Mehrparametrige und nichtlineare Eigenwertaufgaben

K. P. HADELER

Vorgelegt von L. COLLATZ

Die einfachste Form einer Spektraltheorie ergibt sich für eine reelle symmetrische Matrix der endlichen Ordnung n. Es gibt n reelle linear unabhängige, sogar orthonormale Eigenvektoren, die zugehörigen Eigenwerte sind reell. Damit erhält man einen Entwicklungssatz sowie Variationsprinzipien zur Charakterisierung der Eigenwerte (Fischer-Courantsches und Rayleighsches Prinzip). Die Eigenvektoren sind stationäre Punkte eines gewissen Funktionals (des Rayleigh-Quotienten), der Wertebereich dieses Funktionals ist der sog. Wertebereich ("numerical range") der Matrix. Die Eigenwerte können durch Iterationsverfahren, die eng mit den Variationsprinzipien zusammenhängen (v. Mises-Verfahren), approximiert werden.

Bei den mannigfachen Verallgemeinerungen dieser Eigenwertaufgabe, die den Inhalt der Spektraltheorie der linearen Operatoren ausmachen (bis zu unbeschränkten Operatoren in Banachräumen), bleibt das Konzept der Operatorfunktion $T(\lambda)=\lambda E-A$ erhalten, die linear von einem komplexen Parameter λ und einem einzigen Operator A abhängt (mit Ausnahme einiger Aufgabenklassen bei endlichen Matrizen und besonders bei gewöhnlichen Differentialgleichungen, wo zwei von E verschiedene Operatoren eingehen).

Aufgaben mit zwei Parametern wurden von SCHÄFKE [*8*] beim Studium der Mathieuschen Gleichung herangezogen, Aufgaben mit mehreren Parametern wurden von SCHAEFER [*11*] und einer Reihe anderer Autoren (vgl. [*4*, *12*]) zur Lösung technischer Probleme benutzt. Unter den Aufgaben mit einem nichtlinear auftretenden Parameter wurden vor allem polynomiale Büschel, insbesondere vom Grade 2 untersucht. Solche Probleme lassen sich mit Hilfe der „Frobeniusschen Begleitmatrix" in lineare überführen. Wir erwähnen vor allem Arbeiten von WIELANDT (vgl. [*9*]), MÜLLER [*9*], DUFFIN [*2*], KREIN und Schülern [*5*]. Allgemeinere nichtlineare Abhängigkeiten wurden von LANCASTER [*7*], ROGERS [*10*] untersucht. Eine Reihe von Eigenwertaufgaben, die vom Normaltyp abweichen, sind als Beispiele im Buche von COLLATZ [*1*] angeführt.

In dieser Arbeit wird die Frage untersucht, ob mehrparametrige und nichtlineare Operatorfunktionen bzw. Eigenwertaufgaben unter gewissen Symmetrie- und Beschränktheits-Voraussetzungen Eigenschaften aufweisen, die denen der gewöhnlichen einparametrigen linearen Funktion $T(\lambda)=\lambda E-A$ analog sind. Als solche Eigenschaften wollen wir vor allem die verstehen, die eingangs für eine symmetrische Matrix angeführt wurden. Es zeigt sich, daß die Frage positiv beantwortet werden kann.

Wenn wir für die mehrparametrige Funktion

$$F(\lambda)=\sum_{j=1}^{m}\lambda_j B_j - B_0,$$

deren Koeffizienten B_j symmetrische beschränkte Operatoren auf einem Hilbertraum und deren Parameter λ_j reell sind, das Spektrum $\sigma(F)$ und den Wertebereich $W(F)$ in geeigneter Weise definieren (Nr. 2), so ergeben sich unter gewissen Voraussetzungen eine Konvexitätseigenschaft von $W(F)$ (Nr. 3) sowie die Relationen $\overline{W(F)}\supset\sigma(F)$, $\partial W(F)\subset\sigma(F)$ (Nr. 4). Nach der Beschaffenheit von $W(F)$ kann eine Klassifikation der Funktionen F in hyperbolische, parabolische und elliptische vorgenommen werden (Nr. 5). Die hyperbolischen Funktionen stehen in engem Zusammenhang mit nichtlinearen Eigenwertaufgaben. Durch Dualisierung (Nr. 6) gelangt man zu dem von HAUSDORFF [*6*] betrachteten Wertebereich einer Bilinearform.

Bei den nichtlinearen (einparametrigen) Aufgaben

$$T(\lambda)x=0, \qquad x\neq 0$$

beschränken wir uns auf den Fall endlicher Raumdimension n. Hier liegt eine Theorie von ROGERS [*10*] vor, nach der unter gewissen Voraussetzungen die Existenz von genau n linear unabhängigen Eigenvektoren und zugehörigen reellen Eigenwerten gesichert ist. ROGERS zeigt auch eine Verallgemeinerung des Courantschen Minimum-Maximum-Prinzips auf den Fall der nichtlinearen Aufgabe. Wir stellen in Nr. 10 die Ergebnisse von ROGERS (ohne Beweise) dar. In Nr. 11 zeigen wir eine Orthogonalitätseigenschaft der Eigenvektoren mit Hilfe eines verallgemeinerten inneren Produktes, das zwar definit, symmetrisch und homogen, aber im allgemeinen nicht bilinear ist. Eine ähnliche Bildung wurde vor kurzem auch von TURNER [*14*] bei einem speziellen quadratischen Büschel verwendet. In Nr. 12 übertragen wir den Gram-Schmidtschen Orthogonalisierungsprozeß auf dieses innere Produkt. Dann können wir ein dem Rayleighschen Prinzip analoges Minimumprinzip zeigen. Zur Berechnung von Eigenwerten konstruieren wir einige der v. Misesschen Methode verwandte Verfahren und erläutern ihren Zusammenhang mit den Newton- und Gradienten-Verfahren (Nr. 14–17). Wir diskutieren den Spezialfall der polynomialen Büschel (Nr. 18) und bringen einige numerische Beispiele (Nr. 19).

1. Vorbetrachtung

Sei H ein komplexer Hilbertraum mit dem inneren Produkt $(\ ,\ )$ und der Norm $\|\ \|$. Sei $\mathfrak{A}$ die Algebra der beschränkten linearen Operatoren auf H, $\mathfrak{H}$ der reelle Teilraum der symmetrischen Operatoren.

Sei $A\in\mathfrak{A}$. Das Spektrum $\sigma(A)$ ist die Menge aller Punkte λ der komplexen Ebene $\mathbb{C}$, für die die Resolvente $R_\lambda=(\lambda E-A)^{-1}$ nicht existiert oder nicht beschränkt ist oder einen in H nicht dichten Wertebereich besitzt. Durch

$$p(x)=(Ax,x)/(x,x) \tag{1.1}$$

wird ein komplexwertiges Funktional auf $H-\{0\}$ erklärt. Der Wertebereich (Wertevorrat, "numerical range", eigentlich der Wertebereich von p) $W(A)$ des

Operators A ist definiert als

$$W(A)=\{\lambda\colon \exists\, x\in H-\{0\} \text{ mit } \lambda=p(x)\}. \tag{1.2}$$

Es gilt [13, 6]

$$\sigma(A) \text{ ist beschränkt und abgeschlossen}, \tag{1.3}$$

$$W(A) \text{ ist konvex}, \tag{1.4}$$

$$\sigma(A)\subset\overline{W(A)}. \tag{1.5}$$

Ist $A\in\mathfrak{H}$, so ist p reellwertig, $W(A)$ und $\sigma(A)$ sind Teilmengen der reellen Achse, $W(A)$ ist ein Intervall, für dessen Rand $\partial W(A)$ gilt

$$\partial W(A)\subset\sigma(A). \tag{1.6}$$

Jetzt sei H endlichdimensional. Ist $A\in\mathfrak{H}$, so lassen sich die durch

$$\lambda_1\leqq\lambda_2\leqq\cdots\leqq\lambda_n$$

geordneten Eigenwerte von A durch das Courantsche Minimum-Maximum-Prinzip

$$\lambda_j=\min_{R^j\subset H}\ \max_{x\in R^j-\{0\}} p(x), \qquad (j=1,\ldots,n), \tag{1.7}$$

definieren. Dabei bezeichnet R^j einen beliebigen Teilraum der Dimension j.

Sei x_j, $Ax_j=\lambda_j x_j$ $(j=1,\ldots,n)$, ein fest gewähltes Orthonormalsystem von Eigenvektoren. Dann können die Eigenwerte durch das Rayleighsche Extremalprinzip rekursiv definiert werden

$$\lambda_j=\min_{\substack{(x,x_k)=0\\ x\neq 0,\ k=1,\ldots,j-1}} p(x), \qquad (j=1,\ldots,n). \tag{1.8}$$

Ist A positiv definit, so konvergiert die Iteration

$$\tilde{z}_0\neq 0, \qquad z_\nu=\tilde{z}_\nu/\|\tilde{z}_\nu\|, \qquad \tilde{z}_{\nu+1}=A z_\nu, \qquad (\nu=0,1,\ldots)$$

gegen einen Eigenvektor von A. Die Folge der $p(z_\nu)$ konvergiert monoton von unten gegen den zugehörigen Eigenwert.

Die Verallgemeinerungen dieser Sätze auf vollstetige Operatoren in nicht endlichdimensionalen Räumen werden wir im folgenden nicht benötigen.

2. Mehrparametrige Aufgaben

Die Dimension von H sei jetzt beliebig. Durch

$$F(\lambda)=\sum_{j=1}^{m}\lambda_j B_j-B_0, \qquad \lambda=(\lambda_1,\ldots,\lambda_m)\in R^m, \tag{2.1}$$

$$B_j\in\mathfrak{H} \qquad (j=0,\ldots,m), \tag{2.2}$$

wird eine Abbildung des reellen R^m in $\mathfrak{H}$ definiert. Wir nennen F eine mehrparametrige Operatorfunktion (gelegentlich, z.B. im Falle endlicher Raumdimension, sprechen wir auch von mehrparametrigen Eigenwertaufgaben).

Für $m=1$, $B_1=E$, $\lambda_1=\lambda$ erhalten wir die Funktion

$$F(\lambda)=\lambda E-B_0, \tag{2.3}$$

deren Untersuchung der Spektralaufgabe des Operators B_0 äquivalent ist. (2.1) ist also eine Verallgemeinerung der gewöhnlichen „Eigenwertaufgabe".

