§44 Spektralzerlegung normaler Operatoren im endlichdimensionalen Fall

Zusammenfassung

Dieser Paragraf richtet sich im Aufbau weitgehend nach § 42, um den Zerlegungssatz (44.7) analog zum Satz über die Hauptachsentransformation (42.7) zu entwickeln. Statt eines n–dimensionalen euklidischen Vektorraumes (über $\mathbb{R}$) wird jetzt ein n–dimensionaler unitärer Vektorraum (über $\mathbb{C}$) zugrundegelegt. Entsprechend der allgemeineren Situation werden nicht nur selbstadjungierte sondern auch normale Operatoren studiert, insbesondere wird der Hauptsatz 44.7 für normale Operatoren bewiesen.

In diesem Paragrafen ist $(V, <\ ,\ >)$ stets ein n–dimensionaler unitärer Vektorraum, $n > 0$.

A. Dualität und Adjungierte

Für jeden Vektor $w \in V$ ist

$$j_w : V \to \mathbb{R},\ v \mapsto <w, v>,\ v \in V,$$

d.h. $j_w := <w,\ >$, eine Linearform über $\mathbb{C}$, also ein Element von $V^\star = \mathrm{Hom}(V, \mathbb{C})$. Das Skalarprodukt liefert daher eine Abbildung

$$j : V \to V^\star,\ w \mapsto j_w,\ w \in V.$$

(44.1) Satz (Riesz): Die Abbildung $j : V \to V^\star$ ist ein $\mathbb{R}$–Isomorphismus von Vektorräumen, und es gilt $j(\lambda w) = \bar{\lambda} j(w)$ für $\lambda \in \mathbb{C}$ und $w \in V$, d.h. j ist semilinear. (Eine $\mathbb{R}$–lineare Abbildung $g : U \to W$ zwischen komplexen Vektorräumen U, W heißt *semilinear*, wenn neben der $\mathbb{R}$–Linearität auch noch $g(iu) = -ig(u)$ für $u \in U$ und damit $g(\lambda u) = \bar{\lambda} g(u)$ für $\lambda \in \mathbb{C}$ gilt.)

Beweis: j ist $\mathbb{R}$–linear und semilinear. Denn es gilt

$$j(v + w) = <v + w,\ > = <v,\ > + <w,\ > = j(v) + j(w)$$

für $v, w \in V$; und für $\lambda \in \mathbb{C}$ ist $j(\lambda w) = <\lambda w, > = \bar{\lambda}<w, > = \bar{\lambda} j(w)$.

Im endlichdimensionalen Fall ist $\dim_{\mathbb{C}} V = \dim_{\mathbb{C}} V^\star$ (vgl. 19.6 oder 25.1) und dann auch $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} V^\star$, weil allgemein $\dim_{\mathbb{C}} W = 2 \dim_{\mathbb{R}} W$ für komplexe Vektorräume W endlicher Dimension gilt. Es genügt daher – nach dem Äquivalenzsatz 15.10 für lineare Abbildungen zwischen gleichdimensionalen Vektorräumen – zu zeigen, dass j injektiv ist. Seien dazu $w, w' \in V$ mit $j(w) = j(w')$ gegeben. Dann gilt insbesondere $j_w(w - w') = j'_w(w - w')$, also $<w, w - w'> = <w', w - w'>$ und damit $<w - w', w - w'> = 0$. Weil das Skalarprodukt positiv definit ist, folgt $w - w' = 0$, also $w = w'$.

Der Isomorphismus $j : V \to V^\star$ liefert zu jeder linearen Abbildung $f : V \to V$ eine Abbildung $f^\star : V \to V$ durch

$$<f(v), w> = <v, f^\star(w)>, \ \text{für alle} \quad v, w \in V.$$

Die Abbildung $f^\star$ lässt sich auch folgendermaßen verstehen: Für $w \in V$ ist $f^\star(w)$ derjenige Vektor, für den $j_{f^\star(w)}$ die Linearform $v \mapsto <w, f(v)>$ ist, also diese Linearform vermöge j darstellt (und wir wissen nach dem Satz von Riesz (44.1), dass jede Linearform μ sich als ein j_x mit $x \in V$ darstellen lässt). Um Existenz und Eindeutigkeit von $f^\star$ genauer zu beschreiben, vergleichen wir $f^\star$ mit der linearen Abbildung

