

## Orthogonalreihenentwicklung nach linearen Spline-Funktionen

J. NITSCHKE

*Institut für Angewandte Mathematik der Universität Freiburg, Freiburg, Germany*

### 1. EINFÜHRUNG

Unter einem linearen Approximationsschema (vgl. z.B. Browder [9]) zur angenäherten Lösung einer Gleichung

$$Ax = f \tag{1.1}$$

in einem Banach-Raum  $X$  werde eine Folge  $\{X_n\}$  von endlichdimensionalen Teilräumen und eine Folge  $\{P_n\}$  von linearen Projektionsoperatoren  $P_n: D(A) \rightarrow X_n$  verstanden. Fragt man nach der Approximierbarkeit für alle  $x$  einer Menge  $\mathfrak{X} \subset D(A)$  von Lösungen von (1.1), so stellt

$$E(\mathfrak{X}, X_n) = \sup_{x \in \mathfrak{X}} \inf_{y \in X_n} |x - y| \tag{1.2}$$

ein unteres Fehlermass dar. Gilt mit einer von  $n$  unabhängigen Konstanten  $K$

$$\sup_{x \in \mathfrak{X}} |x - P_n x| \leq KE(\mathfrak{X}, X_n), \tag{1.3}$$

so wird das Verfahren, charakterisiert durch  $\{P_n\}$ , quasioptimal genannt (vgl. z.B. Lorentz [12], S. 120). Besteht (1.3) elementweis, d.h. ist

$$|x - P_n x| \leq KE(x, X_n), \tag{1.4}$$

so werde  $\{P_n\}$  nach Alexits [4] fast-best-approximierend genannt, im folgenden mit f.b.a. abgekürzt. Offenbar impliziert (1.4) stets (1.3), während die Umkehrung nicht richtig zu sein braucht.

In vielen Fällen wird die Folge der Projektionsoperatoren aus Minimalprinzipien in einem Hilbert-Raum festgelegt. Sei etwa  $H$  ein solcher Raum mit Skalarprodukt  $(\cdot, \cdot)$  und Norm  $\|\cdot\|$ , der zumindest  $x$  und  $X_n$  enthalte. Dann ist durch die Orthogonalprojektion auf  $X_n$  in  $H$ :

$$\|x - P_n x\| = \inf_{y \in X_n} \|x - y\| \tag{1.5}$$

ein linearer Operator  $P_n: X \cap H \rightarrow X_n$  definiert. Werden die  $X_n$  speziell von den  $n$  ersten Elementen einer Folge  $\{\varphi_i | \varphi_i \in X \cap H\}$  aufgespannt, die o.B.d.A.

als ortho-normiert angenommen werden kann, so ist  $P_n x$  gerade die  $n$ -te Partialsumme der Fourier-Reihe

$$\sum_1^{\infty} (\varphi_i, x) \varphi_i.$$

Der Fall  $X = C(0, 1)$ ,  $H = L_2(0, 1)$  ist eingehend, insbesondere für die trigonometrischen Funktionen, untersucht (vgl. z.B. Alexits [3]).

Zur angenäherten Berechnung einer Lösung von (1.1) ist die Orthogonalreihen-Entwicklung (1.5) nicht brauchbar, da ja dann die Lösung nicht explizit bekannt ist. In dem Fall ist das Ritzsche Verfahren als Orthogonalprojektion bezüglich der weiteren Hilbert-Norm ( $A$  positiv-definit und selbstadjungiert in  $H$ )

$$\|x\|_A = (x, Ax)^{1/2} \quad (1.6)$$

wohl die einfachste Orthogonalreihen-Entwicklung, zu deren Berechnung nur die gegebenen Daten  $A$  und  $f$  benötigt werden.

Die Folge  $\{P_n\}$  ist genau dann in ganz  $X$  f.b.a., wenn die Normen  $\|P_n\|$  gleichmässig beschränkt sind (Alexits [4]). Es ist bekannt, dass bei  $X = C$  für die Teilräume der trigonometrischen Polynome  $n$ -ten Grades als  $X_n$  keine solche Folge existiert (vgl. z.B. Lorentz [12], S. 96). Insofern kann die Wahl von  $X_n$  wesentlich sein. Von Haar [11] wurde erstmals eine in ganz  $C$  f.b.a. Orthogonalreihen-Entwicklung angegeben.<sup>1</sup> Die Teilräume  $X_n$  werden dabei von den stückweise konstanten Funktionen mit vorgegebenen Sprungstellen erzeugt. Durch ein- oder mehrmaliges Integrieren ergeben sich Funktionensysteme, die mit dem Namen Polynom-Spline-Funktionen bezeichnet werden. Derartige Funktionen sind in den letzten 20 Jahren vielfach untersucht worden,<sup>2</sup> allerdings wohl unabhängig von Haar im Zusammenhang mit Interpolationsproblemen. Nach dem Ergebnis von Ahlberg-Nilson-Walsh [1] minimiert die durch Interpolation gewonnene Spline-Funktion  $Jx$  gerade die Abweichung  $x - Jx$  in der Pseudo-Hilbert-Norm