Wir definieren die Teilmengen des reellen R^m

$$\sigma(F)=\{\lambda: 0\in\sigma(F(\lambda))\}, \tag{2.4}$$

$$\sigma_p(F)=\{\lambda: 0\in\sigma_p(F(\lambda))\}, \tag{2.5}$$

$$\sigma_c(F)=\{\lambda: 0\in\sigma_c(F(\lambda))\}. \tag{2.6}$$

Wir nennen $\sigma(F)$ das Spektrum, $\sigma_p(F)$ das Punktspektrum und $\sigma_c(F)$ das kontinuierliche Spektrum von F. Die Punkte von $\sigma(F)$ heißen Spektralwerte, die von $\sigma_p(F)$ Eigenwerte von F. Für ein festes $\lambda\in R^m$ ist $F(\lambda)\in\mathfrak{H}$, also besitzt $F(\lambda)$ kein Restspektrum. Daher gilt

$$\sigma_c(F)\cup\sigma_p(F)=\sigma(F). \tag{2.7}$$

Sei $\lambda_{(j)}\in\sigma(F)$ $(j=1, 2, \ldots)$, $\lambda_{(j)}\to\lambda$. Dann ist $0\in(F(\lambda_{(j)}))$ und $\|F(\lambda_{(j)})-F(\lambda)\|\to 0$, folglich $0\in\sigma(F(\lambda))$. Also ist $\sigma(F)$ abgeschlossen.

Sei $x\in H-\{0\}$. Durch

$$\varphi_x(\lambda)=(F(\lambda)x, x) \tag{2.8}$$

ist eine Abbildung des R^m in den R^1 definiert. Im allgemeinen ist

$$W_x=\{\lambda: \varphi_x(\lambda)=0\} \tag{2.9}$$

eine Hyperebene im R^m. Wir bezeichnen W_x als Rayleigh-Hyperebene des Vektors x. Diese Benennung soll auch dann gelten, wenn W_x leer oder gleich dem ganzen R^m ist. Die Menge

$$W(F)=\bigcup_{x\in H-\{0\}} W_x \tag{2.10}$$

heiße Wertebereich der Funktion F, der topologische Abschluß $\overline{W(F)}$ heiße abgeschlossener Wertebereich.

Im Falle der Funktion (2.3) ist $\sigma(F)=\sigma(B_0)$, $W(F)=W(B_0)$, die Definitionen stimmen also mit den üblichen überein. Im folgenden sollen die Mengen $\sigma(F)$ und $W(F)$ sowie ihre Beziehung zueinander untersucht werden.

3. Eigenschaften des Wertebereichs

Wir definieren die Mengen

$$W_+ \doteq \{\lambda: \varphi_x(\lambda)>0 \text{ für alle } x\in H-\{0\}\}, \tag{3.1}$$

$$W_- \doteq \{\lambda: \varphi_x(\lambda)<0 \text{ für alle } x\in H-\{0\}\}. \tag{3.2}$$

Bei vorgegebenem F kann jede dieser Mengen leer sein. Wir zeigen

Lemma 3.1. $W(F)\cup W_+\cup W_-=R^m$.

Beweis. Sei $\lambda \in R^m$. Wenn es $x \neq 0$ mit $\varphi_x(\lambda)=0$ gibt, so ist $\lambda \in W(F)$. Wenn $\varphi_x(\lambda)>0$ für alle $x \neq 0$ oder $\varphi_x(\lambda)<0$ für alle $x \neq 0$ ist, so ist $\lambda \in W_+$ bzw. $\lambda \in W_-$. Es bleibt der Fall, daß es ein Paar $x, y \in H-\{0\}$ mit $\varphi_x(\lambda)>0$, $\varphi_y(\lambda)<0$ gibt. Für $F(\lambda) \in \mathfrak{H}$ gilt $(F(\lambda)\,x,x)>0$, $(F(\lambda)\,y,y)<0$. Also ist F nicht definit, wegen der Konvexität von $W(F(\lambda))$ gibt es $z \neq 0$ mit $(F(\lambda)z,z)=0$.

Lemma 3.2. *Wenn eine der Mengen W_+ oder W_- nicht leer ist, so ist sie konvex.*

Beweis. Für ein festes $x \in H-\{0\}$ ist $W_x^+=\{\lambda\colon \varphi_x(\lambda)>0\}$ leer oder gleich dem ganzen R^m oder ein offener Halbraum. Dann ist

$$W_+ = \bigcap_{x \in H-\{0\}} W_x^+$$

leer oder konvex als Durchschnitt konvexer Mengen.

Sei $m=1$. Seien $W(F)$, $W_+, W_- \neq \emptyset$. Sei $\lambda \in W(F)$. Es gibt $x \neq 0$ mit $\varphi_x(\lambda)=\lambda(B_1x,x)-(B_0x,x)=0$. Aus $(B_1x,x)=0$ folgt $(B_0x,x)=0$ und damit $\varphi_x \equiv 0$, $W(F)=R^1$. Daher ist $(B_1x,x) \neq 0$. φ_x besitzt genau eine Nullstelle, nämlich λ. Ist $\lambda^+ \in W_+$, $\lambda^- \in W_-$, $\varphi_x(\lambda^+)>0$, $\varphi_x(\lambda^-)<0$, so folgt $\lambda^-<\lambda<\lambda^+$ oder $\lambda^+<\lambda<\lambda^-$. $W(F)$ liegt also „zwischen" W_+, W_-. Mit den Lemmata 3.1, 3.2 folgt

Lemma 3.3. *Sei $m=1$. $W(F)$ erfüllt eine der drei Bedingungen* 1) $W(F)=\emptyset$, 2) *$W(F)$ ist konvex,* 3) *es gibt zwei Halbgeraden h_1, h_2 mit $h_1 \cap h_2=\emptyset$, $W(F)=h_1 \cup h_2 \neq R^1$.*

Anders gewendet: $W(F)$ ist auf dem „Torus" $R^1 \cup \{\infty\}$ zusammenhängend. Sei nun m wieder beliebig.

Lemma 3.4. *Wenn eine der Mengen W_+, W_- beschränkt und nicht leer ist, so ist die andere leer.*

Beweis. Sei $m \geqq 2$. Seien W_+, W_- nicht leer, sei etwa W_+ beschränkt. Sei $\lambda^+ \in W_+$, $\lambda^- \in W_-$ und

$$\lambda_j = a_j + \alpha\, b_j \qquad (j=1,\dots,m),\ \alpha \in R^1$$

eine Parameterdarstellung der Geraden g durch λ^-, λ^+. Mit der neuen Variablen α erhalten wir für die Restriktion von F auf g die Darstellung

$$\tilde{F}(\alpha)=F(\lambda)=\alpha \sum_{j=1}^{m} b_j B_j - \left(\sum_{j=1}^{m} a_j B_j - B_0\right) \mathrel{\dot{=}} \alpha \tilde{B}_1 - \tilde{B}_0 .$$

Die zu $\tilde{F}$ gehörenden Mengen $W(\tilde{F})$, $\tilde{W}_+$, $\tilde{W}_-$ sind die Bilder von $W(F) \cap g$, $W_+ \cap g$, $W_- \cap g$ bei der Abbildung $\lambda \to \alpha$. Also sind $\tilde{W}_+$, $\tilde{W}_-$ nicht leer und $\tilde{W}_+$ ist beschränkt. Nach Lemma 3.3 ist das unmöglich.

4. Wertebereich und Spektrum

Lemma 4.1. *Auf dem R^m sei eine Folge von Linearformen f_k $(k=1,2,\dots)$, gegeben, die gegen eine Form $f \neq 0$ konvergiert. Sei $\tilde{\lambda} \in R^m$ und $f(\tilde{\lambda})=c$. Dann gibt es eine Folge von Punkten $\lambda^{(k)}$ $(k=k_0,\ k_0+1,\ \dots)$, mit*

$$f_k(\lambda^{(k)})=c \quad \textit{und} \quad \lim_{k\to\infty} \lambda^{(k)}=\tilde{\lambda}.$$

Beweis. Es gibt k_0, so daß $f_k \neq 0$ für alle $k \geqq k_0$ ist. Der Abstand der Hyperebene $f_k(\lambda)=c$ von $\tilde{\lambda}$ wird durch $|f_k(\tilde{\lambda})-c|/\|f_k\|$ gegeben. Es ist

$$|f_k(\tilde{\lambda})-c| \leqq |f_k(\tilde{\lambda})-f(\tilde{\lambda})|+|f(\tilde{\lambda})-c| \leqq \|f-f_k\| \cdot \|\tilde{\lambda}\|.$$

Also streben die Abstände der Hyperebenen von $\tilde{\lambda}$ gegen Null. Hier und beim Beweis des Satzes 4.1 bezeichnet $\| \; \|$ die euklidische Norm des R^m.

Sei H nicht endlichdimensional. Sei $m=1$ und $B_0=B_1$ vollstetig und (strikt) positiv definit. Dann gilt $0 \in \sigma(B_0-\lambda B_1)$ für jedes $\lambda \in R^1$, also ist $\sigma(F)=R^1$. Andererseits ist $W_x=\{1\}$ für jedes $x \in H-\{0\}$, folglich $W(F)=\{1\}$. Dies Beispiel zeigt, daß die Voraussetzungen noch keine vernünftige Relation (etwa im Sinne einer Inklusion) zwischen Wertebereich und Spektrum garantieren.

Für den folgenden Satz benötigen wir die gleichmäßige Definitheit von B_0, d.h.

Es gibt $\alpha>0$ *mit* $(B_0 x, x) \geqq \alpha(x,x)$ *für alle* $x \in H$. (4.1)

Satz 4.1. *Sei* (4.1) *erfüllt. Dann gilt* $\sigma(F) \subset \overline{W(F)}$.

Beweis. Sei $\lambda \in \sigma(F)$. Wegen (4.1) ist $\lambda \neq 0$, daher $c := \|\lambda\| > 0$. $\lambda \in \sigma(F) \Leftrightarrow 0 \in \sigma(F(\lambda))$. Also gibt es eine Folge $x_k \in H$ $(k=1,2,\ldots)$, mit

$$|(F(\lambda)x_k, x_k)| < c/k, \qquad (B_0 x_k, x_k)=c,$$

$$\left|1-\sum_{j=1}^{m} b_{jk}\lambda_j/\|\lambda\|\right| < 1/k, \qquad b_{jk}=(B_j x_k, x_k) \quad \begin{matrix}(j=1,\ldots,m)\\(k=1,2,\ldots).\end{matrix}$$

Der Abstand des Punktes $b_k := (b_{jk}) \in R^m$ von der Hyperebene

$$1=\sum_{j=1}^{m} \beta_j \lambda_j/\|\lambda\|$$

ist also kleiner als $1/k$. Nach (4.1) ist

$$|b_{jk}|=|(B_j x_k, x_k)| \leqq \|B_j\| \cdot (x_k, x_k) \leqq \|B_j\| \cdot c/\alpha.$$

Die Punkte b_k liegen also alle in einem endlichen Quader, es gibt eine Teilfolge b_{k_l} mit $b_{k_l} \to \tilde{b}$ und

$$1=\sum_{j=1}^{m} \tilde{b}_j \lambda_j/\|\lambda\| \;\Rightarrow\; c=\sum_{j=1}^{m} \tilde{b}_j \lambda_j.$$

Lemma 4.1 liefert eine Teilfolge $\lambda^{(s)}$ mit

$$c=\sum_{j=1}^{m} b_{js}\lambda_j^{(s)} \qquad (s=1,2,\ldots), \;\; \lambda^{(s)} \to \lambda.$$

Diese $\lambda^{(s)}$ liegen alle in $W(F)$, also λ in $\overline{W(F)}$.

