11.31 *Ein Vektorraum* $\mathbf{V}$ *über* $K = \mathbb{R}$ *oder* $K = \mathbb{C}$ *heißt* **normierter Raum**, *wenn es eine Abbildung* $\| \cdot \| : \mathbf{V} \to \mathbb{R}$ *mit den folgenden Eigenschaften gibt:*

1. $\forall \mathbf{a} \in \mathbf{V}: \ \|\mathbf{a}\| \geq 0.$

2. $\|\mathbf{a}\| = 0 \Longleftrightarrow \mathbf{a} = 0.$

3. $\forall \mathbf{a} \in \mathbf{V}\ \forall x \in K: \ \|x\mathbf{a}\| = |x|\|\mathbf{a}\|.$

4. $\forall \mathbf{a}, \mathbf{b} \in \mathbf{V}: \ \|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|.$

Demnach ist jeder euklidische bzw. unitäre Vektorraum mit $\|\mathbf{a}\| = \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle}$ ein normierter Raum. Andere Normen im $\mathbb{R}^{n\times 1}$ bzw. $\mathbb{C}^{n\times 1}$ sind z.B. die sogenannten p-Normen

$$\|(x_1, x_2, \ldots, x_n)^T\|_p := \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}$$

mit $p \geq 1$ und die ∞–Norm

$$\|(x_1, x_2, \ldots, x_n)^T\|_\infty := \max_{1\leq i\leq n} |x_i|.$$

In normierten Räumen wird durch

$$d(\mathbf{a}, \mathbf{b}) := \|\mathbf{a} - \mathbf{b}\|$$

ein Abstandsbegriff eingeführt. Die Abbildung $d : \mathbf{V} \times \mathbf{V} \to \mathbb{R}$ erfüllt die Eigenschaften eines *metrischen Raums.*

Definition 11.32 *Sei X eine nichtleere Menge. Eine Abbildung $d : X \times X \to \mathbb{R}$ heißt* **Metrik** *auf X, wenn folgende Eigenschaften erfüllt sind:*

- $\forall a, b \in X: \ d(a, b) \geq 0.$
- $d(a, b) = 0 \Longleftrightarrow a = b.$
- $\forall a, b \in X: \ d(a, b) = d(b, a).$
- $\forall a, b, c \in X: \ d(a, c) \leq d(a, b) + d(b, c).$

Das Paar (X, d) heißt **metrischer Raum***, wenn $d : X \times X \to \mathbb{R}$ eine Metrik ist.*

Demnach ist jeder normierte Raum in kanonischer Weise ein metrischer Raum.

In metrischen Räumen, also insbesonder in normierten Räumen, kann wie auf den reellen Zahlen ein Konvergenzbegriff eingeführt werden. Das besondere an den reellen Zahlen ist, daß jede Cauchyfolge konvergent ist. Diese Eigenschaft gilt nicht in allen normierten Räumen. Normierte Räume, die diese Eigenschaft haben, nämlich daß jede Cauchyfolge konvergent ist, heißen **Banachräume**. Ein euklidischer bzw. unitärer Vektorraum, der Banachraum ist, heißt **Hilbertraum**. Insbesondere ist es in Hilberträumen möglich, allgemeine Fourierreihen zu betrachten.

11.2.3 Fourierreihen

In euklidischen bzw. unitären Vektorräumen gibt es eine verallgemeinerte Version des **Satzes von Pythagoras**.

Satz 11.33 *Sei* $(\mathbf{V}, \sigma)$ *ein euklidischer bzw. unitärer Vektorraum. Dann gilt für orthogonale Vektoren* $\mathbf{a}, \mathbf{b} \in \mathbf{V}$,

$$\|\mathbf{a} + \mathbf{b}\|^2 = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2.$$

Aus dieser Eigenschaft folgt, daß die Orthogonalprojektion ein Approximationsproblem löst.

Satz 11.34 *Sei* $(\mathbf{V}, \sigma)$ *ein euklidischer bzw. unitärer Vektorraum und* $\mathbf{U} \leq \mathbf{V}$ *ein Unterraum, zu dem es ein orthogonales Komplement* $\mathbf{U}^\perp$ *gibt. Dann gilt für jeden Vektor* $\mathbf{a} \in \mathbf{V}$

$$\|\mathbf{a} - p_{\mathbf{U}}(\mathbf{a})\| = \min_{\mathbf{u} \in \mathbf{U}} \|\mathbf{a} - \mathbf{u}\|$$

und

$$\|p_{\mathbf{U}}(\mathbf{a})\| \leq \|\mathbf{a}\|.$$

$p_{\mathbf{U}}(\mathbf{a})$ ist also jener Vektor aus $\mathbf{U}$, der $\mathbf{a}$ am besten von allen Vektoren aus $\mathbf{U}$ approximiert.

Diese Eigenschaft kann beispielsweise für die Polynomapproximation verwendet werden.

Beispiel 11.35 Sei $\mathbf{V}$ der Vektorraum aller stetigen Funktionen $f : [0,1] \to \mathbb{R}$. Mit dem Skalarprodukt

$$\langle f, g \rangle := \int_0^1 f(x) g(x)\, dx$$

wird $\mathbf{V}$ zu einem euklidischen Vektorraum.

Die Polynome $f_j(x) = x^j$, $0 \leq j \leq m$, spannen einen $(m+1)$-dimensionalen Unterraum $\mathbf{V}_m \leq \mathbf{V}$ auf. Nach dem Orthogonalisierungverfahren von Gram–Schmidt gibt es auch eine Orthonormalbasis $\{p_j(x) \,|\, 0 \leq j \leq m\}$, von $\mathbf{V}_m$, wobei $p_j(x)$ Polynom von Grad j ist, und es gilt

$$\int_0^1 p_i(x) p_j(x)\, dx = \delta_{ij} \qquad (0 \leq i, j \leq m).$$

Ist nun $f \in \mathbf{V}$ eine beliebige stetige Funktion, so ist

$$p(x) := \sum_{j=0}^{m} \langle f, p_j \rangle\, p_j(x)$$

jenes Polynom vom Grad $\leq m$, für das das Integral

$$\int_0^1 (f(x) - p(x))^2 dx$$

am kleinsten ist.

Beispiel 11.36 Sei $\mathbf{V}$ der Vektorraum aller stetigen Funktionen $f : [-\pi, \pi] \to \mathbb{C}$. Mit dem Skalarprodukt

$$\langle f, g \rangle := \int_{-\pi}^{\pi} f(x)\overline{g}(x)\, dx$$

wird $\mathbf{V}$ zu einem unitären Vektorraum.

Die Funktionen

$$t_j(x) := \frac{1}{\sqrt{2\pi}} e^{ijx} = \frac{1}{\sqrt{2\pi}} (\cos(jx) + i \sin(jx)) \qquad (-m \leq j \leq m)$$

bilden ein Orthonormalsystem in $\mathbf{V}$. Der von den Funktionen $t_j(x)$, $-m \leq j \leq m$, aufgespannte Unterraum von $\mathbf{V}$ wird auch als Raum der trigonometrischen Polynome von Grad $\leq m$ bezeichnet. Wie im vorigen Beispiel ist das trigonometrische Polynom

$$t(x) = \sum_{j=-m}^{m} \langle f, t_j \rangle\, t_j(x)$$

jedes trigonometrische Polynom vom Grad $\leq m$, für das das Integral

$$\int_{-\pi}^{\pi} |f(x) - t(x)|^2 dx$$

kleinstmöglich ist.

Wie schon erwähnt, nennt man bei einem Orthogonalsystem $\{\mathbf{b}_i \,|\, i \in I\}$ die formale Summe

$$\sum_{i \in I} \frac{\langle \mathbf{a}, \mathbf{b}_i \rangle}{\langle \mathbf{b}_i, \mathbf{b}_i \rangle}\, \mathbf{c}_i$$

Fourierreihe von $\mathbf{a}$. Jede endliche Teilsumme dieser *Reihe* entspricht einer Orthogonalprojektion auf einen endlichdimensionalen Teilraum, und je mehr Summanden man aufnimmt desto besser wird die *Approximation*. Daß man (im geeigneten Sinn) wirklich von einer Reihe sprechen kann, sichert die folgende Eigenschaft.

Satz 11.37 *Sei $\{\mathbf{b}_i \,|\, i \in I\}$ ein Orthogonalsystem eines euklidischen bzw. unitären Vektorraums $\mathbf{V}$. Dann ist für jeden Vektor $\mathbf{a} \in \mathbf{V}$ die Menge*

$$\{i \in I \,|\, \langle \mathbf{a}, \mathbf{b}_i \rangle \neq 0\}$$

höchstens abzählbar.

11.3 Adjungierte und metrische Abbildungen

11.3.1 Adjungierte Abbildungen

Definition 11.38 *Seien $(\mathbf{V}, \sigma)$ und $(\mathbf{W}, \tau)$ zwei symmetrische bzw. ζ–symmetrische Vektorräume über demselben Körper K (wobei ζ einen Körperautomorphismus von K bezeichnet)* *und* $f \in L(\mathbf{V}, \mathbf{W})$. *Eine lineare Abbildung* $g \in L(\mathbf{W}, \mathbf{V})$ *heißt zu* f **adjungierte Abbildung**, *wenn für alle Vektoren* $\mathbf{a} \in \mathbf{V}$, $\mathbf{c} \in \mathbf{W}$

$$\langle f(\mathbf{a}), \mathbf{c}\rangle_\tau = \langle \mathbf{a}, g(\mathbf{c})\rangle_\sigma$$

gilt.

Lemma 11.39 *Existiert zu einer linearen Abbildung* $f \in L(\mathbf{V}, \mathbf{W})$ *eine adjungierte Abbildung* $g \in L(\mathbf{W}, \mathbf{V})$, *so ist diese eindeutig bestimmt. Weiters ist dann* f *die zu* g *adjungierte Abbildung.*

Wegen der Eindeutigkeit bezeichnet man (im Falle seiner Existenz) die adjungierte Abbildung von f auch durch $\hat{f}$.