$$N_{2,m}(x) = \left\{ \int |x^{(m)}|^2 ds \right\}^{1/2}, \quad (1.7)$$

wenn mit Spline-Funktionen  $(2m-1)$ -ten Grades ( $(2m-1)$ -malige Integration der Haarschen Funktionen) gearbeitet wird. Insofern kann die Spline-Interpolation als Orthogonalprojektion bezüglich  $N_{2,m}$  aufgefasst werden. Von Birkhoff und de Boor ([7] und [8]) wurde gezeigt, dass  $J$  im kubischen Fall ( $m = 2$ ) in  $C$  stetig ist und somit auch f.b.a. Weitere Untersuchungen, d.h. für andere  $X_n$ , andere Wahl von  $X$  und insbesondere andere Normen als (1.7), stehen aus.

<sup>1</sup> Haar beweist die gleichmässige Konvergenz  $P_n x \rightarrow x$  für jedes  $x \in C$ , woraus nach dem Uniform-Boundedness-Principle die Aussage folgt.

<sup>2</sup> Hinsichtlich Literatur sei auf das Buch von Ahlberg-Nilson-Walsh [2] wie auch die Zusammenstellung in der Arbeit von Schultz-Varga [15] verwiesen.

In der vorliegenden Arbeit wird der einfachste Fall  $m = 1$ , also lineare Spline-Funktionen bzw. Polygonzüge, behandelt. Neben der Wahl  $H = L_2$  werden wir auch den Raum  $H_A$  mit der Norm

$$\|x\|_A = \left\{ \int_0^1 (px'^2 + qx^2) ds \right\}^{1/2} \quad (1.8)$$

heranziehen, was (1.6) für das Randwertproblem

$$\begin{aligned} Ax &= -(px')' + qx = f(s) \quad s \in (0, 1) \\ x(0) &= x(1) = 0 \\ (p \geq p > 0, q \geq 0) \end{aligned} \quad (1.9)$$

entspricht, sofern die zugelassenen Funktionen gemäss den Randbedingungen (1.9<sub>2</sub>) eingeschränkt sind. Die lineare Spline-Interpolation ist darin bei  $p = 1$  und  $q = 0$  enthalten. Wie sich zeigt, sind diese Projektionen unter geringen Voraussetzungen an die Koeffizienten  $p, q$  f. b. a., und zwar in  $C$  als auch in  $C_1$ , dem Raum der Funktionen mit beschränkter messbarer 1. Ableitung. Daraus ergeben sich Fehlerabschätzungen der Gestalt ( $R$  gibt den Projektionsoperator bei der Norm (1.8) an)

$$|x - Rx| \leq k_1 |f| n^{-2}, \quad (1.10)$$

$$|x - Rx|_1 \leq k_2 |f| n^{-1} \quad (1.11)$$

allein unter der Voraussetzung  $p', q, f$  beschränkt und messbar.

Wie auch schon von Birkhoff und de Boor [8] bemerkt, führt das Ritz-Verfahren bei Verwenden linearer Spline-Funktionen auf Differenzengleichungen für die Funktionswerte der Näherung in den Stützstellen. Fehlerabschätzungen der obigen Art waren bislang nur unter der Voraussetzung beschränkter 4. Ableitungen anstelle der zweiten hier bekannt (vgl. z.B. Aziz und Hubbard [5]; Babuska, Prager und Vitasek [6], S. 154; Collatz [10], S. 178; Zlamal [16]). Die Untersuchung von Schultz und Varga [15] liefert bei den Voraussetzungen von oben an  $p, q$  und bei quadratisch integrierbarem  $f$  den Konvergenzfaktor  $n^{-1}$  in (1.10), in [13] wurde der Exponent auf  $3/2$  verbessert.

## 2. BEZEICHNUNGEN

Im folgenden sei  $C = C_0$  der Raum der in  $[0, 1]$  stetigen und am Rande verschwindenden Funktionen und  $C_1$  der Teilraum der totalstetigen Funktionen, versehen mit den Normen

$$|x| = |x|_0 = \text{Max}_{0 \leq s \leq 1} |x(s)|, \quad |x|_1 = \sup_{0 < s < 1} |x'(s)|. \quad (2.1)$$

Daneben führen wir Hilbert-Räume,  $H, H_1$  mit den Skalarprodukten

$$(x, y) = \int_0^1 x(s)y(s)ds, \quad (x, y)_1 = \int_0^1 x'(s)y'(s)ds \quad (2.2)$$

ein; es gilt dann  $C_1 \subset H_1 \subset C \subset H = L_2(0, 1)$ .