Sei $\partial W(F)$ der Rand von $W(F)$. Für den folgenden Satz benötigen wir die Voraussetzung (4.1) nicht.

Satz 4.2. *Es ist* $\partial W(F) \subset \sigma(F)$.

Beweis. I. Sei $m=1$, also $F(\lambda)=\lambda B_1-B_0$. Sei etwa $W_+ \neq \emptyset$ und $\lambda_0 \in \overline{W}_+ \cap \overline{W(F)}$. Wir können annehmen, daß $\lambda_0=0$ ist. Bis auf Spiegelung sind folgende drei Fälle möglich.

$$\begin{array}{ccc}
 & W & W_+ \\
W_+ \qquad W & W_+ \ \downarrow \ W_- & W \ \downarrow \ W \\
\text{———+———}, & \text{———+———}, & \text{———+———}. \\
0 & 0 & 0
\end{array}$$

Fall 1. Für $\lambda < 0$ ist $F(\lambda)$ positiv definit. $F(0)$ ist als gleichmäßiger Limes positiv definiter Operatoren positiv semidefinit. Für $\lambda > 0$ gibt es x_λ, $\|x_\lambda\| = 1$ mit $(F(\lambda)\, x_\lambda, x_\lambda) = 0$, also $0 \in W(F(\lambda))$. Folglich ist $0 \in \overline{W(F(0))}$ und damit $0 \in \sigma(F(0))$.

Fall 2. Wie bei Fall 1 erkennt man, daß $F(0)$ positiv semidefinit ist, ebenso negativ semidefinit, also $F(0) = 0$.

Fall 3. Für alle $x \neq 0$ ist $(F(0)x, x) > 0$, andererseits wird $F(0)$ durch nicht definite Operatoren approximiert, $0 \in \sigma(F(0))$. In allen drei Fällen folgt $0 \in \sigma(F)$.

II. Nun sei $m > 1$. Sei etwa $W_+ \neq \emptyset$ und $\lambda_0 \in \overline{W}_+ \cap \overline{W(F)}$. λ_0 ist Randpunkt der konvexen Menge W_+. Also gibt es eine Gerade g durch λ_0 derart, daß nicht beide Halbumgebungen von λ_0 bezüglich g zu W_+ gehören. Die Restriktion von F auf g liefert eine einparametrige Operatorfunktion $\tilde{F}(\alpha) = \alpha \tilde{B}_1 - \tilde{B}_0$ mit $\tilde{W}_+$, $\tilde{W}_-$, $W(\tilde{F})$. Der dem Punkte λ_0 entsprechende Parameterwert α_0 gehört zu $\overline{\tilde{W}}_+ \cap \overline{W(\tilde{F})}$, nach I ist $\alpha_0 \in \sigma(\tilde{F})$, also $\lambda_0 \in \sigma(F)$.

5. Klassifikation der Operatorfunktionen

Die durch (2.1) definierte Funktion F heiße nicht ausgeartet, wenn $W(F) \neq \emptyset$, $W(F) \neq R^m$ ist und die beiden Implikationen

$$W_+ \neq \emptyset \Rightarrow \mathring{W}_+ \neq \emptyset, \qquad W_- \neq \emptyset \Rightarrow \mathring{W}_- \neq \emptyset \tag{5.1}$$

gelten. ($\mathring{M}$ bezeichnet das Innere der Menge M.) Sei F nicht ausgeartet. Wir definieren: F heiße

hyperbolisch, falls $W_+ \neq \emptyset$, $W_- \neq \emptyset$,

parabolisch, falls eine der Mengen W_+, W_- leer und die andere unbeschränkt ist,

elliptisch, falls eine der Mengen W_+, W_- leer und die andere beschränkt ist.

F sei nicht ausgeartet. Sei $W_- \neq \emptyset$ (andernfalls betrachte $-F$). Sei $\lambda^- \in \mathring{W}_-$. Mit $\lambda = \mu + \lambda^-$ setzen wir

$$\tilde{F}(\mu) = \sum_{j=1}^{m} \mu_j B_j - \left(B_0 - \sum_{j=1}^{m} \lambda_j^- B_j\right) = \sum_{j=1}^{m} \mu_j \tilde{B}_j - \tilde{B}_0 .$$

Für $\tilde{F}$ ist jetzt $0 \in \mathring{\tilde{W}}_-$, also ist $\tilde{B}_0$ gleichmäßig positiv definit. Sei F hyperbolisch, dann auch $\tilde{F}$ und $\mathring{\tilde{W}}_+ \neq \emptyset$. Also kann man Punkte $\mu^{(k)}$ $(k = 1, \ldots, m)$, in $\mathring{\tilde{W}}_+$ finden, so daß die Vektoren $\overline{0\, \mu^{(k)}}$ linear unabhängig sind. Wir transformieren den R^m so, daß diese Vektoren Einheitsvektoren des Koordinatensystems werden. Dann erhalten wir eine neue hyperbolische Funktion $\hat{F}$,

$$\hat{F}(\omega) = \sum_{j=1}^{m} \omega_j \hat{B}_j - \hat{B}_0 ,$$

derart, daß $0 \in \mathring{\hat{W}}_-$ und

$$\omega^{(k)} = (0, \ldots, \underset{k}{1}, \ldots, 0) \in \mathring{\hat{W}}_+$$

sind. Sei D_k der Durchschnitt der k-ten Koordinatenachse mit $W(F)$. Eine Komponente von D_k liegt zwischen 0 und $\omega^{(k)}$. Weitere kann es nach Lemma 3.4 nicht geben. Also gilt

Lemma 5.1. *Wenn F nicht ausgeartet ist, gibt es eine Translation im Parameter-Raum, so daß in der neuen Funktion B_0 gleichmäßig positiv definit ist. Wenn F hyperbolisch ist, kann man durch eine affine Transformation erreichen, daß alle B_j ($j=0, \ldots, m$), gleichmäßig positiv definit werden. Insbesondere sind bei einer hyperbolischen Aufgabe alle Rayleigh-Hyperebenen eigentliche Hyperebenen.*

6. Erläuterung und Beispiel

Sei H der komplexe R^2, $m=2$, $B_0=E$. Die Funktion

$$F(\lambda) = \lambda_1 B_1 + \lambda_2 B_2 - E$$

erzeugt eine zweiparametrige Eigenwert-Aufgabe. Die Eigenwerte $\lambda = (\lambda_1, \lambda_2)$ sind die Nullstellen der Determinante

$$|\lambda_1 B_1 + \lambda_2 B_2 - E| = 0.$$

Die Lösungsgesamtheit ist leer oder bildet einen Kegelschnitt $\mathscr{K}$. Der Wertebereich ist

$$W(F) = \{\lambda\colon (B_1 x, x)\, \lambda_1 + (B_2 x, x)\, \lambda_2 = 1 \text{ für ein } x,\ \|x\| = 1\}.$$

Wenn $\mathscr{K}$ eine Hyperbel, Parabel bzw. Ellipse ist, so liegt der hyperbolische, parabolische oder elliptische Fall vor. Entsprechend lassen sich die ausgearteten Kegelschnitte einordnen.

7. Dualisierung und Rayleigh-Punkte

Wir betrachten wieder die Funktion (2.1) mit gleichmäßig definitem B_0 (vgl. (4.1)). Jedem Vektor $x \in H - \{0\}$ ordnen wir einen Punkt

$$\psi_x = \big((B_j x, x)/(B_0 x, x)\big) \in R^m \tag{7.1}$$

zu. ψ_x heiße Rayleigh-Punkt von x. Die Menge

$$W_d(F) = \{\psi\colon \psi = \psi_x \text{ für ein } x \in H - \{0\}\} \tag{7.2}$$

heiße dualer Wertebereich von F.

Sei $A \in \mathfrak{A}$. A besitzt die eindeutige Zerlegung $A = B_1 + i B_2$, mit $B_1, B_2 \in \mathfrak{H}$. Setzen wir

$$F(\lambda) = \lambda_1 B_1 + \lambda_2 B_2 - E, \tag{7.3}$$

so liefert die übliche Identifikation der komplexen Ebene mit dem reellen R^2

$$W(A) = W_d(F), \tag{7.4}$$

wobei $W(A)$ durch (1.1) definiert ist. Umgekehrt läßt sich jeder Funktion (7.3) ein Operator A so zuordnen, daß (7.4) gilt. Ein bekanntes Ergebnis von HAUSDORFF [6] besagt

Lemma 7.1. *Sei $m=2$ und B_0 gleichmäßig positiv definit. Dann ist $W_d(F)$ konvex.*

Für $m \geqq 3$ ist diese Aussage nicht allgemein richtig.