Die Adjungierte Abbildung $\hat{f}$ hat formale Ähnlichkeit mit der *adjungierten* bzw. *transponierten* Abbildung $f^* : \mathbf{W}^* \to \mathbf{V}^*$, die durch

$$\langle \mathbf{c}^*, f(\mathbf{a})\rangle = \langle f^*(\mathbf{c}^*), \mathbf{a}\rangle$$

definert werden kann. Tatsächlich besteht ein Zusammenhang dieser beiden Begriffe mittels der (ζ–semi–)linearen injektiven Abbildungen $f_\sigma : \mathbf{V} \to \mathbf{V}^*$ und $f_\tau : \mathbf{W} \to \mathbf{W}^*$.

Satz 11.40 *Seien* $(\mathbf{V}, \sigma)$ *und* $(\mathbf{W}, \tau)$ *zwei symmetrische bzw.* ζ*–symmetrische Vektorräume über demselben Körper* K *(wobei* ζ *einen Körperautomorphismus von* K *bezeichnet) und* $f \in L(\mathbf{V}, \mathbf{W})$. *Dann existiert die adjungierte Abbildung* $\hat{f} \in L(\mathbf{W}, \mathbf{V})$ *genau dann, wenn*

$$(f^* \circ f_\tau)(\mathbf{W}) \subseteq f_\sigma(\mathbf{V}).$$

Weiters gilt in diesem Fall

$$\hat{f} = f_\sigma^{-1} \circ f^* \circ f_\tau.$$

Korollar 11.41 *Ist* $\dim(\mathbf{V}) < \infty$, *so existiert für jede lineare Abbildung* $f \in L(\mathbf{V}, \mathbf{W})$ *die adjungierte Abbildung* $\hat{f} \in L(\mathbf{W}, \mathbf{V})$.

Weitere Eigenschaften der adjungierten Abbildung sind im folgenden Satz aufgelistet.

Satz 11.42 *Seien* $(\mathbf{V}, \sigma)$ *und* $(\mathbf{W}, \tau)$ *zwei symmetrische bzw.* ζ*–symmetrische Vektorräume über demselben Körper* K *(wobei* ζ *einen Körperautomorphismus von* K *bezeichnet) und es existiere zu* $f \in L(\mathbf{V}, \mathbf{W})$ *die adjungierte Abbildung* $\hat{f} \in L(\mathbf{W}, \mathbf{V})$. *Dann gelten die folgenden Eigenschaften:*

- $f(\mathbf{V})^\perp = \mathrm{kern}(\hat{f})$ *resp.* $\hat{f}(\mathbf{W})^\perp = \mathrm{kern}(f)$.
- *Ist* f *surjektiv, so ist* $\hat{f}$ *injektiv.*
- *Sei zusätzlich* $\dim(\mathbf{V}) < \infty$. *Ist dann* f *injektiv, so ist* $\hat{f}$ *surjektiv.*
- *Ist* $\dim(\mathbf{V}) < \infty$ *oder* $\dim(\mathbf{W}) < \infty$, *dann gilt* $\mathrm{rg}(f) = \mathrm{rg}(\hat{f})$.

Schließlich besteht im endlichdimensionalen Fall ein Zusammenhang zwischen den Koordinatenmatrizen von f und $\hat{f}$.

Satz 11.43 *Seien* $(\mathbf{V}, \sigma)$ *und* $(\mathbf{W}, \tau)$ *zwei symmetrische bzw.* ζ*–symmetrische Vektorräume über demselben Körper* K *(wobei* ζ *einen Körperautomorphismus von* K *bezeichnet) mit* $\dim(\mathbf{V}) = n$ *und* $\dim(\mathbf{W}) = m$*. Weiters sei* $\mathbf{B}$ *eine Basis von* $\mathbf{V}$ *und* $\mathbf{C}$ *eine Basis von* $\mathbf{W}$*. Dann gilt für adjungierte Abbildungspaare* $f \in L(\mathbf{V}, \mathbf{W})$*,* $\hat{f} \in L(\mathbf{W}, \mathbf{V})$

$$\Phi_{\hat{\mathbf{C}}\hat{\mathbf{B}}}(\hat{f}) = \zeta\left(\Phi_{\mathbf{BC}}(f)\right)^T .$$

Insbesondere gilt für Orthonormalbasen $\mathbf{B}, \mathbf{C}$

$$\Phi_{\mathbf{CB}}(\hat{f}) = \zeta\left(\Phi_{\mathbf{BC}}(f)\right)^T .$$

11.3.2 Isometrische Abbildungen

Definition 11.44 *Seien* $(\mathbf{V}, \sigma)$ *und* $(\mathbf{W}, \tau)$ *zwei symmetrische bzw.* ζ*–symmetrische Vektorräume über demselben Körper* K *(wobei* ζ *einen Körperautomorphismus von* K *bezeichnet). Eine Abbildung* $f \in L(\mathbf{V}, \mathbf{W})$ *heißt* **isometrische Abbildung** *oder* **Isometrie***, wenn für alle Vektoren* $\mathbf{a}, \mathbf{b} \in \mathbf{V}$

$$\langle f(\mathbf{a}), f(\mathbf{b})\rangle_\tau = \langle \mathbf{a}, \mathbf{b}\rangle_\sigma$$

gilt.

Der Betriff *Isometrie* stammt daher, daß im euklidischen bzw. unitären Fall eine Isometrie durch die Längentreue charakterisiert werden kann.

Satz 11.45 *Seien* $(\mathbf{V}, \sigma)$ *und* $(\mathbf{W}, \tau)$ *zwei euklidische bzw. unitäre Vektorräume. Dann ist eine Abbildung* $f \in L(\mathbf{V}, \mathbf{W})$ *genau dann eine Isometrie, wenn für alle* $\mathbf{a} \in \mathbf{V}$

$$\|f(\mathbf{a})\| = \|\mathbf{a}\|$$

gilt.

Es ist interessant, daß die Linearität in der Definition von isometrischen Abbildungen im surjetkiven Fall nicht gefordert werden muß.

Satz 11.46 *Seien* $(\mathbf{V}, \sigma)$ *und* $(\mathbf{W}, \tau)$ *zwei symmetrische bzw.* ζ*–symmetrische Vektorräume über demselben Körper* K *(wobei* ζ *einen Körperautomorphismus von* K *bezeichnet). Dann ist eine surjektive Abbildung* $f : \mathbf{V} \to \mathbf{W}$ *mit der Eigenschaft, daß für alle Vektoren* $\mathbf{a}, \mathbf{b} \in \mathbf{V}$

$$\langle f(\mathbf{a}), f(\mathbf{b})\rangle_\tau = \langle \mathbf{a}, \mathbf{b}\rangle_\sigma$$

gilt, eine isometrische Abbildung.

Die Injektivität muß nicht gefordert werden.

Satz 11.47 *Jede isometrische Abbildung* $f : \mathbf{V} \to \mathbf{W}$ *zwischen symmetrischen bzw.* ζ*–symmetrischen Vektorräumen* $(\mathbf{V}, \sigma)$ *und* $(\mathbf{W}, \tau)$ *über demselben Körper* K *ist injektiv.*

Für lineare Abbildungen kann die Isometriebedingung leicht überprüft werden.

Satz 11.48 *Seien* $(\mathbf{V}, \sigma)$ *und* $(\mathbf{W}, \tau)$ *zwei symmetrische bzw.* ζ*–symmetrische Vektorräume über demselben Körper* K *(wobei* ζ *einen Körperautomorphismus von* K *bezeichnet) und bezeichne* $\mathbf{B} = \{\mathbf{b}_i \,|\, i \in I\}$ *eine Basis von* $\mathbf{V}$*. Dann ist eine lineare Abbildung* $f \in L(\mathbf{V}, \mathbf{W})$ *genau dann eine Isometrie wenn*

$$\forall i, j \in I: \ \langle f(\mathbf{b}_i), f(\mathbf{b}_j)\rangle_\tau = \langle \mathbf{b}_i, \mathbf{b}_j\rangle_\sigma$$

gilt.

Insbesondere werden Orthonormalsysteme auf Orthonormalsysteme abgebildet.

Für bijektive Abbildungen besteht ein Zusammenhang zwischen Isometrien und adjungierten Abbildungen.

Satz 11.49 *Seien* $(\mathbf{V}, \sigma)$ *und* $(\mathbf{W}, \tau)$ *zwei symmetrische bzw.* ζ*–symmetrische Vektorräume über demselben Körper* K *(wobei* ζ *einen Körperautomorphismus von* K *bezeichnet). Dann ist ein Vektorraumisomorphismus* $f \in L(\mathbf{V}, \mathbf{W})$ *genau dann eine Isometrie, wenn die inverse Abbildung* $f^{-1} \in L(\mathbf{W}, \mathbf{V})$ *die zu* f *adjungierte Abbildung ist.*

Korollar 11.50 *Sind die Vektorräume* $\mathbf{V}$*,* $\mathbf{W}$ *aus Satz 11.49 zusätzlich endlichdimensional und sind* $\mathbf{B}, \mathbf{C}$ *Orthonormalbasen von* $\mathbf{V}$ *bzw.* $\mathbf{W}$*, so gilt für jede bijektive Isometrie* $f \in L(\mathbf{V}, \mathbf{W})$

$$\Phi_{\mathbf{BC}}(f)^{-1} = \zeta(\Phi_{\mathbf{BC}}^T).$$

11.3.3 Orthogonale Matrizen

Definition 11.51 *Eine reguläre Matrix* $A \in K^{n\times n}$ *heißt* **orthogonal** *bzw.* ζ–**orthogonal** *(wobei* ζ *einen Körperautomorphismus von* K *mit* $\zeta^2 = \mathrm{id}_K$ *bezeichnet), wenn*

$$A^{-1} = \zeta(A^T)$$

ist.