Mit  $\pi$  wollen wir eine Unterteilung

$$\pi = \{s_i | 0 = s_0 < s_1 < \dots < s_n < s_{n+1} = 1\}$$

bezeichnen und mit  $\bar{\pi}$  bzw.  $\underline{\pi}$  das Maximum bzw. Minimum der Gitterbreiten  $\Delta_i = s_{i+1} - s_i$ . Eine Menge von Unterteilungen heie  $c$ -regulr, wenn  $\bar{\pi} \leq c\underline{\pi}$  fr alle  $\pi$  der Menge gilt.

Die Elemente von  $C$ , die in den Teilintervallen  $(s_i, s_{i+1})$  stckweise linear sind, bilden den  $n$ -dimensionalen Raum<sup>3</sup>  $S(\pi)$  der linearen Spline-Funktionen zur Unterteilung  $\pi$ , wobei  $S(\pi) \subset C_1$  gilt. Die Funktionen

$$\varphi_i = \begin{cases} \Delta_{i-1}^{-1}(s - s_{i-1}) & s \in [s_{i-1}, s_i] \\ \Delta_i^{-1}(s_{i+1} - s) & s \in [s_i, s_{i+1}] \\ 0 & \text{sonst} \end{cases} \quad (2.3)$$

bilden eine Basis in  $S(\pi)$ . Jedem  $x \in S(\pi)$  lsst sich ber die Darstellung<sup>4</sup>  $x = \sum \xi_i \varphi_i$  umkehrbar eindeutig ein Vektor  $\bar{x} = \{\xi_i\}$  zuordnen. Dabei gilt

$$|\sum \xi_i \varphi_i| = |\bar{x}|, \quad (2.4)$$

wobei wir mit  $|\bar{x}|$  oder spter auch  $|\mathfrak{X}|$  die Tschebyscheff-Normen fr Vektoren und Matrizen

$$|\bar{x}| = \text{Max } |\xi_i|, \quad |\mathfrak{X}| = \text{Max}_i \sum_k |a_{ik}|$$

bentzen.

Indizierte Normen wie  $|T|_{k, l}$  geben die Operatornorm  $T: C_l \rightarrow C_k$  an:

$$|T|_{k, l} = \sup_{|x|_l \leq 1} |Tx|_k. \quad (2.5)$$

Ist  $P_\pi: C_k \rightarrow S(\pi)$  ( $k = 0, 1$ ) fr jedes  $\pi$  ein linearer Projektionsoperator, so ist nach Alexits [4] wie erwhnt die Bedingung

$$|P_\pi|_{k, k} \leq K_{k, k} \quad (2.6)$$

notwendig und hinreichend dafr, dass  $P_\pi$  in  $C_k$  f.b.a. ist, d.h. dass fr alle  $x \in C_k$

$$|x - P_\pi x|_k \leq K_k \inf_{y \in S(\pi)} |x - y|_k \quad (2.7)$$

<sup>3</sup> Ohne die Randbedingungen wre die Dimension  $n + 2$ .

<sup>4</sup> Soweit bei  $\sum$  oder auch bei Max die Grenzen 1 und  $n$  sind, wird deren Angabe unterdrckt.

gilt. Speziell kann dann

$$K_k = 1 + K_{k,k} \quad (2.8)$$

gewählt werden.

### 3. ORTHOGONALPROJEKTION IN DER METRIK $H$

In diesem Paragraphen wollen wir die Approximationseigenschaften des durch

$$\|x - Px\| = \inf_{y \in S(\pi)} \|x - y\| \quad (3.1)$$

bzw.

$$(x - Px, y) = 0 \quad \text{für } y \in S(\pi)$$

definierten Operators  $P = P_\pi$  untersuchen. Die Koeffizienten  $\xi_i$  bei  $Px = \sum \xi_i \varphi_i$  berechnen sich aus dem linearen Gleichungssystem

$$\sum_k a_{ik} \xi_k = \eta_i = (\varphi_i, x) \quad (3.2)$$

mit

$$a_{ik} = \frac{1}{6} \begin{cases} \Delta_{i-1} & k = i-1 \\ 2(\Delta_{i-1} + \Delta_i) & k = i \\ \Delta_i & k = i+1 \\ 0 & \text{sonst.} \end{cases} \quad (3.3)$$