$\mathscr{R}$ und $\mathscr{R}_d$ seien zwei Exemplare des reellen R^m, die in folgender „Dualität" zueinander stehen (vgl. [3])

$$a=(a_j)\in\mathscr{R}_d,\quad a\neq 0 \Leftrightarrow \left\{\lambda\colon \sum_{j=1}^m a_j\lambda_j=1\right\}\subset\mathscr{R}.$$

Wir denken uns $W(F)$ in $\mathscr{R}$, $W_d(F)$ in $\mathscr{R}_d$ gelegen. Wir betrachten den (hyperbolischen) Fall, daß alle B_j $(j=0,\ldots,m)$, gleichmäßig positiv definit sind. Dann ist jede Rayleigh-Hyperebene eine echte Hyperebene. Dualisierung des Hausdorffschen Ergebnisses liefert

Lemma 7.2. *Sei $m=2$. Sind $\sum a_j\lambda_j=1$, $\sum b_j\lambda_j=1$ Rayleigh-Hyperebenen, so auch jede Konvex-Kombination $\sum[\alpha a_j+(1-\alpha)b_j]\lambda_j=1$, $0\leqq\alpha\leqq 1$.*

Für $m>2$ erhält man das schwächere

Lemma 7.3. *Sei $m\geqq 2$. Liegen zwei Hyperebenen ganz in $W(F)$, so auch jede Konvex-Kombination.*

Der Funktion F läßt sich als Verallgemeinerung des Rayleigh-Quotienten das homogene Funktional q auf $H-\{0\}$,

$$q(x)=\sum_{j=1}^m p_j(x)^2,\qquad p_j(x)=(B_jx,x)/(x,x)\qquad (j=1,\ldots,m),$$

zuordnen. $q(x)^{-1}$ ist das Quadrat des Abstands der Rayleigh-Hyperebene W_x vom Nullpunkt, andererseits ist $q(x)$ das Quadrat des Abstands des Punktes $\psi_x\in W_d(F)$ vom Nullpunkt. Für $m\geqq 3$ ist das Variationsproblem $q_0=\inf q(x)$ im allgemeinen keine konvexe Optimierungsaufgabe, wogegen

$$Q_0=\inf_{\lambda\in W_+} Q(\lambda),\qquad Q(\lambda)=\sum_{j=1}^m \lambda_j^2$$

für alle m eine konvexe Optimierungsaufgabe ist. Es gilt $q(x)^{-1}\leqq Q(\lambda)$ für alle $x\neq 0$, $\lambda\in W_+$. Durch zusätzliche Voraussetzungen (z.B. $\dim H<\infty$, alle Randpunkte von W_+ sind einfache Eigenwerte) kann man $q_0^{-1}=Q_0$ sichern.

8. Nichtlineare Eigenwertaufgaben

Im folgenden sei H der reelle R^n. $\mathfrak{H}$ sei der reelle lineare Raum der symmetrischen Matrizen der Ordnung n. (a,b) sei ein Intervall des R^1. T sei eine einmal stetig differenzierbare Abbildung von (a,b) in $\mathfrak{H}$,

$$T\colon\ \alpha\to T(\alpha)\in\mathfrak{H}\qquad \alpha\in(a,b). \tag{8.1}$$

Die Ableitung von T heiße T'.

p sei ein stetiges reellwertiges Funktional auf $H-\{0\}$ mit den folgenden Eigenschaften:

$$p:\ x \to p(x) \in (a,b), \tag{8.2}$$

$$p(c\,x)=p(x) \qquad \text{für } c \neq 0, \tag{8.3}$$

$$\big(T(p(x))\,x,x\big)=0, \tag{8.4}$$

$$\big(T'(p(x))\,x,x\big)>0. \tag{8.5}$$

Gelegentlich werden wir (8.5) durch die stärkere Forderung

$$\big(T'(p(x))\,y,y\big)>0 \qquad \text{für alle } x,y \in H-\{0\} \tag{8.6}$$

ersetzen.

Das Funktional p bezeichnen wir als Rayleigh-Funktional der Funktion T. Die Menge

$$W(T)=\{\alpha \in R^1 :\ \alpha=p(x) \text{ für ein } x \in H-\{0\}\}. \tag{8.7}$$

heiße Wertebereich von T. Das Spektrum $\sigma(T)$ von T definieren wir als

$$\sigma(T)=\{\alpha \in W(T):\ \exists\ x \in H-\{0\} \text{ mit } T(\alpha)\,x=0\}. \tag{8.8}$$

$\alpha \in \sigma(T)$ heißt Eigenwert, ein Vektor $x \neq 0$ mit $T(\alpha)x=0$ heißt Eigenvektor zu α.

9. Beispiele

1) Gewöhnliche Eigenwertaufgabe: Sei $A \in \mathfrak{H}$, $(a,b)=(-\infty,\infty)$,

$$T(\alpha)=\alpha E-A, \qquad p(x)=(A\,x,x)/(x,x). \tag{9.1}$$

p ist der gewöhnliche Rayleigh-Quotient (vgl. Nr. 1). Wegen $T'=E$ sind (8.5), (8.6) erfüllt. Das Funktional p ist durch (8.4) allein eindeutig bestimmt. Es ist

$$W(T)=W(A), \qquad \sigma(T)=\sigma(A). \tag{9.2}$$

2) Eine wichtige Klasse nichtlinearer Aufgaben ergibt sich durch Restriktion mehrparametriger Funktionen. Sei

$$F(\lambda)=\sum_{j=1}^{m} \lambda_j B_j - B_0 \tag{9.3}$$

eine hyperbolische Funktion. Sei $\lambda^- \in \mathring{W}_-$, $\lambda^+ \in \mathring{W}_+$, sei $\mathfrak{s}$ ein λ^- und λ^+ verbindender stetig differenzierbarer, doppelpunktfreier Kurvenbogen (vgl. Fig. 1),

$$\lambda=\lambda(\alpha), \quad \alpha \in [\alpha_-,\alpha_+], \quad \lambda(\alpha_-)=\lambda^-, \quad \lambda(\alpha_+)=\lambda^+, \quad \sum_{j=1}^{m}\left|\frac{d\lambda_j}{d\alpha}\right|>0. \tag{9.4}$$

Mit

$$T(\alpha)=F\big(\lambda(\alpha)\big) \tag{9.5}$$

erhalten wir eine nichtlineare Funktion vom Typ (8.1).

Die geometrische Veranschaulichung läßt vermuten: Im allgemeinen trifft $\mathfrak{s}$ gerade n Eigenwerthyperflächen, die zugehörigen Eigenvektoren bilden eine Basis von H. Anders gewendet: $\sigma(T)$ besteht bei geeigneter Interpretation der Vielfachheit aus n Eigenwerten, man kann in H eine Basis von Eigenvektoren zu T finden. Diese Vermutung läßt sich in dem Fall, daß $\mathfrak{s}$ eine Strecke ist, leicht bestätigen.

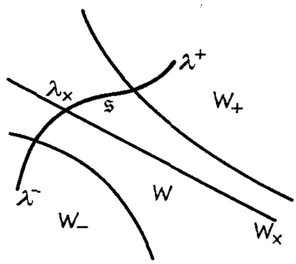

Fig. 1

Wir fordern

a) *$\mathfrak{s}$ trifft jede Rayleigh-Hyperebene von F genau einmal.* Dem Schnittpunkt λ_x der Rayleigh-Hyperebene des Vektors $x \neq 0$ mit $\mathfrak{s}$ entspricht nach (9.4) eindeutig ein Wert $\alpha = \alpha(\lambda_x) \in [\alpha_-, \alpha_+]$. Wir setzen $p(x) = \alpha(\lambda_x)$. Dann ist wegen

$$\sum_{j=1}^{m} \lambda_j(p(x))(B_j x, x) - (B_0 x, x) = 0 \tag{9.6}$$

die Voraussetzung (8.4) erfüllt.

b) *Der Tangentenvektor an die Kurve $\mathfrak{s}$ im Punkte λ_x sei nicht parallel zur Rayleigh-Hyperebene von x*, d. h.

$$\sum_{j=1}^{m} \lambda_j(p(x))(B_j x, x) > 0 \tag{9.7}$$

(wegen $\lambda(\alpha_-) = \lambda^- \in \mathring{W}_-$). Damit ist auch (8.5) erfüllt. Unter diesen Voraussetzungen läßt sich der vermutete Sachverhalt beweisen (vgl. Nr. 10).

Nach Nr. 5 können wir uns auf die Normalform der hyperbolischen Funktion mit positiv definiten B_j $(j=0, \ldots, m)$, beschränken. Beispiele sind etwa

$$T(\alpha) = \sum_{j=1}^{m} \alpha^j B_j - B_0, \qquad (a, b) = (0, \infty), \tag{9.8}$$

$$T(\alpha) = (e^\alpha - 1) B_1 + \alpha^2 B_2 - B_0, \qquad (a, b) = (0, \infty). \tag{9.9}$$

In beiden Fällen ist p durch die Angabe von (a, b) bestimmt. Im folgenden Beispiel sind die B_j nur semidefinit, trotzdem sind (8.4), (8.5) erfüllt. Sei

$$D = (d_i \delta_{ij}), \qquad d_i > 1 \qquad (i, j = 1, \ldots, n),$$

$$A \in \mathfrak{H}, \qquad (Ax, x) \geqq (x, x) \qquad \text{für alle } x,$$

$$T(\alpha) = D^\alpha - A, \qquad (a, b) = (0, \infty). \tag{9.10}$$

3) Seien $B_0, B_1, B_2 \in \mathfrak{H}$, B_0 und B_2 seien positiv definit. Sei $(a,b)=(-\infty,\infty)$ und

$$T(\alpha)=\alpha B_1+\alpha^2 B_2-B_0, \tag{9.11}$$

$$p(x)=\frac{1}{2(B_2 x,x)}\left[{}_+\sqrt{(B_1 x,x)^2+4(B_2 x,x)(B_0 x,x)}-(B_1 x,x)\right]. \tag{9.12}$$

Dann ist

$$T'(\alpha)=B_1+2\alpha B_2. \tag{9.13}$$

$$(T'(p(x))\,x,x)={}_+\sqrt{(B_1 x,x)^2+4(B_2 x,x)(B_0 x,x)}>0. \tag{9.14}$$

Also ist (8.5) erfüllt. Im allgemeinen gilt (8.6) nicht. Ist jedoch B_1 positiv definit, so ist $p(x)>0$ für alle $x \neq 0$, und aus (9.13) folgt (8.6).

T ist die Restriktion der zweiparametrigen Funktion (vgl. Fig. 2)

$$F(\lambda)=\lambda_1 B_1+\lambda_2 B_2-B_0 \tag{9.15}$$

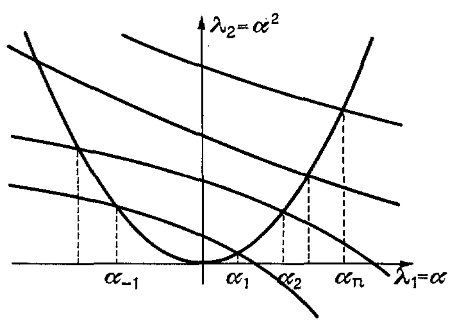

Fig. 2

auf den Kurvenbogen $\lambda_1=\alpha$, $\lambda_2=\alpha^2$, $0\leqq\alpha<\infty$. Durch die Wahl des positiven Vorzeichens in (9.12) werden n der zu erwartenden $2n$ Eigenwerte ausgezeichnet, und zwar die n positiven. Wählt man in (9.12) das andere Vorzeichen der Wurzel, so ist $(T(p(x))x,x)=-\sqrt{\cdot\,\bar{\,}\cdot}<0$. Dann erfüllt $\tilde{T}=-T$ wieder die Voraussetzung (8.5).