Im unitären Fall nennt man eine Matrix $A \in \mathbb{C}^{n\times n}$ mit $A^{-1} = A^H = \overline{A}^T$ **unitäre Matrix**.

Insbesondere sind die Koordinatenmatrizen von isometrischen Abbildungen bezüglich Orthonormalbasen und Basistransformationsmatrizen zwischen zwei Othonormalbasen ζ-orthogonal.

Satz 11.52 *Sei* $\zeta : K \to K$ *ein Körperautomorphismus mit* $\zeta^2 = \mathrm{id}_K$*, und im Vektorraum* $\mathbf{V} = K^{n\times 1}$ *bezeichne*

$$\langle \mathbf{a}, \mathbf{b}\rangle := \mathbf{a}^T\zeta(\mathbf{b})$$

das gewöhnliche Skalarprodukt.
Dann sind für eine Matrix $A \in K^{n\times n}$ *folgende drei Bedingungen äquivalent.*

- A *ist* ζ*–orthogonal.*
- *Die Spaltenvektoren von* A *bilden eine Orthonormalbasis von* $K^{n\times 1}$*.*
- *Die Zeilenvektoren von* A *bilden eine Orthonormalbasis von* $K^{n\times 1}$*.*

Definition 11.53 *Sei $\zeta : K \to K$ ein Körperautomorphismus mit $\zeta^2 = \mathrm{id}_K$. Zwei Matrizen $A, B \in K^{n\times n}$ heißen ζ–***orthogonal–ähnlich***, wenn es eine ζ–orthogonale Matrix $P \in K^{n\times n}$ mit*

$$B = P^{-1}AP$$

gibt.

Wegen $P^{-1} = \zeta(P^T)$ sind in diesem Fall die Matrizen A, B auch ζ-kongruent.

Insbesondere sind Koordinatenmatrizen isometrisch–konjugierter Abbildungen ζ–orthogonal–ähnlich:

Definition 11.54 *Sei $(\mathbf{V}, \sigma)$ ein symmetrischer bzw. ζ–symmetrischer Vektorraum über einem Körper K (wobei ζ einen Körperautomorphismus von K bezeichnet). Zwei lineare Abbildungen $f_1, f_2 \in L(\mathbf{V}, \mathbf{V})$ heißen ζ–***orthogonal–ähnlich***, wenn es eine bijektive Isometrie $g \in L(\mathbf{V}, \mathbf{V})$ mit*

$$f_2 = g^{-1} \circ f_1 \circ g$$

gibt.

11.4 Normale Abbildungen

11.4.1 Spektralsatz

Definition 11.55 *Sei $(\mathbf{V}, \sigma)$ ein symmetrischer bzw. ζ–symmetrischer Vektorraum über einem Körper K (wobei ζ einen Körperautomorphismus von K bezeichnet). Eine lineare Abbildung $f \in L(\mathbf{V}, \mathbf{V})$ heißt* **normal***, wenn die adjungierte Abbildung $\hat{f} \in L(\mathbf{V}, \mathbf{V})$ existiert und f mit $\hat{f}$ kommutiert, d.h.*

$$f \circ \hat{f} = \hat{f} \circ f.$$

Definition 11.56 *Sei $\zeta : K \to K$ ein Körperautomorphismus mit $\zeta^2 = \mathrm{id}_K$. Eine Matrix $A \in K^{n\times n}$ heißt ζ–***normal***, wenn*

$$A\,\zeta(A^T) = \zeta(A^T)\,A$$

gilt.

Offensichtlich sind Koordinatenmatrizen normaler Abbildungen bezüglich Orthogonalbasen ζ-normal.

Normale Abbildungen sind im wesentlichen jene, wo es eine Orthogonalbasis aus Eigenvektoren gibt. Zunächst gilt allgemein folgende Eigenschaft.

Satz 11.57 *Sei $(\mathbf{V}, \sigma)$ ein symmetrischer bzw. ζ–symmetrischer Vektorraum über einem Körper K (wobei ζ einen Körperautomorphismus von K bezeichnet) und $f \in L(\mathbf{V}, \mathbf{V})$. Gibt es eine Orthogonalbasis $\mathbf{B} = \{\mathbf{b}_i \,|\, i \in I\}$ von $\mathbf{V}$, die nur aus Eigenvektoren von f besteht, dann ist f normal.*

Bezeichnet t_i den zu $\mathbf{b}_i$ gehörigen Eigenwert, d.h. $f(\mathbf{b}_i) = t\mathbf{b}_i$, so ist die adjungierte Abbildung $\hat{f}$ durch

$$\hat{f}(\mathbf{b}_i) = \zeta(t_i)\mathbf{b}_i$$

gegeben.

Die Umkehrung gilt unter natürlichen Voraussetzungen im Endlichdimensionalen und wird Inhalt des Spektralsatzes sein.

Lemma 11.58 *Für normale Abbildungen $f \in L(\mathbf{V}, \mathbf{V})$ gilt*

$$\langle f(\mathbf{a}), f(\mathbf{b})\rangle = \langle \hat{f}(\mathbf{a}), \hat{f}(\mathbf{b})\rangle \qquad (\mathbf{a}, \mathbf{b} \in \mathbf{V}).$$

Lemma 11.59 *Sei $(\mathbf{V}, \sigma)$ ein anisotroper ζ–symmetrischer Vektorraum. Ist t ein Eigenwert einer normalen Abbildung $f \in L(\mathbf{V}, \mathbf{V})$ mit Eigenraum $\mathbf{V}_t$, so ist $\zeta(t)$ Eigenwert der adjungierten Abbildung $\hat{f} \in L(\mathbf{V}, \mathbf{V})$ mit demselben Eigenraum.*

Lemma 11.60 *Sei $(\mathbf{V}, \sigma)$ ein anisotroper ζ–symmetrischer Vektorraum und $f \in L(\mathbf{V}, \mathbf{V})$ eine normale Abbildung. Sind $\mathbf{a}, \mathbf{b}$ zwei Eigenvektoren von f mit verschiedenen Eigenwerten $t_1 \neq t_2$, dann sind $\mathbf{a}, \mathbf{b}$ orthogonal.*

Lemma 11.61 *Sei $(\mathbf{V}, \sigma)$ ein anisotroper ζ–symmetrischer Vektorraum und $f \in L(\mathbf{V}, \mathbf{V})$ eine normale Abbildung. Ist $\mathbf{U}$ ein f- und $\hat{f}$-invarianter Unterraum von $\mathbf{V}$, so ist $\mathbf{U}^{\perp}$ ebenfalls f- und $\hat{f}$-invariant.*

Der **Spektralsatz** für normale Abbildungen lautet nun:

Satz 11.62 *Sei $(\mathbf{V}, \sigma)$ ein anisotroper, endlichdimensionaler, ζ–symmetrischer Vektorraum über dem Körper K (wobei ζ ein Körperautomorphismus auf K bezeichnet) und $f \in L(\mathbf{V}, \mathbf{V})$ eine normale Abbildung, deren charakteristisches Polynom $\chi_f(x)$ über K in Linearfaktoren zerfällt. Dann gibt es eine orthogonale Basis $\mathbf{B}$ von $\mathbf{V}$, die nur aus Eigenvektoren von f besteht, d.h.*

$$\Phi_{\mathbf{BB}}(f) = \mathrm{diag}(t_1, t_2, \ldots, t_n),$$

wobei $t_1, t_2, \ldots, t_n$ die Eigenwerte von f sind.

Korollar 11.63 *Eine ζ–normale Matrix $A \in K^{n\times n}$, deren charakteristisches Polynom $\chi_A(x)$ über K in Linearfaktoren zerfällt, ist diagonalisierbar, insbesondere gibt es eine orthogonale Basis aus Eigenwerten von A bzw. es gibt eine ζ-orthogonale Matrix $P \in K^{n\times n}$ (d.h. $P^{-1} = \zeta(P^T)$), so daß $P^{-1}AP$ Diagonalmatrix ist.*

11.4.2 Normale Abbildungen in unitären Vektorräumen

Der Spektralsatz kann nun sofort auf unitäre Vektorräume angewandt werden.

Satz 11.64 *Sei $\mathbf{V}$ ein endlichdimensionaler unitärer Vektorraum und $f \in L(\mathbf{V}, \mathbf{V})$. Dann gelten die folgenden Charakterisierungen:*

1. *f ist genau dann normal, wenn es eine Orthonormalbasis $\mathbf{B}$ von $\mathbf{V}$ gibt, die nur aus Eigenvektoren von f besteht.*
2. *f ist genau dann selbstadjungiert, d.h. $\hat{f} = f$, wenn es eine Orthonormalbasis $\mathbf{B}$ von $\mathbf{V}$ gibt, die nur aus Eigenvektoren von f mit reellen Eigenwerten besteht.*
3. *f ist genau dann antiselbstadjungiert, d.h. $\hat{f} = -f$, wenn es eine Orthonormalbasis $\mathbf{B}$ von $\mathbf{V}$ gibt, die nur aus Eigenvektoren von f mit rein imaginären Eigenwerten besteht.*
4. *f ist genau dann eine Isometrie, wenn es eine Orthonormalbasis $\mathbf{B}$ von $\mathbf{V}$ gibt, die nur aus Eigenvektoren von f mit Eigenwerten t vom Betrag $|t| = 1$ besteht.*

Der entsprechende Satz für Matrizen lautet folgendermaßen.