Nach (2.4) sind zum Berechnen von  $|Px|$  die  $\xi_i$  abzuschätzen. Da die Matrix  $(a_{ik})$  eine überwiegende Hauptdiagonale besitzt, es ist

$$\nu_i = |a_{ii}| - \sum_{k \neq i} |a_{ik}| = \frac{1}{6}(\Delta_{i-1} + \Delta_i) > 0,$$

gilt

$$\text{Max } |\xi_i| \leq \text{Max } (\nu_i^{-1} |\eta_i|). \quad (3.4)$$

Weiter ist

$$|\eta_i| = |(\varphi_i, x)| \leq |x| \int_0^1 \varphi_i ds = \frac{1}{2}(\Delta_{i-1} + \Delta_i) |x|.$$

Daher erhalten wir

$$|Px| \leq 3|x|. \quad (3.5)$$

Hinsichtlich der 1. Ableitung ist<sup>5</sup>

$$|Px|_1 = \text{Max}_{0 \leq i \leq n} \Delta_i^{-1} |\xi_{i+1} - \xi_i|. \quad (3.6)$$

<sup>5</sup> Komponenten  $\xi_i$  usw. sind bei  $i \leq 0$  bzw.  $i \geq n+1$  durch 0 zu ersetzen.

Aus (3.2) kann durch Multiplikation der Zeilen mit  $6(\Delta_{i-1} + \Delta_i)^{-1}$  für  $i = 1, 2, \dots, n$  und Subtraktion benachbarter Zeilen ein Gleichungssystem für die Differenzen  $\xi'_i = \xi_{i+1} - \xi_i$  hergeleitet werden, das wieder eine überwiegende Hauptdiagonale besitzt.<sup>6</sup>

$$(1 - \lambda_i) \xi'_{i-1} + (2 - \lambda_i + \lambda_{i+1}) \xi'_i + \lambda_{i+1} \xi'_{i+1} = \mu_{i+1} \eta_{i+1} - \mu_i \eta_i$$

$$(\lambda_i = \Delta_i / (\Delta_{i-1} + \Delta_i), \quad \mu_i = 6 / (\Delta_{i-1} + \Delta_i)).$$

Die analoge Schlussweise führt hier zu

$$\text{Max}_{0 \leq i \leq n} |\xi'_i| < \text{Max}_{0 \leq i \leq n} |\mu_{i+1} \eta_{i+1} - \mu_i \eta_i|.$$

Mit Hilfe der Darstellung

$$x(s) = x(s_i) + \int_{s_i}^s x' d\sigma$$

lässt sich daraus nach kurzer Rechnung

$$\text{Max}_{0 \leq i \leq n} |\xi'_i| \leq 8\bar{\pi} |x|_1$$

ableiten. Schliesslich ist  $\Delta_i \geq \bar{\pi}$ , so dass (3.6)

$$|Px|_1 \leq 8 \frac{\bar{\pi}}{\pi} |x|_1 \tag{3.7}$$

liefert. Damit ist gezeigt

**SATZ 1.** Die Orthogonalprojektion in  $H$  ist in  $C$  stets f.b.a., für gleichmässig  $c$ -reguläre Unterteilungen gilt dasselbe in  $C_1$ .

#### 4. ORTHOGONALPROJEKTION IN $H_A$

Unter den Voraussetzungen

- (i)  $p \geq \bar{p} > 0, q \geq 0$  in  $[0, 1]$ ,
- (ii)  $p, q$  messbar und  $p \leq \bar{p}, q \leq \bar{q}$  in  $[0, 1]$

ist der durch (1.8) festgelegte Hilbert-Raum  $H_A$  identisch mit  $H_1$ , die unterschiedliche Bezeichnung deutet auf die verschiedenen Normen hin.

Für den durch

$$\|x - Rx\|_A = \inf_{y \in S(\pi)} \|x - y\|_A \tag{4.1}$$

<sup>6</sup> Für  $i = 1$  und  $i = n$  sind die Gleichungen abzuändern, was hier nicht angegeben ist. Formel (3.7) bleibt richtig.

definierten Ritz-Operator  $R = R_n$  machen wir wieder den Ansatz  $Rx = \sum \xi_i \varphi_i$  und erhalten jetzt das System

$$\sum_k (a_{ik} + b_{ik}) \xi_k = \eta_i + \zeta_i \quad (4.2)$$

mit

$$\begin{aligned} a_{ik} &= \int p \varphi_i' \varphi_k' ds, & \eta_i &= \int p \varphi_i' x' ds, \\ b_{ik} &= \int q \varphi_i \varphi_k ds, & \zeta_i &= \int q \varphi_i x ds. \end{aligned} \quad (4.3)$$

Für (4.2) wollen wir auch

$$(\mathfrak{A} + \mathfrak{B}) \mathbf{x} = \boldsymbol{\eta} + \boldsymbol{\zeta} \quad (4.4)$$

schreiben.