10. Die Minimax-Theorie von E. H. Rogers

Rogers [*10*] untersucht die Eigenwertaufgabe $T(\alpha)x=0$ für eine Funktion T, die die Voraussetzungen (8.1) bis (8.5) erfüllt. Grundlegend ist sein Ergebnis, daß die Eigenschaften (8.1) bis (8.5) das Funktional p im wesentlichen eindeutig bestimmen. Wie das Beispiel 3) in Nr. 9 zeigt, gibt es durchaus Funktionen T, die zwei oder mehr Rayleigh-Funktionale zulassen. Deren Wertebereiche sind dann notwendig verschieden.

Lemma 10.1 (Rogers). *Sei* $\alpha \in W(T)$ *und* $x \neq 0$. *Es gilt*

$$p(x)<\alpha \Leftrightarrow (T(\alpha)\,x,x)>0, \tag{10.1}$$

$$p(x)=\alpha \Leftrightarrow (T(\alpha)\,x,x)=0, \tag{10.2}$$

$$p(x)>\alpha \Leftrightarrow (T(\alpha)\,x,x)<0. \tag{10.3}$$

Die Voraussetzung $\alpha \in W(T)$ ist wesentlich.

Folgerung. $T(\alpha)x=0,\ x \neq 0,\ \alpha \in W(T) \Rightarrow \alpha = p(x).$ (10.4)

Lemma 10.2 (ROGERS). *$x_1, \ldots, x_r$ seien Eigenvektoren mit*

$$p(x_1) < p(x_2) < \cdots < p(x_r).$$

Dann sind sie linear unabhängig. Es gilt

$$p(x_1) < p(x_1 + x_2 + \cdots + x_r) < p(x_r). \tag{10.5}$$

Satz 10.1 (ROGERS). *Wenn man die Eigenwerte von T ihrer geometrischen Vielfachheit entsprechend zählt, so gibt es in $\sigma(T)$ genau n Eigenwerte $\alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_n$. Man kann in H eine Basis von Eigenvektoren x_j, $T(\alpha_j)x_j = 0$ $(j=1, \ldots, n)$, finden. Die Eigenwerte können durch das Minimum-Maximum-Prinzip*

$$\alpha_j = \min_{R^j} \max_{x \in R^j - \{0\}} p(x) \qquad (j=1, \ldots, n) \tag{10.6}$$

charakterisiert werden. (R^j bezeichnet einen beliebigen Teilraum der Dimension j von H.)

Folgerung. *$W(T)$ ist konvex.*

Die Beweise finden sich in [*10*]. Entsprechend lassen sich die Eigenwerte in absteigender Reihenfolge charakterisieren.

11. Orthogonalität der Eigenvektoren

Für $\alpha, \beta \in (a, b)$ definieren wir

$$\Delta(\alpha, \beta) = \begin{cases} \dfrac{1}{\alpha - \beta}[T(\alpha) - T(\beta)], & \text{falls } \alpha \neq \beta \\ T'(\alpha), & \text{falls } \alpha = \beta. \end{cases} \tag{11.1}$$

Durch

$$[x, y] = \begin{cases} (\Delta(p(x), p(y))\, x, y), & \text{falls } x, y \neq 0 \\ 0, & \text{falls } x = 0 \text{ oder } y = 0 \end{cases} \tag{11.2}$$

wird auf ganz H ein verallgemeinertes inneres Produkt eingeführt. Nach (11.2), (11.1), (8.5) gilt

$$[x, x] \geqq 0; \quad [x, x] = 0 \Leftrightarrow x = 0, \tag{11.3}$$

$$[x, y] = [y, x], \tag{11.4}$$

$$[c\,x, y] = c[x, y]. \tag{11.5}$$

Das Produkt ist also definit, symmetrisch und homogen, wenn auch im allgemeinen nicht bilinear.

Seien x_i, x_j Eigenvektoren zu Eigenwerten $\alpha_i \neq \alpha_j$. Aus $T(\alpha_i)x_i = 0$, $T(\alpha_j)x_j = 0$ folgt $[x_i, x_j] = 0$ wegen $T(\alpha) \in \mathfrak{H}$.

Sind andererseits $x_j, x_{j+1}, \ldots, x_{j+k-1}$ linear unabhängige Eigenvektoren zum Eigenwert α_j, so kann man in dem von ihnen aufgespannten k-dimensionalen Raum H_k eine bezüglich [,] orthonormierte Basis wählen, da ja für $x, y \in H_k$

das innere Produkt $[x,y]=(\Delta(\alpha_j,\alpha_j)x,y)=(T'(\alpha_j)x,y)$ nach (8.5) ein gewöhnliches definites Skalarprodukt ist. Also gilt

Lemma 11.1. *Man kann ein vollständiges System von Eigenvektoren x_j von T derart auswählen, daß*

$$[x_i, x_j]=\delta_{ij} \qquad (j=1,\dots,n) \tag{11.6}$$

gilt.

12. Orthogonalisierung

Sei $y_1, \dots, y_n$ eine Basis von H. Für

$$y=\sum_{j=1}^{n} c_j y_j$$

ist

$$(T(\alpha)\,y,y)=\sum_{i,j=1}^{n} c_i c_j (T(\alpha)\,y_i,y_j). \tag{12.1}$$

Sei $\alpha_0\in W(T)$ eine Nullstelle der stetig differenzierbaren Funktion

$$f(\alpha)=\det((T(\alpha)\,y_i,y_j)) \qquad (i,j=1,\dots,n). \tag{12.2}$$

Dann gibt es einen Vektor $(d_1,\dots,d_n)\neq 0$ mit

$$\sum_{j=1}^{n} (T(\alpha_0)\,y_i,y_j)\,d_j=0 \qquad (i=1,\dots,n), \tag{12.3}$$

also gelten für

$$y_0=\sum_{j=1}^{n} d_j y_j\neq 0$$

die Beziehungen

$$(T(\alpha_0)\,y_i,y_0)=0 \quad (i=1,\dots,n), \qquad T(\alpha_0)\,y_0=0, \tag{12.4}$$

nach (10.4) ist $\alpha_0=p(y_0)$. Umgekehrt ist auch jeder Eigenwert $\alpha\in\sigma(T)$ eine Nullstelle von f, also gilt

Lemma 12.1. $\alpha\in\sigma(T) \Leftrightarrow \alpha\in W(T),\ f(\alpha)=0.$

Im allgemeinen ist es nicht sinnvoll, von algebraischer Vielfachheit zu sprechen, die Vielfachheit eines Eigenwertes kann nur geometrisch, d.h. als Dimension des zugehörigen Eigenraumes definiert werden.

Sei $\tilde H\subset H$ ein Teilraum der Dimension $k<n$, sei $\tilde y_1,\dots,\tilde y_k$ eine Basis von $\tilde H$, sei P der orthogonale Projektor auf $\tilde H$. Die auf $\tilde H$ definierte Funktion

$$\tilde T(\alpha)=P\,T(\alpha)\,P \tag{12.5}$$

erfüllt (8.1) bis (8.5) mit dem Funktional $\tilde p$=Restriktion von p auf $\tilde H-\{0\}$. Die min-max-Werte $\tilde\alpha_j$ $(j=1,\dots,k)$, von $\tilde T$ sind die Eigenwerte von $\tilde T$ und diese wieder die Nullstellen von

$$\tilde f(\alpha)=\det((\tilde T(\alpha)\,\tilde y_i,\tilde y_j)), \quad (i,j=1,\dots,k), \quad \alpha\in W(\tilde T)\subset W(T). \tag{12.6}$$

Eine bezüglich des Produkts [,] orthonormierte Basis von $\tilde H$ finden wir, indem wir die Eigenvektoren von $\tilde T$ aufsuchen und gemäß Lemma 11.1 orthonormieren.

Sei speziell $\tilde{H}$ das Erzeugnis von $k-1$ Eigenvektoren

$$x_1, \dots, x_{k-1}, \quad [x_i, x_j] = \delta_{ij}, \quad T(\alpha_j) x_j = 0 \quad (i, j = 1, \dots, k-1) \tag{12.7}$$

von T und eines davon linear unabhängigen Vektors x. Wir setzen $\tilde{y}_j = x_j$ ($j=1, \dots, k-1$), $\tilde{y}_k = x$. In diesem Fall hat $\tilde{T}$ die Eigenwerte

$$\tilde{\alpha}_1 \leqq \tilde{\alpha}_2 \leqq \cdots \leqq \tilde{\alpha}_k, \quad \tilde{\alpha}_j = \alpha_j \quad (j = 1, \dots, k-1). \tag{12.8}$$

Zu $\tilde{\alpha}_k$ läßt sich nach Lemma 11.1 ein Eigenvektor $\tilde{x}_k \subset \tilde{H}$ von $\tilde{T}$ finden, so daß

$$[x_j, \tilde{x}_k] = 0, \quad (j = 1, \dots, k-1), \quad [\tilde{x}_k, \tilde{x}_k] = 1 \tag{12.9}$$

gilt. Man hat also in der üblichen Sprechweise den Vektor x „zu den x_j orthogonalisiert".

Praktisch hat man so vorzugehen, daß man den größten Eigenwert $\tilde{\alpha}_k$ von $\tilde{T}$, d.h. die größte Nullstelle von f bestimmt. Über $W(T)$ braucht nichts bekannt zu sein, je nach der geometrischen Vielfachheit ist $\tilde{\alpha}_k = \alpha_{k-1}$ oder $\tilde{\alpha}_k$ ist die nächstgrößere Nullstelle von f in (a, b). Im Falle $\tilde{\alpha}_k > \alpha_{k-1}$ ist der zugehörige Eigenvektor von T von selbst orthogonal zu $x_1, \dots, x_{k-1}$, andernfalls kann im Eigenraum von α_{k-1} ein bezüglich $(T'(\alpha_{k-1})\ ,\ )$ orthogonaler Vektor ausgewählt werden.

13. Das Rayleighsche Extremalprinzip

Wir wollen eine Verallgemeinerung des Rayleighschen Prinzips (1.8) herleiten. Sei $x \in H$ mit

$$[x_j, x] = 0, \quad (j = 1, \dots, k-1), \quad [x, x] = 1 \tag{13.1}$$

gegeben, wobei die x_j ein durch (12.7) orthonormiertes System von Eigenvektoren zu den ersten $k-1$ Eigenwerten sind. Für $0 \leqq l \leqq k-1$ sei V_l der von $x_1, \dots, x_l$ und W_l der von $x, x_1, \dots, x_l$ aufgespannte Raum. Dabei sei $V_0 = \{0\}$.

