JOURNAL OF APPROXIMATION THEORY 9, 126–148 (1973)

Über rationale Tschebyscheff-Approximation mehrerer Funktionen*

HANS-PETER BLATT

Institut für Angewandte Mathematik I der Universität Erlangen-Nürnberg, 852 Erlangen, Martensstraße 1, Deutschland

Communicated by G. Meinardus

Received July 15, 1971

B sei ein kompakter Raum und $S(B)$ der lineare Raum der auf B beschränkten, reellwertigen Funktionen. Durch $\|f\| := \sup_{x \in B} |f(x)|$ wird $S(B)$ ein normierter Raum. $C(B)$ sei die in $S(B)$ liegende Teilmenge der stetigen Funktionen, V eine Teilmenge von $C(B)$.

In [1] wurde die Tschebyscheff-Approximation einer Menge von gleichmäßig beschränkten, komplexwertigen Funtionen durch Funktionen aus V auf die Approximation einer oberhalbstetigen Abbildung $H: B \to \mathscr{K}(\mathbb{C})$ zurückgeführt (B mußte zusätzlich das 1. Abzählbarkeitsaxiom erfüllen). Dabei bedeutet, falls X ein topologischer Raum ist,

$$\mathscr{K}(X) := \{E \subset X \mid E \text{ kompakt und } E \neq \varnothing\}.$$

Außerdem heißt für topologische Räume X und Y eine Abbildung $H: X \to \mathscr{K}(Y)$ *oberhalbstetig in* $x \in X$, wenn zu jeder in Y offenen Menge G mit $H(x) \subset G$ eine Umgebung U von x existiert, so daß $H(U) \subset G$ ist. Unter $H(E)$ verstehen wir dabei für $E \subset X$ die Bildmenge

$$H(E) := \{y \in Y \mid \text{es gibt ein } x \in E \text{ mit } y \in H(x)\}.$$

Aus der Definition von H in [1] ist natürlich klar, daß man sich auf die Approximation einer oberhalbstetigen Abbildung $H: B \to \mathscr{K}(\mathbb{R})$ beschränken kann, falls alle betrachteten Funktionen, wie in unserm Fall, reellwertig sind. Anderseits beweisen Diaz und McLaughlin in [8] für metrisches B, daß man sich auf den Fall der gleichzeitigen Approximation einer oberhalbstetigen Funktion $f^+: B \to \mathbb{R}$ und einer unterhalbstetigen Funktion $f^-: B \to \mathbb{R}$ mit $f^+(x) \geqslant f^-(x)$ für alle $x \in B$ zurückziehen kann (vgl. Rémès [13], Golomb [10], Dunham [9]). Dabei heißt eine Funktion $f: B \to \mathbb{R}$ *oberhalbstetig* (*unterhalbstetig*) *in* $x \in B$, wenn zu jeder reellen

* Zweiter Teil der am Institut für Angewandte Mathematik der Universität des Saarlandes in Saarbrücken angefertigten Dissertation des Verfassers [1970].

Zahl $h > f(x)$ $(h < f(x))$ eine Umgebung U von x existiert mit $f(z) < h$ $(f(z) > h)$ für alle $z \in U$.

Wir betrachten daher zunächst zwei formal verschiedene Aufgabenstellungen:

(A1) Zu der oberhalbstetigen Abbildung $H : B \to \mathscr{K}(\mathbb{R})$ suchen wir ein Element $v_0 \in V$ der Art, daß

$$\sup_{x \in B} \bar{d}(v_0(x), H(x)) \leqslant \sup_{x \in B} \bar{d}(v(x), H(x))$$

für jedes Element $v \in V$.
Dabei setzten wir $\bar{d}(v(x), H(x)) := \sup_{z \in H(x)} | v(x) - z |$.

(A2) Zu der oberhalbstetigen Funktion $f^+ : B \to \mathbb{R}$ und der unterhalbstetigen Funktion $f^- : B \to \mathbb{R}$ mit $f^+(x) \geqslant f^-(x)$ für alle $x \in B$ suchen wir ein Element $v_0 \in V$ mit

$$\max\{\| f^+ - v_0 \|, \| f^- - v_0 \|\} \leqslant \max\{\| f^+ - v \|, \| f^- - v \|\}$$

für jedes $v \in V$.