$$\mathrm{Ad}(f) : V^\star \longrightarrow V^\star, \quad \mu \mapsto \mu \circ f,$$

die gelegentlich auch die *Adjungierte* von f genannt wird. $\mathrm{Ad}(f)$ ist in der Tat komplex-linear:

$$\mathrm{Ad}(f)(\mu + \nu) = (\mu + \nu) \circ f = \mu \circ f + \nu \circ f = \mathrm{Ad}(f)(\mu) + \mathrm{Ad}(f)(\nu)$$

und

$$\mathrm{Ad}(f)(\lambda\mu) = (\lambda\mu) \circ f = \lambda(\mu \circ f) = \lambda \mathrm{Ad}(f)(\mu).$$

(44.2) Definition–Satz: Sei $f \in \mathrm{Hom}(V, V)$.

1° $f^\star$ ist wohldefiniert und linear über $\mathbb{C}$. $f^\star$ heißt die *Adjungierte* zu f (bzw. der zu f *adjungierte Operator*). Es gilt: $j \circ f^\star = \mathrm{Ad}(f) \circ j$.

2° Die Abbildung $f \mapsto f^\star$, $f \in \mathrm{Hom}(V, V)$ ist ein $\mathbb{R}$–Isomorphismus

$$^\star : \mathrm{Hom}(V, V) \longrightarrow \mathrm{Hom}(V, V)$$

von Vektorräumen mit $(\lambda f)^\star = \bar{\lambda} f^\star$, das bedeutet $^\star(\lambda f) = \bar{\lambda}(^\star f)$. * ist also semilinearer Isomorphismus. Außerdem gilt $^{\star\star} = \mathrm{id}$, i.e. $(f^\star)^\star = f^{\star\star} = f$, sowie $(f \circ g)^\star = g^\star \circ f^\star$ für $f, g \in \mathrm{Hom}(V, V)$.

3° f heißt *selbstadjungiert*, wenn $f = f^\star$ gilt.

4° f heißt *normal*, wenn $f \circ f^\star = f \circ f^\star$ gilt.

Beweis: 1°: Für $\tilde{f} := j^{-1} \circ \mathrm{Ad}(f) \circ j : V \to V$,

$$\begin{tikzcd}
V \arrow[r, "\tilde{f}"] \arrow[d, "j"'] & V \arrow[d, "j"] \\
V^\star \arrow[r, "\mathrm{Ad}(f)"'] & V^\star
\end{tikzcd}$$

also $j \circ \tilde{f} = \mathrm{Ad}(f) \circ j$, gilt wegen $j_{\tilde{f}(w)} = \mathrm{Ad}(f)(j_w) = j_w \circ f$ für alle $w \in V$:

$$<\tilde{f}(w), v> = <w, f(v)> \quad \text{für alle} \quad v \in V.$$

Also ist $\tilde{f}$ die gesuchte Abbildung $f^\star$, d.h. $f^\star = \tilde{f}$ ist wohldefiniert und als Komposition von drei $\mathbb{R}$–linearen Abbildung auch $\mathbb{R}$–linear. Ferner ist

$$\begin{array}{rcl} f^\star(\lambda v) & = & j^{-1}\circ \mathrm{Ad}(f)\circ j(\lambda v) & = & j^{-1}\circ \mathrm{Ad}(f)(\bar\lambda j(v)) \\ & = & j^{-1}(\bar\lambda \mathrm{Ad}(f)(j(v))) & = & j^{-1}(\bar\lambda(\mathrm{Ad}(f)\circ j)(v)) \\ & = & \bar\lambda(j^{-1}\circ(\mathrm{Ad}(f)\circ j)(v)) & = & \bar\lambda f^\star(v) \end{array}$$