Es ist zu erwarten, dass die Beiträge von  $\mathfrak{B}$  gegenüber denen von  $\mathfrak{A}$  wie die von  $\boldsymbol{\zeta}$  gegenüber denen von  $\boldsymbol{\eta}$  gering sind. Jedenfalls lässt sich unmittelbar abschätzen

$$|\mathfrak{B}| = \text{Max}_i \int q \varphi_i \left( \sum_k \varphi_k \right) ds \leq \bar{q} \bar{\pi}, \quad (4.5)$$

$$|\boldsymbol{\zeta}| \leq \bar{q} \bar{\pi} |\mathbf{x}|. \quad (4.6)$$

Ausgerechnet ergibt sich für die Elemente von  $\mathfrak{A}$ :

$$a_{ik} = \begin{cases} -\Delta_{i-1}^{-1} p_{i-1} & k = i - 1 \\ \Delta_{i-1}^{-1} p_{i-1} + \Delta_i^{-1} p_i & k = i \\ -\Delta_i^{-1} p_i & k = i + 1 \\ 0 & \text{sonst} \end{cases} \quad (4.7)$$

mit

$$p_i = \Delta_i^{-1} \int_{s_i}^{s_{i+1}} p(s) ds.$$

Daraus ist erkennbar, dass  $\mathfrak{A}$  von monotoner Art ist, vgl. dazu Collatz [10], S. 42 ff. Da ausserdem  $\mathfrak{B}$  eine positive Matrix ist, kann auf

$$|(\mathfrak{A} + \mathfrak{B})^{-1}| \leq |\mathfrak{A}^{-1}|$$

geschlossen werden. Für die Lösung  $\mathbf{x}$  von (4.4) gilt dann

$$|\mathbf{x}| \leq (1 + |\mathfrak{A}^{-1}| |\mathfrak{B}|) (|\mathfrak{A}^{-1} \boldsymbol{\eta}| + |\mathfrak{A}^{-1} \boldsymbol{\zeta}|). \quad (4.8)$$

Analog zur Konstruktion der Greenschen Funktion des Randwertproblems (1.9) kann hier die Inverse  $\mathfrak{G} = \mathfrak{A}^{-1}$  explizit angegeben werden. Wir finden

$$g_{ik} = \gamma \begin{cases} \chi_i \psi_k & i \leq k \\ \chi_k \psi_i & i > k \end{cases} \quad (4.9)$$

mit

$$\chi_i = \sum_0^{i-1} \Delta_k p_k^{-1},$$

$$\psi_i = \sum_i^n \Delta_k p_k^{-1},$$

$$\gamma^{-1} = \sum_0^n \Delta_k p_k^{-1}.$$

Für später merken wir  $0 < \chi_i, \psi_i$  und  $\chi_i + \psi_i = \gamma^{-1}$  sowie  $p \leq \gamma \leq \bar{p}$  an. Dann ist

$$|\mathfrak{A}^{-1}| = |\mathfrak{G}| \leq n\gamma^{-1} \leq 1/(p\pi)$$

und somit

$$1 + |\mathfrak{A}^{-1}||\mathfrak{B}| \leq 1 + \frac{\bar{q}\bar{\pi}}{p\pi} \tag{4.10}$$

$$|\mathfrak{A}^{-1}\mathfrak{z}| \leq \frac{\bar{q}\bar{\pi}}{p\pi} |x| \tag{4.11}$$

erkennbar. Zur Abschätzung des noch ausstehenden Termes  $\bar{x} = \mathfrak{A}^{-1}\mathfrak{y}$  mit den Komponenten  $\bar{\xi}_i$  berücksichtigen wir die Darstellung

$$\eta_i = \rho_{i-1} - \rho_i \quad \text{mit} \quad \rho_i = \Delta_i^{-1} \int_{s_i}^{s_{i+1}} p x' ds.$$

Mit (4.9) ergibt sich so

$$\begin{aligned} \bar{\xi}_i &= \sum_k g_{ik} \eta_k \\ &= \gamma \psi_i \sum_0^{i-1} \Delta_k p_k^{-1} \rho_k - \gamma \chi_i \sum_i^n \Delta_k p_k^{-1} \rho_k. \end{aligned} \tag{4.12}$$

Eine erste Möglichkeit zum Abschätzen von  $\mathfrak{A}^{-1}\mathfrak{y}$  beruht auf der Beziehung

$$|\rho_i| \leq \bar{p} |x|_1,$$

woraus in Verbindung mit (4.12)

$$|\mathfrak{A}^{-1}\mathfrak{y}| \leq \frac{\bar{p}}{p} |x|_1 \tag{4.13}$$

folgt.