Im ersten Teil wird gezeigt, wie man beide Aufgaben ineinander überführt. Auf diese Weise können wir uns auf die für den reellen Fall angemessenere Aufgabenstellung (A2) beschränken und schon bekannte Ergebnisse aus [1] verwenden. Über Charakterisierungssätze, die sich aus entsprechenden in [1] ergeben, gelangt man über einem Intervall zu der Charakterisierung einer Minimallösung durch eine Alternante oder einen "Straddlepunkt" (Dunham [9] führte diese Bezeichnung ein). Für die Eindeutigkeit dieser Minimallösung ist die Existenz einer Alternanten einer bestimmten Länge hinreichend, ein Ergebnis, welches eine Verallgemeinerung eines Eindeutigkeitssatzes von Cheney und Loeb [2] darstellt. Mit Hilfe des strengen Eindeutigkeitssatzes zeigen wir, daß der Bestapproximationsoperator, der jedem (f^+, f^-) die Menge der Minimallösungen zuordnet, stetig ist bei (f^+, f^-), falls $f^+ = f^- \in V$ oder (f^+, f^-) normal ist (Definition 4.1). Zur konstruktiven Berechnung der Simultanapproximation verallgemeinern wir ein linear konvergierendes Verfahren von Cheney und Loeb [4], das auf der konvexen Optimierung beruht.

1. Reelle Simultanapproximation

Zuerst wollen wir noch einige Bezeichnungen einführen:

$$\Delta(v) := \begin{cases} \max\limits_{x \in B} \bar{d}(v(x), H(x)) & \text{bei Aufgabe (A1)} \\ \max\{\| f^+ - v \|, \| f^- - v \|\} & \text{bei Aufgabe (A2)} \end{cases}$$

$$\mathfrak{M}(v) := \begin{cases} \{x \in B \mid \bar{d}(v(x), H(x)) = \Delta(v)\} & \text{bei Aufgabe (A1)} \\ \{x \in B \mid \max\{|f^+(x) - v(x)|, |f^-(x) - v(x)|\} = \Delta(v)\} & \\ & \text{bei Aufgabe (A2)} \end{cases}$$

Unter der Minimalabweichung verstehen wir bei (A1) die Zahl

$$\rho_V(H) := \inf_{v \in V} \Delta(v),$$

bei (A2) bezeichnen wir sie mit $\rho_V(f^+, f^-)$.

$$D[H, v] := \{(x, z) \in B \times \mathbb{R} \mid z \in H(x) \quad \text{und} \quad | v(x) - z | = \Delta(v)\}.$$

Wenn S eine Teilmenge eines linearen Raumes ist, dann bezeichnen wir mit conv(S) die konvexe Hülle von S.

Außerdem benutzen wir die Abkürzung o.B.d.A. für "ohne Beschränkung der Allgemeinheit".

Wir wollen uns zunächst mit der Aufgabe (A1) beschäftigen, also eine oberhalbstetige Abbildung $H : B \to \mathscr{K}(\mathbb{R})$ bezüglich V approximieren. In diesem Fall können wir die Aufgabe auch folgendermaßen umschreiben:

$$\begin{aligned} \max_{x \in B} \bar{d}(v(x), H(x)) &= \max_{x \in B} \max_{z \in H(x)} | v(x) - z |, \\ &= \max_{x \in B} \max\{| v(x) - f^+(x)|, | v(x) - f^-(x)|\}, \\ &= \max\{\| f^+ - v \|, \| f^- - v \|\}. \end{aligned}$$

Dabei sind die Funktionen f^+ und f^- für $x \in B$ definiert als:

$$f^+(x) := \max_{z \in H(x)} z, \tag{1.1}$$

$$f^-(x) := \min_{z \in H(x)} z, \tag{1.2}$$

und haben deshalb die Eigenschaft:

$$f^+(x) \geqslant f^-(x) \qquad \text{für alle } x \in B.$$

HILFSSATZ 1.1. *f^+ wie in (1.1) ist eine oberhalbstetige Funktion.*

Beweis. Sei $x_0 \in B$ und $h > f^+(x_0)$. Setzen wir $\epsilon := h - f^+(x_0)$, dann ist $G :=] - \epsilon + f^-(x_0), f^+(x_0) + \epsilon[$ eine offene Umgebung von $H(x_0)$ in $\mathbb{R}$. Also existiert eine Umgebung $U(x_0)$ mit $H(x) \subset G$ für alle $x \in U(x_0)$. Damit ist $f^+(x) \in G$ für alle $x \in U(x_0)$, d.h. $f^+(x) < h$ für alle $x \in U(x_0)$.