Behauptung. Für $l = 0, \dots, k-1$ gilt

W_l hat die Dimension $l+1$, es gilt $p(x) \geqq \alpha_{l+1}$.

Für $l=0$ ist die Behauptung richtig. Sei sie schon für ein festes $l = \rho - 1$, $0 \leqq \rho - 1 \leqq k-2$, gezeigt. Dann gibt es zwei Fälle.

Fall 1: $x \in V_\rho$. Mit Lemma 10.2 folgt $\alpha_1 \leqq p(x) \leqq \alpha_\rho$, also $p(x) = \alpha_\rho$. Wieder nach Lemma 10.2 folgt, daß x aus dem Durchschnitt U von V_ρ mit dem zu α_ρ gehörenden Eigenraum ist. Nach (13.1) ist x zu U bezüglich $(T'(\alpha_\rho)\ ,\ )$ orthogonal, also $x = 0$.

Fall 2. $x \notin V_\rho$. W_ρ hat die Dimension $\rho + 1$, es gilt

$$\max_{w \in W - \{0\}} p(w) \geqq \alpha_{\rho+1}.$$

Sei $y \in W_\rho - V_\rho$. Bis auf einen Faktor ist $y = x + v$, $v \in V_\rho$, also

$$\big(T(p(x))\, y, y\big) = \big(T(p(x))\, x, x\big) + 2\big(T(p(x))\, x, v\big) + \big(T(p(x))\, v, v\big).$$

Der erste Summand verschwindet nach (8.4), für den zweiten gilt

$$(T(p(x))x, v) = \left(T(p(x))x, \sum_{j=1}^{\rho} b_j x_j\right) = \sum_{j=1}^{\rho} b_j (T(p(x))x, x_j),$$

$$(T(p(x))x, x_j) = ([T(p(x)) - T(p(x_j))]x, x_j) = (p(x) - p(x_j))[x, x_j] = 0.$$

Daher folgt mit Lemma 10.2 und (10.3)

$$(T(p(x))y, y) = (T(p(x))v, v) \geqq 0 \Rightarrow p(x) \geqq \alpha_{\rho+1}.$$

Damit haben wir

Satz 13.1. *Seien die Eigenwerte von T durch $\alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_n$ geordnet. Sei $x_1, \ldots, x_n$ ein bezüglich [,] orthonormiertes System von Eigenvektoren, $T(\alpha_j)x_j = 0$. Dann gilt*

$$\alpha_j = \min_{\substack{[x_k, x] = 0,\ x \neq 0 \\ k = 1, \ldots, j-1}} p(x) \qquad (j = 1, \ldots, n).$$

14. Iterationsverfahren

Wir beginnen mit einem Hilfssatz.

Lemma 14.1. *Seien A, $B \in \mathfrak{H}$ und B positiv definit. Sei $z_0 \neq 0$ und*

$$A z_1 = B z_0. \tag{14.1}$$

a) *Sei v ein Eigenvektor zu einem von Null verschiedenen Eigenwert,*

$$A v = \lambda B v, \qquad \lambda \neq 0,\ v \neq 0. \tag{14.2}$$

Dann gilt die Implikation

$$(B z_0, v) = 0 \Rightarrow (B z_1, v) = 0. \tag{14.3}$$

b) *Sei $r(z) = (Az, z)/(z, z)$ der Rayleigh-Quotient des Vektors $z \neq 0$. Sei A positiv definit. Es gilt*

$$r(z_1) \leqq r(z_0) \tag{14.4}$$

und

$$r(z_1) = r(z_0) \Leftrightarrow z_0 \text{ ist Eigenvektor zum Eigenwert } r(z_0). \tag{14.5}$$

Beweis. a) $\lambda(B z_1, v) = (z_1, \lambda B v) = (z_1, A v) = (A z_1, v) = (B z_0, v) = 0$.

b) Wegen

$$B^{-\frac{1}{2}} A B^{-\frac{1}{2}} (B^{\frac{1}{2}} z_1) = (B^{\frac{1}{2}} z_0)$$

genügt es, den Fall $B = E$ zu betrachten. Dann sind (14.4), (14.5) wegen der Definitheit von A leicht einzusehen.

Wir setzen nur (8.5) (nicht (8.6)) voraus und formulieren ein erstes Iterationsverfahren zur Bestimmung eines Eigenwertes von T. Sei C eine stetige Abbildung von (a, b) in $\mathfrak{H}$,

$$C\colon\ \alpha \to C(\alpha) \in \mathfrak{H}, \tag{14.6}$$

derart, daß

$$C(\alpha) \quad \text{und} \quad C(\alpha) + T(\alpha) \tag{14.7}$$

für alle $\alpha \in W(T)$ positiv definit sind. Durch

$$\begin{gathered}\tilde{z}_0 \neq 0, \quad z_\nu = \tilde{z}_\nu / \|\tilde{z}_\nu\|, \quad \beta_\nu = p(z_\nu), \quad C_\nu = C(\beta_\nu), \\ C_\nu \tilde{z}_{\nu+1} = [C_\nu + T(\beta_\nu)]\, z_\nu \qquad (\nu = 0, 1, 2, \ldots)\end{gathered} \tag{14.8}$$

werde ein Iterationsverfahren definiert.

Satz 14.1. *Die Iteration* (14.8) *ist unbeschränkt ausführbar. Die Folge der* β_ν *konvergiert monoton von oben gegen einen Eigenwert* β. *Eine Teilfolge der* z_ν *konvergiert gegen einen zugehörigen Eigenvektor.*

Beweis. Die Operatorfunktion

$$S(\alpha) = C(\alpha)^{-1}[C(\alpha) + T(\alpha)]$$

ist im abgeschlossenen Intervall $W(T) \subset (a, b)$ stetig und gleichmäßig beschränkt,

$$\alpha \in W(T) \Rightarrow \|S(\alpha)\| \leqq K. \tag{14.9}$$

Die Iteration ist also unbeschränkt ausführbar. Den Rayleigh-Quotienten eines Vektors $z \neq 0$ bezüglich der definiten linearen Eigenwertaufgabe

$$C_\nu x = \lambda [C_\nu + T(\beta_\nu)]\, x \tag{14.10}$$

bezeichnen wir mit $r_\nu(z)$,

$$r_\nu(z) = (C_\nu z, z) / ([C_\nu + T(\beta_\nu)]\, z, z). \tag{14.11}$$

Da $\tilde{z}_{\nu+1}$ iterierter Vektor zu z_ν bei der Aufgabe (14.10) ist, folgt mit (14.4) und $(T(\beta_\nu)\, z_\nu, z_\nu) = 0$

$$1/[1 + (T(\beta_\nu)\, z_{\nu+1}, z_{\nu+1})/(C_\nu z_{\nu+1}, z_{\nu+1})] = r_\nu(z_{\nu+1}) = r_\nu(\tilde{z}_{\nu+1}) \leqq r_\nu(z_\nu) = 1,$$

also mit (10.1)

$$(T(\beta_\nu)\, z_{\nu+1}, z_{\nu+1}) \geqq 0 \Rightarrow \beta_\nu \geqq p(z_{\nu+1}) = \beta_{\nu+1}.$$

Da die Folge der β_ν nach unten durch α_1 beschränkt ist, existiert

$$\lim_{\nu \to \infty} \beta_\nu = \beta \geqq \alpha_1. \tag{14.12}$$

Die Folge der z_ν besitzt (wegen Kompaktheit der Einheitskugel) eine Teilfolge z_{ν_l}, die gegen ein y mit $\|y\| = 1$ konvergiert. Für $w = S(\beta)\, y$ und $\nu = 0, 1, 2 \ldots$ gilt

$$\|w - \tilde{z}_{\nu+1}\| \leqq \|S(\beta) - S(\beta_\nu)\| \cdot \|y\| + \|S(\beta_\nu)\| \cdot \|y - z_\nu\|.$$

Wegen (14.9), (14.12) und der Stetigkeit von S folgt

$$\tilde{z}_{\nu_l+1} \to w, \quad p(z_{\nu_l+1}) \to p(w), \quad \beta_{\nu_l+1} \to p(w), \quad \beta = p(w).$$

Nun ist w iterierter Vektor von y bezüglich der Aufgabe

$$C(\beta)\, x = \lambda [C(\beta) + T(\beta)]\, x. \tag{14.13}$$

Die Rayleigh-Quotienten von y, w bezüglich (14.13) sind beide gleich 1. Nach (14.5) ist y ein Eigenvektor von (14.13) zum Eigenwert 1,

$$C(\beta)\, y = 1 \cdot [C(\beta) + T(\beta)]\, y \Rightarrow T(\beta)\, y = 0.$$

Sei C eine stetige Abbildung von (a, b) in $\mathfrak{H}$, so daß $\hat{C}(\alpha)$ und $\hat{C}(\alpha)-T(\alpha)$ für alle $\alpha \in W(T)$ positiv definit sind. Definiert man ein Iterationsverfahren durch

$$\begin{aligned} &\tilde{z}_0 \neq 0, \quad z_\nu = \tilde{z}_\nu / \|\tilde{z}_\nu\|, \quad \beta_\nu = p(z_\nu), \quad \hat{C}_\nu = \hat{C}(\beta_\nu), \\ &[\hat{C}_\nu - T(\beta_\nu)]\, \tilde{z}_{\nu+1} = \hat{C}_\nu z_\nu \qquad (\nu = 0, 1, 2, \ldots), \end{aligned} \tag{14.14}$$

so gilt auch für dieses Verfahren die Behauptung des Satzes 14.1. Der Beweis ist dem vorigen sehr ähnlich. Im allgemeinen ist der Aufwand bei (14.14) größer als bei (14.8). In (14.8) kann man ja C_ν als Diagonalmatrizen oder einfacher

$$C(\alpha) = c(\alpha)\, E \quad \text{oder} \quad C(\alpha) \equiv c\, E \tag{14.15}$$

wählen.

Zur Erläuterung betrachten wir den linearen Fall $T(\alpha) = \alpha E - A$. Dann lautet die Vorschrift (14.8), durch (14.15) spezialisiert,

$$z_{\nu+1} = \rho_\nu \left[E - \frac{1}{c(\beta_\nu) + \beta_\nu} A \right] z_\nu, \qquad \beta_\nu = (A z_\nu, z_\nu)/(z_\nu, z_\nu)$$

mit gewissen Normierungsfaktoren ρ_ν und hinreichend großen $c(\beta_\nu)$. Im Spezialfall ergibt sich also ein v. Mises-Verfahren zur Berechnung des kleinsten Eigenwertes von A.