Um in (4.13) eine Abschätzung mit  $|x|$  anstelle von  $|x|_1$  zu gewinnen, machen wir nun die weitere Voraussetzung

(iii)  $p'$  messbar und  $|p'(s)| \leq \bar{p}_1$  in  $[0, 1]$ .

Dann lässt sich schreiben

$$\begin{aligned} \rho_k &= \rho_k^1 + \rho_k^2 \\ &= \Delta_k^{-1} \{ p(s_{k+1}) x(s_{k+1}) - p(s_k) x(s_k) \} - \Delta_k^{-1} \int_{s_k}^{s_{k+1}} p' x ds. \end{aligned}$$

Die obige Überlegung liefert für die 2. Komponenten von  $\bar{\xi}_i = \bar{\xi}_i^1 + \bar{\xi}_i^2$  unmittelbar

$$|\bar{\xi}_i^2| \leq \frac{\bar{p}_1}{p} |x|. \quad (4.14)$$

Für die 1. Komponente ergibt (4.12)

$$\begin{aligned} \bar{\xi}_i^1 &= p_i^{-1} p(s_i) x(s_i) \\ &+ \gamma \psi_i \sum_1^i (p_{k-1}^{-1} - p_k^{-1}) p(s_k) x(s_k) \\ &- \gamma \chi_i \sum_{i+1}^n (p_{k-1}^{-1} - p_k^{-1}) p(s_k) x(s_k) \end{aligned}$$

und daher

$$|\bar{\xi}_i^1| \leq \bar{p} (p^{-1} + n \text{Max } |p_{k-1}^{-1} - p_k^{-1}|) |x|. \quad (4.15)$$

Schliesslich ist bei  $\bar{\pi} \bar{p}_1 \leq p$  ohne Schwierigkeit

$$|p_{k-1}^{-1} - p_k^{-1}| \leq 4 \bar{\pi} \frac{\bar{p}_1}{p^2}$$

zu erkennen, so dass wir mit (4.14) und (4.15)

$$|\mathfrak{A}^{-1} \eta| \leq \left\{ \frac{\bar{p}}{p^2} \left( p + 4 \frac{\bar{\pi}}{\bar{\pi}} \bar{p}_1 \right) + \frac{\bar{p}_1}{p} \right\} |x| \quad (4.16)$$

gewonnen haben. Damit liefert (4.8) in Verbindung mit (4.10), (4.11) und (4.13) bzw. (4.16) den

*SATZ 2. Für c-reguläre Unterteilungen existieren nur von c und p,  $\bar{p}$ ,  $\bar{q}$  bzw. zusätzlich von  $\bar{p}_1$  abhängige Konstante  $K_{0.1}$  und  $K_{0.0}$ , so dass für alle  $x \in C_1$  die Abschätzungen*

$$|Rx| \leq K_{0.1} |x|_1 \quad (4.17)$$

$$|Rx| \leq K_{0.0} |x| \quad (4.18)$$

*gültig sind.*

Der Operator R ist zunächst nur in  $C_1$  definiert. Da aber  $K_{0.0}$  von x unabhängig ist und  $C_1$  dicht in C ist, kann R in eindeutiger Weise zu einem in C stetigen

Operator fortgesetzt werden. Für diesen Operator, wir behalten den Buchstaben  $R$  bei, gilt dann  $|R|_{0,0} \leq K_{0,0}$ , so dass also (4.18) auch in der Form ausgesprochen werden kann.

*ZUSATZ 2. Bei den Voraussetzungen (i)–(iii) ist der Ritz-Operator für  $c$ -reguläre Unterteilungen f.b.a. in  $C$ .*

Beim Nachweis der Beschränktheit von  $|R|_{1,1}$  wird uns (4.17) von Nutzen sein, da wir dann ohne die Voraussetzung (iii) auskommen können. Mit der Lösung  $\{\xi_i\}$  von (4.3) gilt nämlich jedenfalls

$$\left| \sum_k b_{ik} \xi_k \right| \leq |\mathfrak{B}| |Rx| \leq \bar{q} K_{0,1} \bar{\pi} |x|_1, \quad (4.19)$$

$$|\zeta_i| \leq \bar{q} \bar{\pi} |x| \leq \bar{q} \bar{\pi} |x|_1. \quad (4.20)$$