Ganz analog beweist man

HILFSSATZ 1.2. *f^- wie in (1.2) ist eine unterhalbstetige Funktion.*

Unser Spezialfall der Aufgabenstellung (A1) ist also übergeführt zu der Aufgabe (A2). Wir fassen das Ergebnis zusammen zu

SATZ 1.1. *Ist $C(B)$ der lineare Raum der auf dem kompakten Raum B stetigen, reellwertigen Funktionen, $V \subset C(B)$ und $H : B \to \mathscr{K}(\mathbb{R})$ oberhalbstetig, dann sind folgende Aufgaben äquivalent*:

(α) *Minimiere* $\max_{x \in B} \bar{d}(v(x), H(x))$ *bezüglich* V!

(β) *Minimiere* $\max\{\|f^+ - v\|, \|f^- - v\|\}$ *bezüglich* V!

Dabei werden f^+ und f^- durch (1.1) *bzw.* (1.2) *definiert und haben die Eigenschaft: $f^+(x) \geqslant f^-(x)$ für $x \in B$. Darüberhinaus ist f^+ oberhalb- und f^- unter halbstetig.*

Gehen wir nun umgekehrt von der Aufgabenstellung (A2) aus, dann definieren wir

$$H(x) := [f^-(x), f^+(x)] \quad \text{for} \quad x \in B. \tag{1.3}$$

H ist eine oberhalbstetige Abbildung von B in $\mathscr{K}(\mathbb{R})$ und es gilt für alle $v \in V$:

$$\max\{\|f^+ - v\|, \|f^- - v\|\} = \max_{x \in B} \bar{d}(v(x), H(x)).$$

Also ergibt sich

SATZ 1.2. *Ist $C(B)$ der lineare Raum der auf dem kompakten Raum B stetigen, reellwertigen Funktionen, $V \subset C(B)$, f^+ eine auf B oberhalbstetige und f^- eine auf B unterhalbstetige Funktion mit $f^+(x) \geqslant f^-(x)$ für $x \in B$, dann sind äquivalent*:

(α') *Minimiere* $\max\{\|f^+ - v\|, \|f^- - v\|\}$ *bezüglich* V!

(β') *Minimiere* $\max_{x \in B} \bar{d}(v(x), H(x))$ *bezüglich* V!

Dabei wird H durch (1.3) *definiert.*

Wir beschränken uns daher im weiteren Verlauf auf die für den reellen Fall angemessenere Aufgabenstellung (A2). In $S(B) \times S(B)$ führen wir durch

$$\|(f, g)\| = \max\{\|f\|, \|g\|\}$$

eine Norm ein. Mit $O(B)$ bezeichnen wir die Menge der auf B oberhalbstetigen Funktionen, mit $U(B)$ die Menge der auf B unterhalbstetigen Funktionen. Weiter setzen wir

$$OU(B) := \{(f^+, f^-) \in O(B) \times U(B) \mid f^+(x) \geqslant f^-(x) \text{ für } x \in B\}$$

Da $OU(B) \subset S(B) \times S(B)$, können wir die Aufgabenstellung (A2) auch so formulieren:

Zu $(f^+, f^-) \in OU(B)$ ist ein Element $v_0 \in V$ gesucht mit

$$\begin{aligned} \|(f^+, f^-) - (v_0, v_0)\| &= \|(f^+ - v_0, f^- - v_0)\| \\ &\leqslant \|(f^+ - v, f^- - v)\| = \|(f^+, f^-) - (v, v)\| \end{aligned}$$

für alle $v \in V$.