15. Zusammenhang mit dem Gradientenverfahren

Lemma 10.1 zeigt, daß das Funktional p durch die Gleichung (8.4) im wesentlichen eindeutig bestimmt ist. Nach dem Satz über implizite Funktionen und (8.5) ist p differenzierbar. Aus (8.4) erhalten wir

$$(T'(p(x))\, x, x) \operatorname{grad} p|_x + 2\, T(p(x))\, x = 0. \tag{15.1}$$

Ein Gradientenverfahren zur Bestimmung eines stationären Punktes des Funktionals p ist eine Vorschrift der Art

$$z_{\nu+1} = z_\nu - d_\nu \operatorname{grad} p|_{z_\nu}, \tag{15.2}$$

wobei die Konstanten d_ν geeignet zu wählen sind. Mit (15.1), (15.2) erhalten wir die Vorschrift

$$z_{\nu+1} = [E + \{2 d_\nu / (T'(p(z_\nu))\, z_\nu, z_\nu)\}\, T(p(z_\nu))]\, z_\nu, \tag{15.3}$$

die bei geeigneter Wahl der Konstanten mit (14.8), C_ν durch (14.15) spezialisiert, übereinstimmt. Man kann danach auch das gewöhnliche v. Mises-Verfahren als ein Gradientenverfahren deuten.

16. Linearisierung der Operatorfunktion

Sei $z_0 \neq 0$, $\beta_0 = p(z_0)$. Die Entwicklung

$$T(\alpha) = T(\beta_0) + (\alpha - \beta_0)\, T'(\beta_0) + \cdots \tag{16.1}$$

legt es nahe, die Funktion T durch

$$\hat{T}(\alpha) = \alpha\, T'(\beta_0) - [\beta_0\, T'(\beta_0) - T(\beta_0)] \tag{16.2}$$

zu ersetzen. Wir machen die Voraussetzung (8.6). T ist linear und erfüllt die Voraussetzungen (8.1) bis (8.6) mit dem gewöhlichen Rayleigh-Quotienten

$$\hat{p}(x) = ([\beta_0 T'(\beta_0) - T(\beta_0)]\, x, x)/(T'(\beta_0)\, x, x).$$

Ein v. Mises-Verfahren zur Bestimmung des kleinsten Eigenwertes von $\hat{T}$ wird durch

$$(c-\beta_0)\, T'(\beta_0)\, z_1 = \{c\, T'(\beta_0) - [\beta_0 T'(\beta_0) - T(\beta_0)]\}\, z_0 \tag{16.3}$$

gegeben. Wir definieren durch

$$\begin{gathered} \tilde{z}_0 \neq 0, \qquad z_\nu = \tilde{z}_\nu / \|\tilde{z}_\nu\|, \qquad \beta_\nu = p(z_\nu), \\ \tilde{z}_{\nu+1} = \left\{E + \frac{1}{c-\beta_\nu} [T'(\beta_\nu)]^{-1}\, T(\beta_\nu)\right\} z_\nu \qquad (\nu = 0, 1, 2, \ldots) \end{gathered} \tag{16.4}$$

ein Iterationsverfahren vom Typ (14.8). Nach Satz 14.1 konvergiert (16.4), falls nur $(c-\alpha)\, T'(\alpha)$ und $(c-\alpha)\, T'(\alpha) + T(\alpha)$ positiv definit für alle $\alpha \in W(T)$ sind.

Das Verfahren (16.4) ist offenbar ein Newton-Verfahren. Es wird nur die Ableitung nach dem Parameter α gebildet, nicht die Ableitung eines nichtlinearen Operators auf einem $(n+1)$-dimensionalen Raum (wie etwa beim Verfahren von UNGER [*15*]).

17. Bestimmung des zweiten Eigenwertes

Sei (8.6) erfüllt. Sei α_1 Eigenwert der Vielfachheit r von T mit dem Eigenraum V. Für jedes $x \neq 0$ gilt mit einem $\alpha = \alpha(x) \in [\alpha_1, \alpha_{r+1}]$

$$([T(\alpha_{r+1}) - T(\alpha_1)]\, x, x) = (\alpha_{r+1} - \alpha_1)(T'(\alpha)\, x, x) > 0, \tag{17.1}$$

also ist der Operator $T(\alpha_{r+1}) - T(\alpha_1)$ positiv definit. Sei $\vartheta > 0$ derart, daß

$$\vartheta [T(\alpha_{r+1}) - T(\alpha_1)] - T(\alpha_n) \tag{17.2}$$

positiv definit ist. Sei

$$c > \max[1, \vartheta]. \tag{17.3}$$

Sei $\tilde{z}_0 \notin V$. Zwei Folgen z_ν, $\tilde{z}_\nu$, $(\nu = 0, 1, 2, \ldots)$, werden rekursiv definiert: Sei W_ν der von V und $\tilde{z}_\nu$ aufgespannte Raum und

$$\beta_\nu = \max_{x \in W_\nu - \{0\}} p(x). \tag{17.4}$$

Bestimme z_ν und $\tilde{z}_{\nu+1}$ durch

$$z_\nu \in W_\nu, \qquad p(z_\nu) = \beta_\nu, \qquad \|z_\nu\| = 1, \tag{17.5}$$

$$[c(T(\beta_\nu) - T(\alpha_1)) - T(\beta_\nu)]\, \tilde{z}_{\nu+1} = [T(\beta_\nu) - T(\alpha_1)]\, z_\nu. \tag{17.6}$$

Satz 17.1. *Die durch* (17.4) *bis* (17.6) *definierte Iteration ist unbeschränkt ausführbar. Die Folge der β_ν konvergiert monoton von oben gegen einen Eigenwert $\beta > \alpha_1$. Eine Teilfolge der z_ν konvergiert gegen einen Eigenvektor zu β.*

Beweis. Es genügt, den Schritt von $\tilde{z}_0$ zu $\tilde{z}_1$ zu betrachten. Nach Nr. 12 folgt aus $\tilde{z}_0 \notin V$ und (17.5), daß z_0 bis auf einen Faktor (-1) eindeutig bestimmt ist und

$$\begin{gathered} p(z_0) \geqq \alpha_{r+1}, \quad (T(p(z_0))\, z_0, z_0) = 0, \quad \|z_0\| = 1, \\ [z_0, v] = 0 \qquad \text{für alle } v \in V \end{gathered} \tag{17.7}$$

erfüllt. Nach Konstruktion ist $c[T(\beta_0) - T(\alpha_1)] - T(\beta_1)$ positiv definit, also ist $\tilde{z}_1$ bis auf einen Faktor (-1) eindeutig bestimmt. Sei $v \in V$, $v \neq 0$. Wegen $T(\alpha_1)\, v = 0$ ist v ein Eigenvektor von

$$[c(T(\beta_0) - T(\alpha_1)) - T(\beta_0)]\, x = \lambda [T(\beta_0) - T(\alpha_1)]\, x \tag{17.8}$$

zum Eigenwert $\lambda = c - 1 > 0$. Nach (14.3) folgt

$$(T(\beta_0)\, \tilde{z}_1, v) = ([T(\beta_0) - T(\alpha_1)]\, \tilde{z}_1, v) = 0 \qquad \text{für alle } v \in V. \tag{17.9}$$

Wäre $\tilde{z}_1 \in V$, so speziell $(T(\beta_0)\, \tilde{z}_1, \tilde{z}_1) = 0 \Rightarrow \beta_0 = p(\tilde{z}_1)$ und $p(\tilde{z}_1) = \alpha_1 \Rightarrow \beta_0 = \alpha_1$ im Widerspruch zur Voraussetzung. Also hat der Raum W_1 die Dimension $r+1$. Die Iteration ist unbeschränkt ausführbar.

Sei $w \in W_1 - V$. Bis auf einen zu vernachlässigenden Faktor ist $w = \tilde{z}_1 + v$, $v \in V$, also mit (17.9), (14.4), (10.1)

$$\begin{aligned} (T(\beta_0)\, w, w) &= (T(\beta_0)\, v, v) + 2(T(\beta_0)\, v, \tilde{z}_1) + (T(\beta_0)\, \tilde{z}_1, \tilde{z}_1) \\ &= (T(\beta_0)\, v, v) + (T(\beta_0)\, \tilde{z}_1, \tilde{z}_1) \geqq 0 \;\Rightarrow\; p(w) \leqq \beta_0 . \end{aligned} \tag{17.10}$$

Durch Orthogonalisierung von $\tilde{z}_1$ zu V erhalten wir einen Vektor $z_1 \in W_1$, der alle Eigenschaften (17.7) von z_0 besitzt und $p(w) \leqq \beta_1 \leqq \beta_0$ für alle $w \in W_1$, $\alpha_{r+1} \leqq \beta_1$ erfüllt. Die Iteration liefert eine Folge $\beta_0 \geqq \beta_1 \geqq \cdots \geqq \alpha_{r+1}$, die gegen ein $\beta \geqq \alpha_{r+1}$ konvergiert. Die Folge der z_ν besitzt eine gegen ein z, $\|z\| = 1$, konvergente Teilfolge. z erfüllt $\beta = p(z)$. Mit den beim Beweise von Satz 14.1 benutzten Schlüssen zeigt man, daß z Eigenvektor der linearen Aufgabe

$$[c(T(\beta) - T(\alpha_1)) - T(\beta)]\, z = \lambda [T(\beta) - T(\alpha_1)]\, z \tag{17.11}$$

ist. Bildet man auf beiden Seiten das innere Produkt mit z, so folgt wegen $\beta = p(z)$ und $(T(\alpha_1)\, z, z) < 0$ die Gleichung $c = \lambda$ und dann $T(\beta)\, z = 0$ aus (17.11).

Bemerkung. Offenbar kann man (17.2) durch die schwächere Forderung ersetzen, daß $\vartheta [T(\alpha) - T(\alpha_1)] - T(\alpha)$ für alle $\alpha \in [\alpha_{r+1}, \beta_0]$ positiv definit sei.