Die Ableitung  $(Rx)'$  ist stückweise konstant und zwar ist

$$(Rx)' = \xi_i' = \Delta_i^{-1}(\xi_{i+1} - \xi_i) \quad \text{für } s \in (s_i, s_{i+1}). \quad (4.21)$$

Diese Grössen  $\xi_i'$  genügen den Gleichungen

$$p_{i-1} \xi_{i-1}' - p_i \xi_i' = \rho_{i-1} - \rho_i - \eta_i' \\ \left( \eta_i' = \sum_k b_{ik} \xi_k - \zeta_i \right).$$

Daher ist

$$p_i \xi_i' = \rho_i + \lambda + \sum_1^i \eta_j'$$

mit einem Parameter  $\lambda$ , der sich aus der Bedingung

$$0 = \xi_{n+1} - \xi_0 = \sum_0^n \Delta_i \xi_i'$$

errechnet. Mit (4.19) und (4.20) ergibt eine einfache Rechnung

$$|\lambda| \leq \frac{\bar{p}}{p} (\bar{p} + n\bar{\pi}(1 + K_{0,1})\bar{q}) |x|_1$$

und damit auch

$$|Rx|_1 = \text{Max}_{0 \leq i \leq n} |\xi_i'| \leq 2 \frac{\bar{p}}{p} \left( \bar{p} + \frac{\bar{\pi}}{p} (1 + K_{0,1}) \bar{q} \right) |x|_1.$$

SATZ 3. Für  $c$ -reguläre Unterteilungen existiert eine nur von  $c$  und  $p, \bar{p}, \bar{q}$  abhängige Konstante  $K_{1,1}$ , so dass für alle  $x \in C_1$  die Abschätzung

$$|Rx|_1 \leq K_{1,1} |x|_1 \quad (4.22)$$

gültig ist.

Wieder können wir das Resultat durch  $|R|_{1,1} \leq K_{1,1}$  bzw. in dem Zusatz ausdrücken.

ZUSATZ 3. Bei den Voraussetzungen (i), (ii) ist  $R$  für  $c$ -reguläre Unterteilungen f.b.a. in  $C_1$ .

In den obigen Abschätzungen ist bei  $p=1, q=0$  der Fall der Spline-Interpolation enthalten. Dann ist  $K_{0,0} = 1$ , was sich nicht verbessern lässt, entsprechend kann in (2.7)  $K=2$  gewählt werden. Wie sich etwa durch Betrachtung konvexer Funktionen  $x$  überlegen lässt, kann diese Konstante bei Spline-Interpolation nicht verbessert werden.

## 5. FEHLERABSCHÄTZUNGEN BEI RANDWERTPROBLEMEN

Wir wenden uns nun der Aufstellung von Fehlerabschätzungen beim Ritz-Verfahren für das Sturm-Liouville-Randwertproblem (1.9) zu. Unter den Voraussetzungen (i)–(iii) aus §4 besitzt die Lösung von (1.9) fast überall eine beschränkte 2. Ableitung, jedenfalls ist die 1. Ableitung Lipschitz-beschränkt, wenn  $f$  beschränkt und messbar ist. Wir wollen die Konstante explizit abschätzen. Aus

$$\|x\|_A^2 = (x, f)$$

folgt zunächst

$$\|x\|_1 \leq \frac{1}{p} \|f\| \leq \frac{1}{p} |f|$$

und daraus

$$|x| \leq \frac{1}{p} |f|.$$

Weiter hat  $x'$  wegen der Randbedingungen (1.9<sub>2</sub>) in  $(0,1)$  mindestens eine Nullstelle, es sei  $x'(\sigma) = 0$ . Dann kann wegen

$$x'(s) = \frac{1}{p(s)} \int_{\sigma}^s (-f + qx) dt$$

für den Stetigkeitsmodul  $\omega_1(x', h)$  auf

$$\omega_1(x', h) \leq K_0 |f| h \quad \left( K_0 = \frac{1}{p^3} (p + \bar{q})(p + \bar{p}_1) \right)$$

geschlossen werden. Die Beziehung

$$\omega_2(x, h) \leq \int_0^h \omega_1(x', 2t) dt$$

zwischen dem 2. Stetigkeitsmodul von  $x$  und dem 1. von  $x'$  liefert daher

$$\omega_2(x, h) \leq K_0 |f| h^2.$$

Die Beziehung

$$E_0(x, S(\pi)) = \inf_{y \in S(\pi)} |x - y| \leq \frac{1}{\ln 2} \int_0^{\bar{\pi}} t^{-1} \omega_2(x, t) dt$$

aus [14] gibt uns daher für den Approximationsgrad  $E_0$

$$E_0(x, S(\pi)) \leq \frac{1}{2 \ln 2} K_0 |f| \bar{\pi}^2,$$

und daher haben wir

$$|x - Rx| \leq \frac{1}{2 \ln 2} (1 + K_{0.0}) K_0 |f| \bar{\pi}^2.$$

Entsprechend finden wir

$$|x - Rx|_1 \leq (1 + K_{1.1}) K_0 |f| \bar{\pi}.$$

Für  $c$ -reguläre Unterteilungen ist  $\bar{\pi} \leq c\pi \leq cn^{-1}$ .