Wir können die Aufgabe (A2) noch etwas umschreiben, indem wir für $x \in B$

$$f(x) := (f^+(x) + f^-(x))/2 \tag{1.4}$$

$$Z(x) := (f^+(x) - f^-(x))/2 \text{ setzen.} \tag{1.5}$$

Also ist $Z(x) \geqslant 0$ in B und es gilt:

$$\max\{\| f^+ - v \|, \| f^- - v \|\} = \max_{x \in B}\{| f(x) - v(x)| + Z(x)\}. \tag{1.6}$$

Diese Umschreibung werden wir in einigen Sätzen und Definitionen benutzen. Als erste Anwendung von (1.6) ergibt sich sofort:

SATZ 1.3. $\rho_V(f^+, f^-) \geqslant \| Z \|$.

2. CHARAKTERISIERUNG DER MINIMALLÖSUNG BEI RATIONALER SIMULTANAPPROXIMATION

Wir wollen die reelle Simultanapproximation noch weiter spezialisieren und jetzt die rationale Simultanapproximation betrachten:

$P = \operatorname{span}(u_1, u_2, ..., u_n)$ sei ein n-dimensionaler und

$Q = \operatorname{span}(u_{n+1}, u_{n+2}, ..., u_{n+m})$ ein m-dimensionaler Unterraum des reellen linearen Raums $C(B)$. Die Menge der Approximationsfunktionen ist dann

$$R := \left\{ r(a, \cdot) = \frac{p(a, \cdot)}{q(a, \cdot)} = \frac{\sum_{i=1}^{n} a_i u_i}{\sum_{i=n+1}^{n+m} a_i u_i} \;\middle|\; q(a, x) > 0 \text{ für alle } x \in B \right\}.$$

Dabei ist $a = (a_1, a_2, ..., a_{n+m}) \in \mathbb{R}^{n+m}$. Damit R nicht leer ist, setzen wir zusätzlich voraus, daß Q mindestens eine Funktion enthält, die in ganz B größer als null ist. Sind nun $r(a, \cdot)$ und $r(b, \cdot)$ aus R, und setzt man $a(t) := a + t(b - a)$, so wird

$$r(a(t), x) = ((1 - t)\, p(a, x) + tp(b, x))/((1 - t)\, q(a, x) + tq(b, x)).$$

Für $0 \leqslant t \leqslant 1$ gehört $r(a(t), \cdot)$ zu R und mit

$$g(x, t) := q(b, x)/((1 - t)\, q(a, x) + tq(b, x))$$

wird

$$r(a(t), x) = (1 - tg(x, t))\, r(a, x) + tg(x, t)\, r(b, x).$$

Also ist R asymptotisch konvex (Meinardus und Schwedt [12]), $a(t)$ im Punkt 0 Fréchet-differenzierbar und $a(0) = a$. Darüberhinaus ist die Parametermenge A, in der a variieren kann, offen in $\mathbb{R}^{n+m}$. Jedes Element $r(a, \cdot)$ aus R besitzt eine Fréchet-Ableitung nach a, nämlich

$$F[b; a, \cdot] = \sum_{i=1}^{n+m} b_i \frac{\partial r(a, \cdot)}{\partial a_i} \qquad \text{für} \quad b \in \mathbb{R}^{n+m}.$$

Setzen wir

$$\operatorname{grad} r(a, \cdot) := \left(\frac{\partial r(a, \cdot)}{\partial a_1}, \ldots, \frac{\partial r(a, \cdot)}{\partial a_{n+m}}\right),$$

dann können wir mit dem Skalarprodukt $\langle . \rangle$ schreiben:

$$F[b; a, \cdot] = \langle b, \operatorname{grad} r(a, \cdot)\rangle. \tag{2.1}$$

Für spätere Zwecke wollen wir $\operatorname{grad} r(a, \cdot)$ noch ausrechnen:

$$\frac{\partial r(a, \cdot)}{\partial a_i} = \begin{cases} \dfrac{u_i}{q(a, \cdot)} & \text{für} \quad i = 1, 2, \ldots, n \\[2ex] -\dfrac{p(a, \cdot)}{q(a, \cdot)^2}\, u_i & \text{für} \quad i = n + 1, \ldots, n + m, \end{cases}$$

also

$$\operatorname{grad} r(a, \cdot) = \frac{1}{q(a, \cdot)} (u_1, \ldots, u_n, -r(a, \cdot)\, u_{n+1}, \ldots, -r(a, \cdot)\, u_{n+m}). \tag{2.2}$$

Zur Abkürzung wollen wir in Zukunft auch $\mathfrak{M}(a)$ anstatt $\mathfrak{M}(r(a, \cdot))$ und $\Delta(a)$ anstatt $\Delta(r(a, \cdot))$ für $r(a, \cdot) \in R$ schreiben.