2° sieht man folgendermaßen: Zunächst ist die Abbildung

$$\mathrm{Ad} : \mathrm{Hom}(V,V) \longrightarrow \mathrm{Hom}(V,V),\ f \mapsto \mathrm{Ad}(f)$$

komplex-linear: Für $f, g \in \mathrm{Hom}(V,V)$ und $\lambda \in \mathbb{C}$ ist

$$\begin{array}{rcl} \mathrm{Ad}(f+\lambda g)(\mu) & = & \mu\circ(f+\lambda g) & = & \mu\circ f+\mu\circ(\lambda g) \\ & = & \mathrm{Ad}(f)(\mu)+\lambda\mathrm{Ad}(g)(\mu) & = & (\mathrm{Ad}(f)+\lambda\mathrm{Ad}(g))(\mu), \end{array}$$

also $\mathrm{Ad}(f+\lambda g) = \mathrm{Ad}(f)+\lambda\mathrm{Ad}(g)$. Daher gilt:

$$\begin{array}{rcl} (f+g)^\star & = & j^{-1}\circ \mathrm{Ad}(f+g)\circ j & = & j^{-1}\circ(\mathrm{Ad}(f)+\mathrm{Ad}(g))\circ j \\ & = & j^{-1}\circ\mathrm{Ad}(f)\circ j + j^{-1}\circ\mathrm{Ad}(g)\circ j & = & f^\star + g^\star \end{array}$$

sowie

$$\begin{array}{rcl} (\lambda f)^\star & = & j^{-1}\circ\mathrm{Ad}(\lambda f)\circ j & = & j^{-1}\circ\lambda\mathrm{Ad}(f)\circ j \\ & = & \bar\lambda(j^{-1}\circ\mathrm{Ad}(f)\circ j & = & \bar\lambda f^\star. \end{array}$$

Also ist * semilinear.

$(f^\star)^\star = f$ ergibt sich aus

$$<v,(f^\star)^\star(w)> = <f^\star(v),w> = \overline{<w,f^\star(v)>} = \overline{<f(w),v>} = <v,f(w)>.$$

Deshalb ist * bijektiv und ein semilinearer $\mathbb{R}$–Isomorphismus.

Schließlich folgt $(f\circ g)^\star = g^\star\circ f^\star$ aus $<v,(f\circ g)^\star(w)> = <(f\circ g)(v),w> = <g(v),f^\star(w)> = <v,g^\star\circ f^\star(w)>$.

Damit ist 44.2.2° bewiesen.

B. Normale Operatoren

(44.3)Bemerkungen und Beispiele:

1° Jeder selbstadjungierter Operater $f \in \mathrm{Hom}(V,V)$ ist normal, denn $f = f^\star$ impliziert unmittelbar $f\circ f^\star = f^\star\circ f$.

2° Auch jeder unitäre Operator $f \in \mathrm{Hom}(V,V)$ ist normal, denn es gilt $f\circ f^\star = \mathrm{id}_V = f^\star\circ f$.

3° Bekanntlich liefert eine unitäre Matrix $U \in \mathbb{C}^{n\times n}$ in Bezug auf des Standard-Skalarprodukt (vgl. 43.2.1°) eine unitäre und daher normale lineare Abbildung $\mathbb{C}^n \to \mathbb{C}^n$ durch $z \mapsto Uz$.

4° Eine *hermitesche Matrix* $A \in \mathbb{C}^{n\times n}$ (d.h. definitionsgemäß $A = \bar{A}^\top =: A^\star$) liefert ganz entsprechend eine selbstadjungierte lineare Abbildung $\mathbb{C}^n \to \mathbb{C}^n$ durch

$z \mapsto Az$.

5° Schließlich liefert eine *normale Matrix* $A \in \mathbb{C}^{n\times n}$ (d.h. definitionsgemäß $AA^\star = A^\star A$) eine normale lineare Abbildung $\mathbb{C}^n \to \mathbb{C}^n$ durch $z \mapsto Az$.