18. Polynomiale Büschel

Unter einem polynomialen Büschel verstehen wir eine in der komplexen Ebene $\mathbb{C}$ definierte Funktion

$$P(\lambda) = \sum_{j=1}^{m} \lambda^j A_j, \tag{18.1}$$

deren Koeffizienten A_j komplexe Matrizen der Ordnung n sind. Das Spektrum $\sigma(P)$ wird durch

$$\sigma(P) = \{\lambda \in \mathbb{C} \colon 0 \in \sigma(P(\lambda))\} \tag{18.2}$$

erklärt. Durch Übergang zur Frobenius-Matrix erhält man das lineare Büschel

$$P_1(\lambda)=\lambda\begin{pmatrix} E & & & \\ & E & 0 & \\ & & \ddots & \\ & 0 & E & \\ & & & -A_m\end{pmatrix}-\begin{pmatrix} 0 & E & & 0 \\ & & E & \\ & & & \ddots \\ 0 & & & E \\ A_0 & A_1 & \dots & A_{m-1}\end{pmatrix}, \tag{18.3}$$

dessen Werte Matrizen der Ordnung $m \cdot n$ sind. Es gilt

$$\sigma(P_1)=\sigma(P). \tag{18.4}$$

Im Falle dim $H=n=1$ ist P_1 einfach, d.h. die Anzahl linear unabhängiger Eigenvektoren ist gleich der Anzahl verschiedener Eigenwerte. Daher ist eine Zurückführung auf symmetrische oder normale Operatoren selbst im Falle symmetrischer A_j schwierig. Sie gelingt jedoch im Falle $m=2$ unter speziellen Voraussetzungen. Wir benutzen die Bezeichnungen von Nr. 9, 3). Sei H wieder der reelle R^n. Die auf $\mathbb{C}$ definierte Funktion P,

$$P(\lambda)=\lambda^2 B_2+\lambda B_1-B_0 \tag{18.5}$$

$B_0, B_1, B_2 \in \mathfrak{H}$, B_0, B_2 positiv definit, wird beim Übergang (18.1), (18.3) zu

$$P_1(\lambda)=\lambda\begin{pmatrix} E & 0 \\ 0 & B_2\end{pmatrix}-\begin{pmatrix} 0 & E \\ B_0 & -B_1\end{pmatrix}. \tag{18.6}$$

P_1 wird durch C symmetrisiert (WIELANDT, vgl. [9]):

$$C=\begin{pmatrix} B_0 & 0 \\ 0 & E\end{pmatrix}, \quad P_2(\lambda)=C\,P_1(\lambda)=\lambda\begin{pmatrix} B_0 & 0 \\ 0 & B_2\end{pmatrix}-\begin{pmatrix} 0 & B_0 \\ B_0 & -B_1\end{pmatrix}. \tag{18.7}$$

Da der Koeffizient von λ definit ist, sind alle Eigenwerte von P_2 und damit auch von P reell. Der Rayleigh-Quotient von P_2 ist

$$\tilde{R}(u,v)=\frac{2(B_0 u, v)-(B_1 v, v)}{(B_0 u, u)+(B_2 v, v)}, \qquad 0 \neq \{u, v\} \in H^2. \tag{18.8}$$

Da die Eigenvektoren von P_2 die Gestalt $\{u, \lambda u\}$ haben, kann man sich auf

$$R(u,\lambda)=\frac{2\lambda(B_0 u, u)-\lambda^2(B_1 u, u)}{(B_0 u, u)+\lambda^2(B_2 u, u)}, \qquad 0 \neq u \in H,\ \lambda \text{ reell}, \tag{18.9}$$

beschränken. Sei $g(\lambda)=R(u,\lambda)$ bei festem $u \neq 0$. Die Extremwerte von g sind die Lösungen der Gleichung

$$(B_2 u, u)\,g^2+(B_1 u, u)\,g-(B_0 u, u)=0. \tag{18.10}$$

Die größere Lösung ist das nach (9.12) bestimmte $p(u)$. Wir numerieren die Eigenwerte von P_2 durch

$$\lambda_{-n} \leqq \lambda_{-n+1} \leqq \cdots \leqq \lambda_{-1} < 0 < \lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_n. \tag{18.11}$$

Die positiven Eigenwerte λ_j sind gerade die Eigenwerte von Nr. 9, 3). Mit dem üblichen Variationsprinzip für P_2 und (18.8) bis (18.10) erkennt man

$$\max_{x \neq 0} p(x) = \lambda_n; \qquad \min_{x \neq 0} p(x) = \lambda_1 \tag{18.12}$$

sieht man leicht, wenn man in (18.5) λ durch λ^{-1} ersetzt. Für den ersten Eigenwert $\lambda_1 = \alpha_1$ ergibt sich also dasselbe Variationsprinzip wie in Nr. 9, Nr. 10.

Seien x, y Eigenvektoren zu (nicht notwendig positiven) Eigenwerten λ, μ von $P, \lambda \neq \mu$,

$$\lambda^2 B_2 x + \lambda B_1 x - B_0 x = 0, \qquad \mu^2 B_2 y + \mu B_1 y - B_0 y = 0. \tag{18.13}$$

Multipliziert man die erste Gleichung skalar mit y, die zweite mit x, subtrahiert die Gleichungen und dividiert durch $\lambda - \mu$, so folgt

$$0 = (\lambda + \mu)(B_2 x, y) + (B_1 x, y) = [x, y]. \tag{18.14}$$

Andererseits ergibt Multiplikation der ersten Gleichung mit $\mu\, y$, der zweiten mit λx

$$0 = \lambda \mu (B_2 x, y) + (B_0 x, y). \tag{18.15}$$

Das Orthogonalitätsprinzip (18.14) entspricht (11.2), das Prinzip (18.15) ist das gewöhnliche Orthogonalitätsprinzip von P_2 mit Berücksichtigung der speziellen Gestalt der Eigenvektoren in H^2. Man kann also bei einer nichtlinearen Eigenwertaufgabe im allgemeinen mehrere wesentlich verschiedene (sinnvolle) innere Produkte einführen, bezüglich denen die Eigenvektoren zu verschiedenen Eigenwerten orthogonal sind.

19. Numerisches Beispiel

Wir betrachten die Funktion

$$T(\alpha) = (e^{\alpha} - 1) B_1 + \alpha^2 B_2 - B_0 \tag{9.9}$$

mit den positiv definiten Matrizen

$$\begin{aligned} B_0 &= b_0 E, & b_0 &> 0, \\ B_1 &= (b_{jk}^{(1)}), & b_{jk}^{(1)} &= [n + 1 - \max(j, k)] \cdot j \cdot k, \\ B_2 &= (b_{jk}^{(2)}), & b_{jk}^{(2)} &= n\, \delta_{jk} + 1/(j + k). \end{aligned}$$

Wir setzen etwa $n = 8$ und $b_0 = 100$. Sei $\tilde{z}_0 = (1, \ldots, 1)$. Dann ist $p(z_0) = 0.24306$. Wir iterieren nach (14.8) mit $C(\alpha) = c(\alpha)\, E$, wobei $c(\alpha)$ die Zeilensummen-Norm von $T(\alpha)$ ist. Es ist

$$\beta_1 = 0.22262, \qquad \beta_2 = 0.21849, \ldots, \qquad \beta_{12} = 0.217\,461\,138.$$

Als Näherung für den Eigenvektor erhalten wir

$$z_{12} = (0.2209, 0.4402, 0.6519, 0.8393, 0.9708, 1.000, 0.8719, 0.5405).$$

Für den zweiten Eigenwert wird nach (17.5), (17.6) iteriert. Als Ausgangsvektor wählen wir $\tilde{z}_0 = (0.2209, -0.4402, 0.6519, - \cdots)$. Dann ist $\beta_0 = 2.1729$. Als Näherung für den zweiten Eigenwert ergibt sich $\beta_{31} = 0.8850$ mit dem Eigenvektor

$$z_{31} = (0.226, 0.441, 0.610, 0.635, 0.377, -0.218, -0.899, -1.000).$$

In dieser Arbeit sind die Ergebnisse des zweiten Teils der Habilitationsschrift des Autors zusammengefaßt. Der erste Teil, der inverse Eigenwertprobleme behandelt, erscheint demnächst [16]. Herrn Professor Dr. Dr. h.c. L. COLLATZ möchte ich sehr herzlich für sein Interesse an diesen Arbeiten und für seine Unterstützung danken. Herrn Professor Dr. E. SPERNER in Hamburg und Herrn Professor Dr. H. SCHAEFER in Tübingen bin ich ebenfalls zu besonderem Dank verpflichtet.

Literatur

[1] COLLATZ, L., Eigenwertaufgaben mit technischen Anwendungen. Leipzig 1963.
[2] DUFFIN, R. J., A minimax theory for overdamped networks. J. Rational Mech. Anal. **4**, 221–223 (1955).
[3] EGGLESTON, H. G., Convexity. Cambridge 1958.
[4] GUT, J., Kipp-Probleme als zusammengesetzte Stabilitätsaufgaben. Schweizerische Bauzeitung **84**, 38–41 (1966).
[5] GOCHBERG, I. Z., & M. G. KREIN, Einführung in die Theorie der linearen nichtselbstadjungierten Operatoren. Moskau 1965 (russ.).
[6] HAUSDORFF, F., Der Wertevorrat einer Bilinearform. Math. Zeitschr. **3**, 314–316 (1919).
[7] LANCASTER, P., Lambda-matrices and vibrating systems. Pergamon Press 1966.
[8] MEIXNER, J., & F. W. SCHÄFKE, Mathieusche Funktionen und Sphäroidfunktionen. Berlin-Göttingen-Heidelberg: Springer 1960.
[9] MÜLLER, P. H., Eine neue Methode zur Behandlung nichtlinearer Eigenwertaufgaben. Math. Zeitschr. **70**, 381–406 (1959).
[10] ROGERS, E. H., A minimax theory for overdamped systems. Arch. Rational Mech. Anal. **16**, 89–96 (1964).
[11] SCHAEFER, H., Beitrag zur Berechnung des kleinsten Eigenwertes eindimensionaler Eigenwertprobleme. Diss. Hannover 1934, erschienen 1937.
[12] STRIGL, G., Der Stahlbau **24**, 33 (1955).
[13] TAYLOR, A. E., Introduction to Functional Analysis. New York 1958.
[14] TURNER, R. E. L., Some variational principles for a nonlinear eigenvalue problem. J. Math. Anal. and Appl. **17**, 151–160 (1967).
[15] UNGER, H., Nichtlineare Behandlung von Eigenwertaufgaben. ZAMM **30**, 281–282 (1950).
[16] HADELER, K. P., Ein inverses Eigenwertproblem. Linear Algebra and Applications **1**, (1967).

Institut für Angewandte Mathematik
der Universität Hamburg
2 Hamburg 13, Rothenbaumchaussee 67/69

(Eingegangen am 31. Juli 1967)