Satz 11.65 *Sei $A \in \mathbb{C}^{n\times n}$ eine Matrix. Dann gelten die folgenden Charakterisierungen.*

1. *A ist genau dann normal, d.h. $AA^H = A^H A$, wenn A unitär–ähnlich zu einer Diagonalmatrix ist, d.h. es gibt eine unitäre Matrix $P \in \mathbb{C}^{n\times n}$ mit*
$$P^{-1}AP = P^H AP = \operatorname{diag}(t_1, t_2, \ldots, t_n),$$
wobei $t_1, t_2, \ldots, t_n \in \mathbb{C}$ die Eigenwerte von A bezeichnen.
2. *A ist genau dann Hermitesch, d.h. $A = A^H$, wenn A unitär–ähnlich zu einer rellen Diagonalmatrix ist.*
3. *A ist genau dann anti-Hermitesch, d.h. $A = -A^H$, wenn A unitär–ähnlich zu einer rein imaginären Diagonalmatrix ist.*
4. *A ist genau dann unitär, d.h. $A^{-1} = A^H$, wenn A unitär–ähnlich zu einer Diagonalmatrix ist, deren Diagonalelemente alle Betrag 1 haben.*

11.4.3 Normale Abbildungen in euklidischen Vektorräumen

Im euklidischen Fall ist die Situation ein wenig komplizierter, da bei allgemeinen normalen Matrizen die Eigenwerte nicht unbedingt reell sein müssen, allerdings kann ein euklidischer Raum $(\mathbf{V}, \sigma)$ sofort in einen unitären Raum $(\mathbf{V}_{\mathbb{C}}, \sigma_{\mathbb{C}})$ eingebettet werden, wobei $\sigma_{\mathbb{C}}$ die positiv definite Hermitesche Form ist, die die positiv definite Bilinearform σ fortsetzt.

Satz 11.66 *Sei $\mathbf{V}$ ein endlichdimensionaler euklidischer Vektorraum und $f \in L(\mathbf{V}, \mathbf{V})$. Dann gelten die folgenden Charakterisierungen:*

1. *f ist genau dann normal, wenn es eine Orthonormalbasis $\mathbf{B}$ von $\mathbf{V}$ gibt, so daß $\Phi_{\mathbf{B}}(f)$ erweiterte Diagonalgestalt der Form*
$$\Phi_{\mathbf{B}}(f) = \operatorname{diag}\left(t_1, t_2, \ldots, t_r, \begin{pmatrix} \sigma_1 & \tau_1 \\ -\tau_1 & \sigma_1 \end{pmatrix}, \begin{pmatrix} \sigma_2 & \tau_2 \\ -\tau_2 & \sigma_2 \end{pmatrix}, \ldots, \begin{pmatrix} \sigma_s & \tau_s \\ -\tau_s & \sigma_s \end{pmatrix}\right)$$
hat.

2. *f ist genau dann selbstadjungiert, wenn es eine Orthonormalbasis $\mathbf{B}$ von $\mathbf{V}$ gibt, die nur aus Eigenvektoren von f besteht.*

3. *f ist genau dann antiselbstadjungiert, d.h. $\hat{f} = -f$, wenn es eine Orthonormalbasis $\mathbf{B}$ von $\mathbf{V}$ gibt, so daß $\Phi_{\mathbf{B}}(f)$ erweiterte Diagonalgestalt der Form*

$$\Phi_{\mathbf{B}}(f) = \operatorname{diag}\left(0, 0, \ldots, 0, \begin{pmatrix} 0 & \tau_1 \\ -\tau_1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & \tau_2 \\ -\tau_2 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & \tau_s \\ -\tau_s & 0 \end{pmatrix}\right)$$

hat.

4. *f ist genau dann eine Isometrie, wenn es eine Orthonomalbasis $\mathbf{B}$ von $\mathbf{V}$, so daß $\Phi_{\mathbf{B}}(f)$ erweiterte Diagonalgestalt der Form*

$$\begin{aligned} &\Phi_{\mathbf{B}}(f) \\ &= \operatorname{diag}\left(1, 1, \ldots, 1, -1, -1 \ldots, -1, \begin{pmatrix} \cos(\varphi_1) & -\sin(\varphi_1) \\ \sin(\varphi_1) & \cos(\varphi_1) \end{pmatrix}, \ldots, \begin{pmatrix} \cos(\varphi_s) & -\sin(\varphi_s) \\ \sin(\varphi_s) & \cos(\varphi_s) \end{pmatrix}\right) \end{aligned}$$

mit $0 < \varphi_1, \ldots, \varphi_s < \pi$ hat.

Der entsprechende Satz für Matrizen lautet folgendermaßen.

Satz 11.67 *Sei $A \in \mathbb{R}^{n \times n}$ eine Matrix. Dann gelten die folgenden Charakterisierungen.*

1. *A ist genau dann normal, d.h. $AA^T = A^T A$, wenn A orthogonal–ähnlich zu einer erweiterten Diagonalmatrix der Form*

$$\operatorname{diag}\left(t_1, t_2, \ldots, t_r, \begin{pmatrix} \sigma_1 & \tau_1 \\ -\tau_1 & \sigma_1 \end{pmatrix}, \begin{pmatrix} \sigma_2 & \tau_2 \\ -\tau_2 & \sigma_2 \end{pmatrix}, \ldots, \begin{pmatrix} \sigma_s & \tau_s \\ -\tau_s & \sigma_s \end{pmatrix}\right)$$

ist, d.h. es gibt eine orthogonale Matrix $P \in \mathbb{R}^{n \times n}$, so daß $P^{-1}AP = P^T AP$ von dieser Gestalt ist.

2. *A ist genau dann symmetrisch, d.h. $A = A^T$, wenn A orthogonal–ähnlich zu einer Diagonalmatrix ist.*

3. *A ist genau dann antisymmetrisch, d.h. $A = -A^T$, wenn A orthogonal–ähnlich zu einer erweiterten Diagonalmatrix der Form*

$$\operatorname{diag}\left(0, 0, \ldots, 0, \begin{pmatrix} 0 & \tau_1 \\ -\tau_1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & \tau_2 \\ -\tau_2 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & \tau_s \\ -\tau_s & 0 \end{pmatrix}\right)$$

ist.

4. *A ist genau dann orthogonal, d.h. $A^{-1} = A^T$, wenn A orthogonal–ähnlich zu einer erweiterten Diagonalmatrix der Form*

$$\operatorname{diag}\left(1, 1, \ldots, 1, -1, -1 \ldots, -1, \begin{pmatrix} \cos(\varphi_1) & -\sin(\varphi_1) \\ \sin(\varphi_1) & \cos(\varphi_1) \end{pmatrix}, \ldots, \begin{pmatrix} \cos(\varphi_s) & -\sin(\varphi_s) \\ \sin(\varphi_s) & \cos(\varphi_s) \end{pmatrix}, \right)$$

mit $0 < \varphi_1, \ldots, \varphi_s < \pi$ ist.

Die isometrischen Abbildungen auf einem euklidischen Vektorraum $\mathbf{V}$ bzw. die reellen orthogonalen Matrizen $A \in \mathbb{R}^{n\times n}$ bilden eine Gruppe, die sogenannte **orthogonale Gruppe**, die durch $O(\mathbf{V})$ bzw. durch $O(n,\mathbb{R})$ bezeichnet wird. Von besonderem Interesse sind die speziellen Abbildungen bzw. Matrizen dieses Typs, nämlich jene mit $\det(f) = 1$ bzw. $\det A = 1$. Diese Gruppen werden mit $SO(\mathbf{V})$ bzw. $SO(n,\mathbb{R})$ bezeichnet.

Definition 11.68 *Eine isometrische Abbildung f mit Determinante* $\det(f) = 1$ *auf einem euklidischen Vektorraum heißt* **Drehung**.

Drehungen werden also im endlichdimensionalen Fall (bezüglich einer ONB) durch Matrizen der Form

$$\operatorname{diag}\left(1,1,\ldots,1,-1,-1\ldots,-1,\begin{pmatrix}\cos(\varphi_1) & -\sin(\varphi_1)\\ \sin(\varphi_1) & \cos(\varphi_1)\end{pmatrix},\ldots,\begin{pmatrix}\cos(\varphi_s) & -\sin(\varphi_s)\\ \sin(\varphi_s) & \cos(\varphi_s)\end{pmatrix},\right)$$

realisiert, wobei die Anzahl der -1 gerade ist. Daraus ergibt sich der folgende Satz.

Satz 11.69 *Jede Drehung f in einem endlichdimensionalen euklidischen Raum* $\mathbf{V}$ *ungerader Dimension besitzt einen Fixvektor* $\mathbf{a} \in \mathbf{V} \setminus \{\mathbf{0}\}$*, d.h.* $f(\mathbf{a}) = \mathbf{a}$.

Der von diesem Fixvektor aufgespannte eindimensionale Unterraum kann als *Drehachse* interpretiert werden.

Abschließend geben wir noch ein weiteres Kriterium für positiv (semi-)definite Matrizen an.

Satz 11.70 *Sei $A \in \mathbb{R}^{n\times n}$ (bzw. $A \in \mathbb{C}^{n\times n}$) eine symmetrische (bzw. Hermitesche) Matrix. Dann ist A genau dann positiv definit, wenn alle Eigenwerte von A positiv sind, und A ist positiv semidefinit, wenn alle Eigenwerte von A nicht negativ (also ≥ 0) sind.*

11.4.4 Singulärwerte

Beliebige lineare Abbildungen $f \in L(\mathbf{V},\mathbf{W})$ können natürlich nicht invertiert werden. In euklidischen bzw. unitären Vektorräumen $\mathbf{V},\mathbf{W}$ ist es aber möglich, eine *inverse Abbildung* zu *simulieren*. Sei also $f \in L(\mathbf{V},\mathbf{W})$ und $\mathbf{b} \in \mathbf{W}$. Es soll ein Vektor $\mathbf{a} \in \mathbf{V}$ angegeben werden, so daß einerseits der Abstand von $f(\mathbf{a})$ zu $\mathbf{b}$ kleinstmöglich ist, d.h.

$$\|f(\mathbf{a}) - \mathbf{b}\| = \min\{\|f(\mathbf{a}') - \mathbf{b}\| \mid \mathbf{a}' \in \mathbf{V}\}$$

und andererseits unter allen Vektoren $\mathbf{a} \in \mathbf{V}$, die diese Eigenschaft erfüllen, die Länge $\|\mathbf{a}\|$ kleinstmöglich ist.