6° Eine Diagonalmatrix $D = \operatorname{diag}(\alpha_1, \alpha_2, \dots, \alpha_n), \alpha_i \in \mathbb{C}$ liefert auf diese Weise stets eine normale lineare Abbildung, denn $D^\star = \operatorname{diag}(\bar\alpha_1, \bar\alpha_2, \dots, \bar\alpha_n)$ und $DD^\star = \operatorname{diag}(\alpha_1\bar\alpha_1, \alpha_2\bar\alpha_2, \dots, \alpha_n\bar\alpha_n) = D^\star D$. Eine solche Abbildung ist aber im allgemeinen weder unitär (nämlich dann, wenn nicht alle α_i die Bedingung $\alpha_i\bar\alpha \neq 1$) noch selbstadjungiert (nämlich dann, wenn nicht alle α_i reell sind).

7° Für ein lineare Abbildung $f \in \operatorname{Hom}(V,V)$ sei $A = A_f$ die darstellende Matrix bezüglich einer ONB. f ist genau dann selbstadjungiert, wenn A hermitesch ist: $A = A^\star$, $(A^\star := \bar{A}^\star)$.

8° Unter denselben Voraussetzungen: f ist genau dann normal, wenn A normal ist: $AA^\star = A^\star A$.

Beweis: (Zu 7°, 8°) Sei $B = (b_1, b_2, \dots, b_n)$ ONB (nach Voraussetzung ist V ja endlichdimensional). Die darstellende Matrix hat die Koeffizienten $A_i^\nu \in \mathbb{C}$, die durch $f(b_i) = A_i^\nu b_\nu$ gegeben sind. Es gilt $<f(b_i), b_j> = <A_i^\nu b_\nu, b_j> = \bar{A}_i^\nu <b_\nu, b_j> = \bar{A}^j{}_i$ und $<f(b_i), b_j> = <b_i, f^\star(b_j)>$, also $f^\star(b_j) = \sum_{i=1}^n \bar{A}^j{}_i b_i$. Deshalb hat die darstellende Matrix von $f^\star$ die Koeffizienten $\bar{A}^j{}_i$, ist also gerade $A^\star$. Daraus folgen die Behauptungen 7° und 8° unmittelbar.

(44.4) Satz: Sei $f \in \operatorname{Hom}(V,V)$ selbstadjungiert. Dann sind die Eigenvektoren von f zu verschiedenen Eigenwerten orthogonal zueinander.

Beweis: Sei $f(v) = \lambda v, f(w) = \mu w$ für Vektoren $v, w \in V \setminus \{0\}$. Aus $<f(v), w> = <v, f(w)>$ folgt $<\lambda v, w> = <v, \mu w>$, also $(\bar\lambda - \mu)<v,w>$. Daraus folgt die Behauptung.

(44.5) Satz: Sei $f \in \operatorname{Hom}(V,V)$ normal.

1° Es gilt $<f(v), f(w)> = <f^\star(v), f^\star(w)>$ für alle $v, w \in V$. (Und diese Eigenschaft charakterisiert die normalen Operatoren.)

2° Es gilt $\operatorname{Ker} f = \operatorname{Ker} f^\star$ und $\operatorname{Im} f = \operatorname{Im} f^\star$, und man hat die orthogonale Zerlegung:

$$V = \operatorname{Ker}(f) \oplus \operatorname{Im}(f) \quad \text{mit} \quad \operatorname{Ker}(f) \perp \operatorname{Im}(f).$$

3° Die Eigenräume von f und $f^\star$ stimmen überein und es gilt für jeden Eigenwert $\lambda \in \mathbb{C}$ von $f : \bar\lambda$ ist Eigenwert von $f^\star$.

Beweis: Zu 1°: $<f(v), f(w)> = <v, f^\star(f(w))> = <v, f(f^\star(w))>$, da f normal ist, und $<v, f(f^\star(w))> = <f(f^\star(w)), v> = <f^\star(w), f^\star(v)> = <f^\star(v), f^\star(w)>$.