Wir beachten noch, dass  $\xi_i$  den Wert von  $Rx$  an der Stützstelle  $s_i$  angibt. Wegen  $\varphi_i \varphi_k = 0$  bei  $|i - k| > 1$  hat das Gleichungssystem (4.2) nur in der Haupt- und den 1. Nebendiagonalen von Null verschiedene Elemente. Schliesslich ist die rechte Seite in (4.2)

$$\eta_i + \zeta_i = \int_0^1 \varphi_i(-(px)') + qx) ds = \int_0^1 \varphi_i f ds$$

als Integralmittel über  $f$  darstellbar. Zusammengefasst haben wir gefunden

**SATZ 4.** *Es seien die Voraussetzungen (i)–(iii) erfüllt. Dann lassen sich die Lösungen des Randwertproblems (1.9) für beschränkte messbare Funktionen  $f$  durch Polygonzüge mit  $n$  Stützstellen derart approximieren, dass die maximale Abweichung der Funktionswerte durch  $k_1 |f| n^{-2}$  und die der 1. Ableitungen durch  $k_2 |f| n^{-1}$  abgeschätzt werden können.—Die Funktionswerte der Näherungen an den Stützstellen bestimmen sich aus einem tridiagonalen System von Differenzgleichungen, dessen Koeffizienten Integralmittel über  $p, q, f$  sind.*

#### LITERATUR

1. J. H. AHLBERG, E. N. NILSON, UND J. L. WALSH, Convergence properties of generalized splines. *Proc. Natl. Acad. Sci. U.S.A.* **54** (1965), 344–350.
2. J. H. AHLBERG, E. N. NILSON, UND J. L. WALSH, "The Theory of Splines and Their Applications." Academic Press, New York, 1967.
3. G. ALEXITS, "Konvergenzprobleme der Orthogonalreihen." VEB Deutscher Verlag der Wissenschaften, Berlin, 1960.

4. G. ALEXITS, Einige Beiträge zur Approximationstheorie. *Acta Sci. Math.* **26** (1965), 212–224.
5. A. K. AZIZ UND B. E. HUBBARD, Bounds for the solution of the Sturm-Liouville problem with application to finite difference methods. *J. Soc. Ind. Appl. Math.* **13** (1964), 163–178.
6. I. BABUSKA, M. PRAGER, UND E. VITASEK, “Numerical Processes in Differential Equations.” Wiley (Interscience), New York, 1966.
7. G. H. BIRKHOFF UND C. DE BOOR, Error bounds for spline interpolation. *J. Math. Mech.* **13** (1964), 827–836.
8. G. H. BIRKHOFF UND C. DE BOOR, Piecewise polynomial interpolation and approximation. “Proc. Symp. Approx. Functions, Michigan, 1964.” Elsevier, Amsterdam, 1965.
9. F. E. BROWDER, Approximation-solvability of nonlinear functional equations in normed linear spaces. *Arch. Rat. Mech. Anal.* **26** (1967), 33–42.
10. L. COLLATZ, “The Numerical Treatment of Differential Equations.” Springer, Berlin, 1960.
11. A. HAAR, Zur Theorie der orthogonalen Funktionensysteme. *Math. Ann.* **69** (1910), 331–371; **71** (1912), 38–53.
12. G. G. LORENTZ, “Approximation of Functions.” Holt, Rinehart, and Winston, New York, 1966.
13. J. NITSCHKE, Verfahren von Ritz und Spline-Interpolation bei Sturm-Liouville-Randwertproblemen. Erscheint in *Num. Math.*
14. J. NITSCHKE, Sätze von Jackson-Bernstein-Typ für die Approximation mit Spline-Funktionen. (Erscheint demnächst.)
15. M. H. SCHULTZ UND R. S. VARGA,  $L$ -Splines. *Num. Math.* **10** (1967), 345–369.
16. M. ZLAMAL, Discretization and error estimates for boundary value problems of the second order. *Calcolo Consiglio Nazl. Ric.* **4** (1967), 541–550.