Dazu muß zunächst jener Vektor $\mathbf{b}' \in f(\mathbf{V}) \subseteq \mathbf{W}$ gesucht werden, für den der Abstand $\|\mathbf{b}' - \mathbf{b}\|$ minimal ist. Das ist aber gerade $\mathbf{b}' = p(\mathbf{b})$, wobei $p : \mathbf{W} \to f(\mathbf{V})$ die Orthogonalprojektion von $\mathbf{W}$ auf $f(\mathbf{V})$ bezeichnet. Nun gibt es sicherlich Vektoren $\mathbf{a} \in \mathbf{V}$ mit $f(\mathbf{a}) = \mathbf{b}'$. Der kürzeste von diesen liegt im Orthogonalraum der Kerns: $\operatorname{kern}(f)^\perp = \hat{f}(\mathbf{W})$.

Satz 11.71 *Seien* $\mathbf{V},\mathbf{W}$ *euklidische bzw. unitäre Vektorräume und* $f \in L(\mathbf{V},\mathbf{W})$ *eine lineare Abbildung, deren adjungierte Abbildung* $\hat{f} \in L(\mathbf{W},\mathbf{V})$ *existiert. Bezeichnet* $p : \mathbf{W} \to f(\mathbf{V})$ *die*

Orthogonalprojektion von $\mathbf{W}$ *auf* $f(\mathbf{V})$ *und* $f_1 : f(\mathbf{V}) \to \hat{f}(\mathbf{W})$ *die Inverse der Einschränkung* $\tilde{f} : \hat{f}(\mathbf{W}) \to f(\mathbf{V})$ *von* f*. Dann ist*

$$\mathbf{a} = f_1(p(\mathbf{b})) \in \mathbf{V}$$

jener Vektor minimaler Länge $\|\mathbf{a}\|$ *mit*

$$\|f(\mathbf{a}) - \mathbf{b}\| = \min\{\|f(\mathbf{a}') - \mathbf{b}\| \mid \mathbf{a}' \in \mathbf{V}\}$$

Die Abbildung $f^+ = f_1 \circ p : \mathbf{W} \to \mathbf{V}$ ist also linear und erfüllt die Eigenschaften einer *pseudoinversen Abbildung.*

Definition 11.72 *Seien* $\mathbf{V}, \mathbf{W}$ *euklidische bzw. unitäre Vektorräume und* $f \in L(\mathbf{V}, \mathbf{W})$*. Eine Abbildung* $g \in L(\mathbf{W}, \mathbf{V})$ *heißt* **Moore-Penrose-Pseudoinverse***, wenn*

$$f \circ g \quad \textit{und} \quad g \circ f$$

selbstadjungiert sind und

$$f = f \circ g \circ f \quad \textit{und} \quad g = g \circ f \circ g$$

gilt.

Satz 11.73 *Seien* $\mathbf{V}, \mathbf{W}$ *euklidische bzw. unitäre Vektorräume und* $f \in L(\mathbf{V}, \mathbf{W})$ *eine lineare Abbildung, deren adjungierte Abbildung* $\hat{f} \in L(\mathbf{W}, \mathbf{V})$ *existiert. Dann gibt es eine eindeutig bestimmte pseudoinverse Abbildung* $g \in L(\mathbf{W}, \mathbf{V})$*, die durch*

$$g = f_1 \circ p$$

gegeben ist, wobei $p : \mathbf{W} \to f(\mathbf{V})$ *die Orthogonalprojektion von* $\mathbf{W}$ *auf* $f(\mathbf{V})$ *und* $f_1 : f(\mathbf{V}) \to \hat{f}(\mathbf{W})$ *die Inverse der Einschränkung* $\tilde{f} : \hat{f}(\mathbf{W}) \to f(\mathbf{V})$ *von* f *ist.*

Im folgenden wird für die Pseudoinverse immer die Bezeichnung f^+ verwendet werden. Man beachte auch, daß im Falle einer bijektiven Abbildung $f^+ = f^{-1}$ ist.

In vielen Fällen kann die Pseudoinverse sehr einfach angegeben werden.

Satz 11.74 *Seien* $\mathbf{V}, \mathbf{W}$ *euklidische bzw. unitäre Vektorräume und* $f \in L(\mathbf{V}, \mathbf{W})$ *eine lineare Abbildung, deren adjungierte Abbildung* $\hat{f} \in L(\mathbf{W}, \mathbf{V})$ *existiert. Ist die Abbildung* $\hat{f} \circ f : \mathbf{V} \to \mathbf{V}$ *bijektiv, so gilt*

$$f^+ = (\hat{f} \circ f)^{-1} \circ \hat{f}.$$

Man beachte, daß im Endlichdimensionalen immer eine pseudoinverse Abbildung existiert. Insbesondere können diese Ergebnisse auf Matrizen übertragen werden. Das einleitend geschilderte Problem transformiert sich nun in ein *Lösen* eines möglicherweise unlösbaren linearen Gleichungssystem. Sei $A \in \mathbb{R}^{m \times n}$ bzw. $A \in \mathbb{C}^{m \times n}$ eine Matrix und $\mathbf{b} \in \mathbb{R}^{n \times 1}$ bzw. $\mathbf{b} \in \mathbb{C}^{n \times 1}$. Gesucht ist ein Vektor $\mathbf{x} \in \mathbb{R}^{n \times 1}$ bzw. $\mathbf{x} \in \mathbb{C}^{n \times 1}$ minimaler Länge $\|\mathbf{x}\|$, so daß der Abstand

$$\|A\mathbf{x} - \mathbf{b}\|$$

kleinstmöglich ist.

Definition 11.75 *Sei $A \in \mathbb{R}^{m\times n}$ bzw. $A \in \mathbb{C}^{m\times n}$. Eine Matrix $B \in \mathbb{R}^{m\times n}$ bzw. $B \in \mathbb{C}^{m\times n}$ heißt* **Moore-Penrose-pseudoinverse Matrix**, *wenn*

$$AB \quad \text{und} \quad BA$$

symmetrisch bzw. Hermitesch sind und

$$A = ABA \quad \text{und} \quad B = BAB$$

gilt.

Es existiert immer eine pseudoinverse Matrix B, und sie ist eindeutig bestimmt. Sie wird im folgenden durch A^+ bezeichnet werden. Bei einer invertierbaren Matrix A gilt natürlich $A^+ = A^{-1}$.

In vielen Fällen kann A^+ sehr einfach angegeben werden.

Satz 11.76 *Sei $A \in \mathbb{R}^{m\times n}$ bzw. $A \in \mathbb{C}^{m\times n}$, so daß A^TA bzw. A^HA invertierbar ist. Dann gilt*

$$A^+ = (A^TA)^{-1}A^T \quad \text{bzw.} A^+ = (A^HA)^{-1}A^H.$$

Im allgemeinen Fall ist es zur Berechnung von A^+ günstig, die *Singulärwerte* von A bzw. f zu bestimmen.

Lemma 11.77 *Seien $\mathbf{V}, \mathbf{W}$ euklidische bzw. unitäre Vektorräume und $f \in L(\mathbf{V}, \mathbf{W})$ eine lineare Abbildung, deren adjungierte Abbildung $\hat{f} \in L(\mathbf{W}, \mathbf{V})$ existiert. Dann ist $\hat{f} \circ f \in L(\mathbf{V}, \mathbf{V})$ selbstadungiert, und die Eigenwerte von $\hat{f} \circ f$ sind nichtnegativ.*

Definition 11.78 *Seien $\mathbf{V}, \mathbf{W}$ endlichdimensionale euklidische bzw. unitäre Vektorräume und $f \in L(\mathbf{V}, \mathbf{W})$. Die positiven Wurzeln*

$$\sqrt{w_1},\ \sqrt{w_2}, \ldots,\ \sqrt{w_n} \geq 0$$

der Eigenwerte von $\hat{f} \circ f$ heißen **Singulärwerte** *von f.*

Definition 11.79 *Sei $A \in \mathbb{R}^{m\times n}$ bzw. $A \in \mathbb{C}^{m\times n}$. Die positiven Wurzeln*

$$\sqrt{w_1},\ \sqrt{w_2}, \ldots,\ \sqrt{w_n} \geq 0$$

der Eigenwerte von A^TA bzw. A^HA heißen **Singulärwerte** *von A.*

Im folgenden werden die Singulärwerte so geordnet, daß

$$w_1 > 0, \ldots, w_r > 0 \quad \text{und} \quad w_{r+1} = \cdots = w_n = 0$$

sind.

Lemma 11.80 *Seien* $\mathbf{V}, \mathbf{W}$ *endlichdimensionale euklidische bzw. unitäre Vektorräume und* $f \in L(\mathbf{V}, \mathbf{W})$. *Sei* $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ *eine Orthonormalbasis aus Eigenvektoren von* $\hat{f} \circ f$ *mit Eigenwerten* $w_1 > 0, \ldots, w_r > 0, w_{r+1} = \cdots = w_n = 0$. *Dann gilt*

$$\langle f(\mathbf{b}_i), f(\mathbf{b}_j)\rangle = w_i \delta_{ij}, \qquad (1 \leq i, j \leq n)$$

d.h. $\|f(\mathbf{b}_i)\| = \sqrt{w_i}$, *und die Bildvektoren* $f(\mathbf{b}_1), \ldots, f(\mathbf{b}_r)$ *bilden ein Orthogonalsystem von* $f(\mathbf{V})$.