Zu 2°: Aus 1° folgt zunächst $\operatorname{Ker} f = \operatorname{Ker} f^\star$. Aus dem nachfolgenden Hilfssatz lässt sich daher $\operatorname{Ker} f = (\operatorname{Im} f)^\perp$ ablesen, und wir haben die gewünschte orthogonale Zerlegung: $\operatorname{Ker} f \perp \operatorname{Im} f$, also ist $\operatorname{Ker} f + \operatorname{Im} f$ direkte Summe: $\operatorname{Ker} f + \operatorname{Im} f = \operatorname{Ker} f \oplus \operatorname{Im} f$. Außerdem gilt $\operatorname{Ker} f \oplus \operatorname{Im} f = V$ nach der Dimensionsformel. Schließlich lässt sich daraus mit Hilfe des Hilfssatzes auch noch $\operatorname{Im} f = \operatorname{Im} f^\star$ ablesen.

Zu 3°: Für $\lambda \in \mathbb{C}$ ist $(f - \lambda\mathrm{id})^\star = f^\star - \bar{\lambda}\mathrm{id}$, wie man leicht nachprüft. Es gilt daher

$$(f - \lambda\mathrm{id})^\star \circ (f - \lambda\mathrm{id}) = f^\star \circ f - \bar{\lambda} f - \lambda f^\star + \bar{\lambda}\lambda$$
$$(f - \lambda\mathrm{id}) \circ (f - \lambda\mathrm{id})^\star = f \circ f^\star - \lambda f^\star - \bar{\lambda} f + \bar{\lambda}\lambda,$$

und es folgt: Weil f normal ist, ist auch $f - \lambda\mathrm{id}$ normal. 2° angewandt auf diesen normalen Operater liefert die Aussage 3°: Denn $\mathrm{Ker}(f - \lambda\mathrm{id}) = \mathrm{Ker}(f^\star - \bar{\lambda}\mathrm{id})$.

(44.5.4°) Hilfssatz: Für Operatoren $f \in \mathrm{Hom}(V, V)$ gilt allgemein:

$$\mathrm{Ker} f^\star = (\mathrm{Im} f)^\perp \quad \text{und ebenso} \quad \mathrm{Im} f^\star = (\mathrm{Ker} f)^\perp.$$

Beweis: Denn $w \in (\mathrm{Im})^\perp \iff 0 = <f(v), w> = <v, f^\star(w)>\ \forall v \in V \iff f^\star(w) = 0$, also $\mathrm{Ker} f^\star = (\mathrm{Im} f)^\perp$. Daher auch $(\mathrm{Im} f^\star)^\perp = \mathrm{Ker} f^{\star\star} = \mathrm{Ker} f$ und dann $(\mathrm{Im} f^\star) = (\mathrm{Ker} f)^\perp$.

(44.6) Satz: Sei $f \in \mathrm{Hom}(V, V)$ normal. Für jeden Eigenvektor v von f ist $v^\perp = \{v\}^\perp$ f–invariant, d.h. $f(v^\perp) \subset v^\perp$.

Beweis: Sei $f(\lambda v) = \lambda f(v)$. Für alle $w \in v^\perp$ ist nach Definition $<v, w> = 0 = <w, v>$, also gilt $<f(w), v> = <w, f^\star(v)> = <w, \bar{\lambda} v> = \bar{\lambda}<w, v> = 0$, daher $f(w) \in v^\perp$.

C. Spektralzerlegung normaler Operatoren

(44.7) Satz: Ein Operator $f \in \mathrm{Hom}(V, V)$ ist genau dann normal, wenn es eine Orthonormalbasis $B = (b_1, b_2, \ldots, b_n)$ von V gibt, die aus lauter Eigenvektoren von f besteht. f hat also bezüglich dieser Basis die Matrixdarstellung

$$\boxed{A_f = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)}$$

mit den Eigenwerten $(\lambda_1, \lambda_2, \ldots, \lambda_n)$.

Insbesondere hat f die folgende Entwicklung

$$\boxed{f(v) = \sum_{i=1}^{n} \lambda_i <b_i, v> b_i}$$

(vgl. 44.8).

Beweis: Wir wissen bereits, dass jede Diagonalmatrix einen normalen Operator definiert (vgl. 44.3.6°).

Die Umkehrung sieht man durch Induktion nach $n = \dim V$ wie in § 42:

Im Falle $n = 1$ gibt es so gut wie nichts zu zeigen: Bezüglich jeder ONB (b), d.h. $\|b\| = 1$ hat jede lineare Abbildung Diagonalform: $f(sb) = \lambda sb$, $s \in \mathbb{C}$, wobei $\lambda = f(b)$.

Im Falle $n > 1$ gibt es aufgrund des Fundamentalsatzes der Algebra eine Nullstelle des des charakteristischen Polynoms χ_f, also einen Eigenwert $\lambda_1 \in \mathbb{C}$ und dazu einen Eigenvektor $v \in V$. Setze $b_1 := v/\|v\|$ und $U = v^\perp = (b_1)^\perp$. Es gilt dann $U \oplus U^\perp = V$ und $U^\perp = \mathbb{C} b_1$. Also ist $U^\perp$ eindimensional und es folgt $\dim U = n - 1$

nach der Dimensionsformel. Nach 44.6 ist $f(U) \subset U$, deshalb kann man die Restriktion $g := f|_U$ von f auf U als eine lineare Abbildung von U nach U definieren. Natürlich ist g normal, weil f normal ist. Wegen $\dim U = n-1$ hat U nach Induktionsvoraussetzung eine ONB $(b_2, \ldots, b_n)$ aus Eigenvektoren von g, die natürlich auch Eigenvektoren von f sind. Wegen $b_1 \in U^\perp$ ist jetzt $(b_1, b_2, \ldots, b_n)$ eine ONB aus Eigenvektoren von f, und der Satz ist bewiesen.

(44.8) Bemerkung:

1° Wir haben wieder eine „Hauptachsentransformation"(vgl. 42.8), die allerdings im komplexen Fall nicht als solche betrachtet wird, bestenfalls als die natürliche Verallgemeinerung vom Reellen zum Komplexen. Es sei eine Quadrik $Q = q^{-1}(0)$ in $\mathbb{C}^n$ mit Standard-Skalarprodukt und Standard-Orthonormalbasis durch eine quadratische Form $q(z) = \bar{z}^\top Az$ zu einer hermiteschen Matrix $A \in \mathbb{C}^{n\times n}$ gegeben. Die Aussage des Satzes bedeutet, dass einen Koordinatenwechsel innerhalb der Orthonormalbasen so gibt, dass diese Quadrik Q in den neuen Koordinaten $\tilde{z}$ die (*Normal*-) Form

$$\boxed{Q = \{\tilde{z} : \sum_{i=1}^{i=n} \lambda_i \left|\tilde{z}^i\right|^2 = 0\}}$$

hat.

2° Unter den Voraussetzungen des Satzes 44.7 seien $\lambda_1, \lambda_2, \ldots, \lambda_k$, $k \leq n$, die paarweise verschiedenen Eigenwerte von f und es seien $E_j := \mathrm{Ker}(f - \lambda_j \mathrm{id})$ die zugehörigen Eigenräume. Sei $P_j : V \to V$ die nach 43.13 existierende und eindeutig bestimmte orthogonale Projektion auf E_j, also $P_j \circ P_j = P_j$, P_j ist selbstadjungiert und es gilt $E_j = \mathrm{Im} P_j$. Dann hat f die Darstellung

$$\boxed{f = \sum_{j=1}^{j=k} \lambda_j P_j.}$$

Man nennt diese Darstellung auch die „*Spektralzerlegung*" von f, denn das *Spektrum* von f ist in dieser (endlichdimensionalen) Situation die Menge $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$. Wir kommen auf Spektralzerlegungen in Paragraf 46 (der unendlichdimensionale Fall) zurück.

(44.9) Korollar: Zu einer normalen $n \times n$–Matrix $A \in \mathbb{C}^{n\times n}$ gibt es stets eine unitäre Matrix $U \in \mathrm{U}(n)$, so dass $UAU^\star$ eine Diagonalmatrix ist.

Das ist genau die Aussage des Satzes 44.7 angewandt auf die durch die Matrix A gegebene normale lineare Abbildung von $\mathbb{C}^n$ nach $\mathbb{C}^n$.