Normiert man die Bildvektoren $f(\mathbf{b}_1), \ldots, f(\mathbf{b}_r)$ und ergänzt diese zu einer Orthonormalbasis $\mathbf{C}$ von $\mathbf{W}$, so hat die Matrix $\Phi_{\mathbf{BC}}(f)$ folgende Gestalt:

$$\Phi_{\mathbf{BC}}(f) = \begin{pmatrix} \sqrt{w_1} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{w_2} & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \sqrt{w_r} & \vdots \\ 0 & \cdots & 0 & 0 & 0 \end{pmatrix}.$$

Interpretiert man diese Eigenschaft für Matrizen, erhält man die sogenannte **Singulärwertzerlegung** eine Matrix A.

Satz 11.81 *Sei* $A \in \mathbb{R}^{m\times n}$ *bzw.* $A \in \mathbb{C}^{m\times n}$. *Dann gibt es orthogonale bzw. unitäre Matrizen* U, V *mit*

$$A = U \begin{pmatrix} \sqrt{w_1} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{w_2} & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \sqrt{w_r} & \vdots \\ 0 & \cdots & 0 & 0 & 0 \end{pmatrix} V,$$

wobei $\sqrt{w_1}, \sqrt{w_2}, \ldots, \sqrt{w_r} > 0$ *die positiven Singulärwerte von* A *bezeichnen.*

Mit Hilfe der Singulärwertzerlegung gelingt es auch, allgemein die pseudoinverse Matrix zu berechnen.

Satz 11.82 *Sei* $A \in \mathbb{R}^{m\times n}$ *bzw.* $A \in \mathbb{C}^{m\times n}$ *und*

$$A = U \begin{pmatrix} \sqrt{w_1} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{w_2} & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \sqrt{w_r} & \vdots \\ 0 & \cdots & 0 & 0 & 0 \end{pmatrix} V,$$

die Singulärwertzerlegung von A *mit orthogonalen bzw. unitären Matrizen* U, V, *dann ist die pseudoinverse Matrix* A^+ *durch*

$$A^+ = V^{-1} \begin{pmatrix} \frac{1}{\sqrt{w_1}} & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{w_2}} & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{1}{\sqrt{w_r}} & \vdots \\ 0 & \cdots & 0 & 0 & 0 \end{pmatrix} U^{-1}$$

gegeben.

11.4.5 Polar- und QR-Zerlegung

In Analogie zu definiten Bilinearformen und definiten Matrizen kann man auch definite selbstadjungierte Abbildungen betrachten.

Definition 11.83 *Sei* $\mathbf{V}$ *ein euklidischer oder unitärer endlichdimensionaler Vektorraum. Eine selbstadjungierte Abbildung* $f \in L(\mathbf{V}, \mathbf{V})$ *heißt positiv definit, wenn alle Eigenwerte von* f *positiv sind und positiv semidefinit, wenn alle Eigenwerte* ≥ 0 *sind.*

Entsprechend kann man auch negativ definite Abbildungen definieren.

Man beachte, daß eine Abbilung $f \in L(\mathbf{V}, \mathbf{V})$ genau dann positiv (semi-)definit ist, wenn die $\Phi_{\mathbf{BB}}(f)$ positiv (semi-)definit ist.

Lemma 11.84 *Sei* $\mathbf{V}$ *ein euklidischer oder unitärer endlichdimensionaler Vektorraum. Eine selbstadjungierte Abbildung* $f \in L(\mathbf{V}, \mathbf{V})$ *ist genau dann positiv semidefinit, wenn es eine positiv semidefinite Abbildung* $g \in L(\mathbf{V}, \mathbf{V})$ *mit* $f = g \circ g$ *gibt. Die Abbildung* g *ist dann auch eindeutig bestimmt.*

Man bezeichnet diese Abbildung g auch als **Quadratwurzel** $\sqrt{f}$ von f. Man beachte auch, daß f genau dann positiv definit ist, wenn $g = \sqrt{f}$ positiv definit ist.

Mit Hilfe dieser Begriffsbildung läßt sich die **Polarzerlegung** einer Abbildung angeben.

Satz 11.85 *Sei* $\mathbf{V}$ *ein euklidischer oder unitärer endlichdimensionaler Vektorraum und* $f \in L(\mathbf{V}, \mathbf{V})$. *Dann gibt es eine Isometrie* $g \in L(\mathbf{V}, \mathbf{V})$ *und eine positiv semidefinite Abbildung* $s \in L(\mathbf{V}, \mathbf{V})$ *mit*

$$f = g \circ s,$$

wobei in jeder solchen Zerlegung $s = \sqrt{\hat{f} \circ f}$ *gilt.*

Entsprechend gilt für Matrizen:

Satz 11.86 *Sei* $A \in \mathbb{R}^{n \times n}$ *(bzw.* $A \in \mathbb{C}^{n \times n}$*). Dann gibt es eine orthogonale Matrix* Q *(bzw. eine unitäre Matrix* Q*) und eine positiv semidefinite Matrix* S *mit*

$$A = QS,$$

wobei in jeder solcher Zerlegung $S = \sqrt{A^H A}$ *gilt.*

Der Name *Polarzerlegung* kommt von der Analogie zu den komplexen Zahlen, die als Produkt $z = e^{i\theta} r$ dargestellt werden können.

Eine ähnlich Zerlegung ist die sogenannte **QR-Zerlegung**.

Satz 11.87 *Sei* $A \in \mathbb{R}^{n \times n}$ *(bzw.* $A \in \mathbb{C}^{n \times n}$*). Dann gibt es eine orthogonale Matrix* Q *(bzw. eine unitäre Matrix* Q*) und eine obere Dreiecksmatrix* R *mit*

$$A = QR.$$

Dieser Satz ist Grundlage des QR-**Verfahrens** zur näherungsweisen Bestimmung der Eigenwerte einer Matrix A. Man beginnt mit der Zerlegung $A_0 = A = Q_0 R_0$ und betrachtet darauf die Matrix $A_1 = Q_0^H A_0 Q_0 = R_0 Q_0$, die natürlich dieselben Eigenwerte wie A hat, und bestimmt die QR-Zerlegung dieser Matrix: $A_1 = Q_1 R_1$. Dieses Verfahren wird nun iteriert. Haben die Eigenwerte von A nicht alle nahezu den selben Absolutbetrag, dann konvergieren die Matrizen A_n zu einer oberen Dreiecksmatrix. Damit sind die Diagonalelemente von A_n Näherungswerte für die Eigenwerte von A.

11.4.6 Gramsche Matrix

Definition 11.88 *Sei σ ein anisotropes Skalarprodukt auf einem Vektorraum* $\mathbf{V}$ *und seien* $\mathbf{a}_1, \ldots, \mathbf{a}_p \in \mathbf{V}$ *Vektoren aus* $\mathbf{V}$. *Dann bezeichnet man die $p \times p$-Matrix*

$$G(\mathbf{a}_1, \ldots, \mathbf{a}_p) := (\langle \mathbf{a}_i, \mathbf{a}_j \rangle)_{1 \leq i,j \leq p}$$

als **Gramsche Matrix** *und die Determinante* $\det G(\mathbf{a}_1, \ldots, \mathbf{a}_p)$ *als* **Gramsche Determinante**.

Die wichtigste Eigenschaft von $G(\mathbf{a}_1, \ldots, \mathbf{a}_p)$ ist die folgende:

Lemma 11.89 *Sei σ ein anisotropes Skalarprodukt auf einem Vektorraum* $\mathbf{V}$ *und seien* $\mathbf{a}_1, \ldots, \mathbf{a}_p \in \mathbf{V}$ *Vektoren aus* $\mathbf{V}$. *Dann sind* $\mathbf{a}_1, \ldots, \mathbf{a}_p$ *genau dann linear abhängig, wenn die Gramsche Determinante* $\det G(\mathbf{a}_1, \ldots, \mathbf{a}_p) = 0$ *ist.*

In einem endlichdimensionalen Vektorraum mit Basis $\mathbf{B}$ besteht auch die Darstellung

$$G(\mathbf{a}_1, \ldots, \mathbf{a}_p) = A^T \Phi_{\mathbf{B}}(\sigma)\, \zeta(A)$$

mit

$$A = \begin{pmatrix} \Phi_{\mathbf{B}}(\mathbf{a}_1) & \Phi_{\mathbf{B}}(\mathbf{a}_2) & \cdots & \Phi_{\mathbf{B}}(\mathbf{a}_p) \end{pmatrix}.$$

Insbesondere gilt in einem euklidischen Vektorraum mit einer ONB $\mathbf{B}$

$$G(\mathbf{a}_1, \ldots, \mathbf{a}_p) = A^T A,$$

woraus auch

$$\det G(\mathbf{a}_1, \ldots, \mathbf{a}_p) \geq 0$$

folgt, da eine Matrix der Form $A^T A$ immer positiv semidefinit ist.

Im euklidischen Fall hat die Gramsche Determinante auch eine interessante Interpretation.

Satz 11.90 *Zu Vektoren* $\mathbf{a}_1, \ldots, \mathbf{a}_p \in \mathbb{R}^n$ *bezeichne*

$$\Delta(\mathbf{a}_1, \ldots, \mathbf{a}_p) := \{t_1 \mathbf{a}_1 + \cdots + t_p \mathbf{a}_p \,|\, 0 \leq t_1, \ldots, t_p \leq 1\}$$

das von $\mathbf{a}_1, \ldots, \mathbf{a}_p$ *aufgespannte Parallelepiped. Das p-dimensionale Lebesguemaß von* $\Delta(\mathbf{a}_1, \ldots, \mathbf{a}_p)$ *ist durch*

$$\mathrm{Vol}_p(\Delta(\mathbf{a}_1, \ldots, \mathbf{a}_p)) = \sqrt{G(\mathbf{a}_1 \ldots, \mathbf{a}_p)}$$

gegeben.

Abschließend wird noch eine interessante Determinatenidentität angegeben, die mit der Gramschen Determinante zusammenhängt.

Satz 11.91 *Sei $p \le n$, $A = (a_{ij})_{1\le i\le p, 1\le j\le n} \in K^{p\times n}$ und $B = (b_{lk})_{1\le l\le n, 1\le k\le p} \in K^{n\times p}$. Dann gilt*

$$\det(AB) = \sum_{1\le j_1<\cdots<j_p\le n} \begin{vmatrix} a_{1j_1} & \cdots & a_{1j_p} \\ \vdots & & \vdots \\ a_{pj_1} & \cdots & a_{pj_p} \end{vmatrix} \begin{vmatrix} b_{j_1 1} & \cdots & b_{j_p 1} \\ \vdots & & \vdots \\ b_{j_1 p} & \cdots & b_{j_p p} \end{vmatrix}.$$

Daraus folgt ein zweites Mal $\det G(\mathbf{a}_1, \dots, \mathbf{a}_p) \ge 0$.

11.5 Euklidische Quadriken

11.5.1 Diagonalisieren symmetrischer Matrizen

Ist τ eine symmetrische Bilinearform (bzw. eine Hermitesche Form) in einem endlichdimensionalen euklidischen (bzw. unitären) Vektorraum $\mathbf{V}$ und bezeichne $\mathbf{B}$ eine Orthonormalbasis von $\mathbf{V}$, dann ist die Matrix $G = \Phi_{\mathbf{B}}(\tau)$ symmetrisch. Es gibt daher eine orthogonale (bzw. unitäre) Matrix P, sodaß

$$P^{-1}GP = \operatorname{diag}(g_1, \dots, g_n)$$

eine Diagonalmatrix ist. Die Matrix $Q = \overline{P}$ ist dann auch orthogonal (bzw. unitär) und es gilt wegen $P^{-1} = Q^T$, daß

$$Q^T G\overline{Q} = P^{-1}GP = \operatorname{diag}(g_1, \dots, g_n)$$

eine Diagonalmatrix ist. Das heißt, die Matrix Q vermittelt einen Basiswechsel von der Basis $\mathbf{B}$ zu einer Basis $\mathbf{C}$ mit

$$\Phi_{\mathbf{C}}(\tau) = Q^T G\overline{Q} = \operatorname{diag}(g_1, \dots, g_n).$$

Da Q orthogonal (bzw. unitär) ist, ist $\mathbf{C}$ wieder eine Orthonormalbasis.

Satz 11.92 *Sei $(\mathbf{V}, \sigma)$ ein endlichdimensionaler euklidischer bzw. unitärer Vektorraum und τ eine symmetrische Bilinearform bzw. Hermitesche Form auf $\mathbf{V}$. Dann gibt es eine Orthonormalbasis $\mathbf{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ ($n = \dim(\mathbf{V})$) von $\mathbf{V}$, sodaß $\Phi_{\mathbf{B}}(\tau)$ Diagonalgestalt hat: d.h.*

$$\tau(\mathbf{b}_i, \mathbf{b}_j) = 0 \quad \textit{für } i \ne j.$$

σ ist ja nichts anderes als eine definite symmetrische Bilinearform bzw. Hermitesche Form auf $\mathbf{V}$, und eine Orthonormalbasis $\mathbf{B}$ ist durch $\Phi_{\mathbf{B}}(\sigma) = E_n$ charakterisiert. Insbesondere ist E_n auch eine Diagonalmatrix. Satz 11.92 kann daher folgendermaßen umformuliert werden (**Hauptachsentransformation**):

Satz 11.93 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum über $\mathbb{R}$ bzw. $\mathbb{C}$ und seien σ, τ zwei symmetrische Bilinearformen bzw. Hermitesche Formen auf $\mathbf{V}$, von denen eine definit ist. Dann gibt es eine Basis $\mathbf{B}$ von $\mathbf{V}$, so daß $\Phi_{\mathbf{B}}(\sigma)$ und $\Phi_{\mathbf{B}}(\tau)$ Diagonalmatrizen sind.*

Eine entsprechende Version dieses Satzes für Matrizen lautet folgendermaßen:

Satz 11.94 *Seien A, B zwei symmetrische bzw. Hermitesche Matrizen aus $\mathbb{R}^{n\times n}$ bzw. $\mathbb{C}^{n\times n}$, so daß A oder B definit ist (d.h. eine der quadratischen Formen $q_A(\mathbf{x}) = \mathbf{x}^T A\overline{\mathbf{x}}$, $q_B(\mathbf{x}) = \mathbf{x}^T B\overline{\mathbf{x}}$ ist definit). Dann gibt es eine reguläre Matrix P ($\in \mathbb{R}^{n\times n}$ bzw. $\mathbb{C}^{n\times n}$), so daß*

$$P^T A\overline{P} \quad \textit{und} \quad P^T B\overline{P}$$

Diagonalmatrizen sind.

Zunächst ermittelt man mit Kongruenzumformungen (gleichzeitiges Spalten- und Zeilenumformen) eine reguläre Matrix P_1 mit $P_1^T A\overline{P}_1 = E_n$. Die Matrix $C = P_1^T BP_1$ ist dann auch symmetrisch (bzw. Hermitesch), und mit Hilfe des Spektralsatzes kann man eine orthogonale (bzw. unitäre) Matrix P_2 bestimmen, so daß $P_2^{-1}CP_2$ Diagonalform hat. Die Matrix $P = P_1\overline{P}_2$ diagonalisiert dann beide Matrizen A, B im Sinn von Satz 11.94.

Es ist auch möglich, die Spalten der Matrix P durch Eigenvektoren von $A^{-1}B$ zusammenzufassen, so daß sie bezüglich des von A induzierten Skalarprodukts orthogonal sind.

11.5.2 Euklidische Geometrie

Sei $\mathbf{N} = \mathbf{c} + \mathbf{U}$ ein Nebenraum eines euklidischen bzw. unitären Vektorraums $\mathbf{V}$. Die affine Geometrie $\mathcal{A}(\mathbf{N})$ gewinnt durch das in $\mathbf{V}$ vorhandene Skalarprodukt eine erweiterte Struktur. Insbesondere kann man in natürlicher Weise den Abstand zweier Punkte, den Winkel zweier einander schneidender Geraden und auf den Elementen von $\mathcal{A}(\mathbf{N})$ eine Orthogonalitätsrelation einführen.

Definition 11.95 *Sei $\mathbf{N} = \mathbf{c}+\mathbf{U}$ ein Nebenraum eines euklidischen bzw. unitären Vektorraums $\mathbf{V}$. Dann heißt die affine Geometrie $\mathcal{A}(\mathbf{N})$* **euklidische Geometrie** *bzw.* **unitäre Geometrie** *auf $\mathbf{N}$.*

Der **Abstand** *zweier Punkte $a = \{\mathbf{a}\}, b = \{\mathbf{b}\} \in \mathbf{U}$ aus $\mathcal{A}(\mathbf{N})$ ist durch*

$$d(a, b) := \|\mathbf{b} - \mathbf{a}\| = \sqrt{\langle \mathbf{b} - \mathbf{a}, \mathbf{b} - \mathbf{a}\rangle}$$

definiert.

Sind $g_1 = \mathbf{c}_1 + [\mathbf{a}_1], g_1 = \mathbf{c}_2 + [\mathbf{a}_2] \in \mathcal{A}(\mathbf{N})$ zwei Geraden mit nichtleerem Schnitt, so ist der **Winkel** *$\varphi \in [0, \frac{\pi}{2}]$ zwischen g_1 und g_1 durch*

$$\cos\varphi = \frac{|\langle \mathbf{a}_1, \mathbf{a}_2\rangle|}{\|\mathbf{a}_1\| \cdot \|\mathbf{a}_2\|}$$

gegeben.

Zwei Elemente $\mathbf{c}_1 + \mathbf{U}_1, \mathbf{c}_2 + \mathbf{U}_2 \in \mathcal{A}(\mathbf{N})$ heißen **orthogonal**, *wenn $\mathbf{U}_1 \perp \mathbf{U}_2$ gilt.*

Insbesondere gibt es in euklidischen bzw. unitären Geometrien Analoga zu Orthonormalbasen.

Definition 11.96 *Ein affines Koordinatensystem $\{\mathbf{u}, \mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ einer n-dimensionalen euklidischen bzw. unitären Geometrie $\mathcal{A}(\mathbf{N})$ heiß* **kartesisches Koordinatensystem**, *wenn*

$$\{\mathbf{p}_1 - \mathbf{u}, \mathbf{p}_2 - \mathbf{u}, \ldots, \mathbf{p}_n - \mathbf{u}\}$$

eine Orthonormalbasis von $\mathbf{U}$ ist.

11.5.3 Normalformen euklidischer Quadriken

Im Abschnitt 10.6 wurde gezeigt, daß es zu jeder affinen Quadrik Q_{aff} auf einem endlichdimensionalen affinen Raum eine affine Basis gibt, so daß bezüglich deren Koordinaten $(x_1, \ldots, x_n)$ die Quadrik Q_{aff} entweder durch die Gleichung

$$g_0 + g_1 x_1^2 + g_2 x_2^2 + \cdots + g_n x_n^2 = 0$$

oder durch die Gleichung

$$x_1 + g_2 x_2^2 + \cdots + g_n x_n^2 = 0$$

bestimmt wird. In euklidischen Geometrien stellt sich natürlich die Frage, ob man auch ein kartesisches Koordinatensystem finden kann, so daß eine Quadrik durch eine so einfache Gleichung wie die oberen beschrieben werden kann. Dies ist tatsächlich der Fall.

Satz 11.97 *Sei* $\mathbf{N} = \mathbf{c}+\mathbf{U}$ *ein endlichdimensionaler Nebenraum eines euklidischen Vektorraums* $\mathbf{V}$ *und*

$$\lambda(\mathbf{a}) = q(\mathbf{a} - \mathbf{c}) + l(\mathbf{a} - \mathbf{c}) + g$$

eine quadratische Funktion auf $\mathbf{N}$. *Dann gibt es ein kartesiches Koordinatensystem* $\{\mathbf{u}, \mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ *der euklidischen Geometrie* $\mathcal{A}(\mathbf{N})$, *so daß einem Vektor* $\mathbf{a} \in \mathbf{N}$ *mit den affinen Koordinaten* $(x_1, \ldots, x_n)$, *d.h.*

$$\mathbf{a} = \mathbf{u} + x_1(\mathbf{p}_1 - \mathbf{u}) + x_1(\mathbf{p}_2 - \mathbf{u}) + \cdots + x_n(\mathbf{p}_n - \mathbf{u})$$

die quadratische Funktion $q(\mathbf{a})$ *entweder durch*

$$q(\mathbf{a}) = g_0 + g_1 x_1^2 + g_2 x_2^2 + \cdots + g_n x_n^2$$

mit $g_0, g_1, \ldots, g_n \in \mathbb{R}$ *oder durch*

$$q(\mathbf{a}) = g_1 x_1 + g_2 x_2^2 + \cdots + g_n x_n^2$$

mit $g_1, \ldots, g_n \in \mathbb{R}$ *und* $g_1 \neq 0$ *gegeben ist.*

In einer endlichdimensionalen euklidischen Geometrie gibt es daher zu jeder affinen Quadrik Q_{aff} ein kartesisches Koordinatensystem, so daß die Punkte $a = \{\mathbf{a}\} \in Q_{aff}$ mit den Koordinaten $(x_1, \ldots, x_n)$ entweder durch eine Gleichung der Form

$$g_1 x_1^2 + \cdots + g_p x_p^2 - g_{p+1} x_{p+1}^2 - \cdots - g_{p+q} x_{p+q}^2 = 0$$

mit $p \geq q$, $p + q \leq n$ und $g_1, g_2, \ldots, g_{p+q} > 0$, durch

$$g_1 x_1^2 + \cdots + g_p x_p^2 - g_{p+1} x_{p+1}^2 - \cdots - g_{p+q} x_{p+q}^2 = \pm 1$$

mit $p \geq q$, $p + q \leq n$ und $g_1, g_2, \ldots, g_{p+q} > 0$, oder durch

$$g_2 x_2^2 + \cdots + g_{p+1} x_{p+1}^2 - g_{p+2} x_{p+2}^2 - \cdots - g_{p+q-1} x_{p+q-1}^2 = \pm x_1$$

mit $p \geq q$, $p + q < n$ und $g_1, g_2, \ldots, g_{p+q} > 0$ beschrieben werden kann.

Praktisch wird diese Transformation folgendermaßen durchgeführt:

1. Man wählt ein kartesisches Koordinatensystem auf $\mathbf{N}$ und stellt die quadratische Funktion λ, die die Quadrik Q_{aff} bescheibt, bezüglich dieser Basis in der Form
$$\lambda(\mathbf{a}) = \sum_{i,j=1}^{n} g_{ij}x_i x_j + \sum_{i=1}^{n} a_i x_i + g$$
mit $g_{ij} = g_{ji}$ dar. Die quadratische Form in der entsprechenden projektiven Einbettung hat dann die Form
$$\overline{q} = \sum_{i,j=1}^{n} g_{ij}x_i x_j + \sum_{i=1}^{n} a_i x_0 x_i + g x_0^2.$$
Definiert man $g_{00} := g$ und $g_{0i} = g_{i0} := a_i$, dann ist die Matrix $\overline{G} = (g_{ij})_{0\leq i,j\leq n}$ die Koeffizientenmatrix jener symmetrischen quadratischen Form τ, der die quadratische Form $\overline{q}$ entspricht. Die Hyperebene $\mathbf{H}$ ist durch die Gleichung $x_0 = 0$ festgelegt und der Nebenraum $\mathbf{N}$ durch die Gleichung $x_0 = 1$.

2. Man untersucht, ob $\mathbf{H}^{\perp_\tau} \not\subseteq \mathbf{H}$ oder $\mathbf{H}^{\perp_\tau} \subseteq \mathbf{H}$ erfüllt ist. Dazu muß nur das lineare Gleichungssystem
$$\tilde{G}\mathbf{x} = \mathbf{0}$$
gelöst werden, wobei $\tilde{G} = (g_{ij})_{1\leq i\leq n, 0\leq j\leq n}$ jene Matrix ist, die aus G durch Streichen der 1. Zeile entsteht.

3. Ist $\mathbf{H}^{\perp_\tau} \not\subseteq \mathbf{H}$, so wählt man einen Vektor
$$\mathbf{b}_0 \in (\mathbf{H}^{\perp_\tau} \setminus \mathbf{H}) \cap \mathbf{N}$$
und bestimmt durch Lösen eines Eigenwertproblems eine Orthonormalbasis $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ von $\mathbf{H}$, die Eigenvektoren der Matrix $\overline{G}$ bzw. $G = (g_{ij})_{1\leq i,j\leq n}$ sind. (Man unterdrückt dabei die erste Komponente der Vektoren $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, die ja 0 ist.) Das System von Vektoren $\{\mathbf{b}_0, \mathbf{b}_0 + \mathbf{b}_1, \ldots, \mathbf{b}_0 + \mathbf{b}_n\}$ ist dann das gewünschte kartesische Koordinatensystem, so daß die quadratische Funktion λ die Gestalt
$$\lambda(\mathbf{a}) = g_0 + g_1 x_1^2 + \cdots + g_n x_n^2$$
hat. $g_1, \ldots, g_n$ sind die Eigenwerte von G und $g_0 = \mathbf{b}_0^T \overline{G}\mathbf{b}_0$.

4. Ist $\mathbf{H}^{\perp_\tau} \subseteq \mathbf{H}$, so bestimmt man zunächst $\mathbf{V}^{\perp_\tau}$ durch Lösen des Gleichungssystems
$$\overline{G}\mathbf{x} = \mathbf{0}$$
und eine Orthonormalbasis $\{\mathbf{b}_{r+1}, \ldots, \mathbf{b}_n\}$ von $\mathbf{V}^{\perp_\tau}$. Daraufhin ergänzt man diese zu einer Orthonormalbasis von $\{\mathbf{b}_1, \mathbf{b}_{r+1}, \ldots, \mathbf{b}_n\}$ von $\mathbf{H}^{\perp_\tau}$ und schließlich mit einem Eigenwertproblem zu einer Orthonormalbasis $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_r, \mathbf{b}_{r+1}, \ldots, \mathbf{b}_n\}$ von $\mathbf{H}$, so daß die Vektoren $\mathbf{b}_2, \ldots, \mathbf{b}_r$ Eigenvektoren der Matrix $\overline{G}$ (bzw. G) sind. Für $\mathbf{b}_0$ wird vorläufig der

Vektor $(1,0,0,\ldots,0)^T$ verwendet. Die Matrix $\Phi_{\mathbf{B}}(\tau) = \mathbf{B}^T\overline{G}\mathbf{B}$ hat nun die Gestalt

$$\Phi_{\mathbf{B}}(\tau) = \begin{pmatrix} h_{00} & h_{01} & h_{02} & h_{03} & \cdots & h_{0r} & 0 & \cdots & 0 \\ h_{10} & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ h_{20} & 0 & g_2 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ h_{30} & 0 & 0 & g_3 & \ddots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \ddots & \vdots & \vdots & & \vdots \\ h_{r0} & 0 & 0 & \cdots & 0 & g_r & 0 & \cdots & 0 \\ 0 & 0 & 0 & & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

mit $h_{01} = h_{10} \neq 0$ und $g_2, g_3, \ldots, g_r \neq 0$. Durch Kongruenzumformungen (d.h. gleichartiges Spalten- und Zeilenumformen) dieser Matrix gelingt es (z.B. in dieser Reiherfolge) die Elemente $h_{r0} = h_{0r}, \ldots, h_{20} = h_{02}$ durch 0 zu ersetzen. Dabei werden nur (entsprechende) Vielfache der $r., \ldots, 2.$ Spalten bzw. Zeilen zur 0. Spalte bzw. Zeile dazuaddiert. Das Element h_{00} wird dadurch natürlich verändert. Schließlich kann noch, da $h_{01} = h_{10} \neq 0$ ist, durch Addition eines entsprechenden Vielfachen der 1. Spalte bzw. Zeile das (neue) Element h_{00} durch 0 ersetzt werden. Alle diese Operationen modifizieren nur den Vektor $\mathbf{b}_0$. Das (neue) System von Vektoren $\{\mathbf{b}_0, \mathbf{b}_0 + \mathbf{b}_1, \ldots, \mathbf{b}_0 + \mathbf{b}_n\}$ ist dann das gewünschte kartesische Koordinatensystem, so daß die quadratische Funktion λ die Gestalt

$$\lambda(\mathbf{a}) = 2h_{10}x_1 + g_2x_2^2 + \cdots + g_rx_r^2$$

hat. $g_2, \ldots, g_r$ sind die Eigenwerte von G.

Literaturverzeichnis

[1] E. Artin, *Analytische Geometrie und Algebra*, Universität Hamburg, 1960.

[2] J. Cigler, *Einführung in die Lineare Algebra und Geometrie*, Manz, Wien, 1977.

[3] W. H. Greub, *Linear Algebra*, Springer, Berlin, 1967.

[4] K. W. Gruenberg and A. J. Weir, *Linear Geometry*, Springer, New York, 1977.

[5] H. Havlicek, *Lineare Algebra und analytische Geometrie I*, Skriptum, TU Wien, 1995.

[6] H. Havlicek, *Lineare Algebra und analytische Geometrie II*, Skriptum, TU Wien, 1996.

[7] H. Mitsch, *Lineare Algebra und Geometrie I*, Prugg, Wien, 1978.

[8] H. Mitsch, *Lineare Algebra und Geometrie II*, Prugg, Wien, 1979.

Index