SATZ 2.1. *$r(a, \cdot) \in R$ ist genau dann eine Minimallösung zu $(f^+, f^-) \in OU(B)$, wenn für jedes $r(b, \cdot) \in R$ ein $x \in \mathfrak{M}(a)$ existiert mit*

$$(f(x) - r(a, x))(r(b, x) - r(a, x)) \leqslant 0$$

(f wie in (1.4)).

Beweis. Wir definieren H wie in (1.3).

(a) *Notwendigkeit.* Nach Satz 1.2 ist $r(a, \cdot)$ Minimallösung für H, also existiert nach Satz 3.5 in [1] ein

$$(x, z) \in D[H, r(a, \cdot)] \qquad \text{mit} \quad (z - r(a, x))(r(b, x) - r(a, x)) \leqslant 0.$$

Dabei ist $z = f^+(x)$ oder $z = f^-(x)$.

(1) $z = f^+(x)$:

Dann ist $|f^-(x) - r(a, x)| \leqslant |f^+(x) - r(a, x)|$ und aus

$$f(x) - r(a, x) = (f^+(x) - r(a, x))/2 + (f^-(x) - r(a, x))/2$$

folgt $\operatorname{sgn}(f(x) - r(a, x)) = \operatorname{sgn}(f^+(x) - r(a, x))$. Daraus ergibt sich für $x \in \mathfrak{M}(a)$:

$$(f(x) - r(a, x))(r(b, x) - r(a, x)) \leqslant 0.$$

(2) $z = f^-(x)$:

Dann ist $|f^+(x) - r(a, x)| \leqslant |f(x) - r(a, x)|$ und

$$\operatorname{sgn}(f(x) - r(a, x)) = \operatorname{sgn}(f^-(x) - r(a, x)).$$

Daraus folgt wieder die Behauptung.

Wie in [1] ist dabei $\operatorname{sgn} h$ als $+1$ oder -1 wählbar, falls $h = 0$ ist.

(b) *Hinlänglichkeit.* Sei also $x \in \mathfrak{M}(a)$ mit

$$(f(x) - r(a, x))(r(b, x) - r(a, x)) \leqslant 0$$

oder

$$\left(\frac{f^+(x) - r(a, x)}{2} + \frac{f^-(x) - r(a, x)}{2}\right)(r(b, x) - r(a, x)) \leqslant 0.$$

Ist nun $|f^+(x) - r(a, x)| = \Delta(a)$ und $|f^-(x) - r(a, x)| < \Delta(a)$, dann ist $(x, z) := (x, f^+(x)) \in D[H, r(a, \cdot)]$ mit

$$(z - r(a, x))(r(b, x) - r(a, x)) \leqslant 0. \tag{2.3}$$

Ist $|f^+(x) - r(a, x)| < |f^-(x) - r(a, x)| = \Delta(a)$, dann erfüllt

$$(x, z) := (x, f^-(x)) \in D[H, r(a, \cdot)]$$

die Ungleichung (2.3). Falls dagegen

$$\Delta(a) = |f^+(x) - r(a, x)| = |f^-(x) - r(a, x)|$$

ist, dann erfüllt entweder $(x, z) := (x, f^-(x)) \in D[H, r(a, \cdot)]$ oder $(x, z) := (x, f^+(x)) \in D[H, r(a,\cdot)]$ die Relation (2.3). Nach Satz 3.2 in [1] ist damit $r(a, \cdot)$ Minimallösung zu H, also auch zu (f^+, f^-).

Als nächsten Satz läßt sich die Charakterisierung mit der Fréchet-Ableitung (Satz 3.11 in [1]) übertragen.

SATZ 2.2. *$r(a, \cdot) \in R$ ist genau dann eine Minimallösung zu $(f^+, f^-) \in OU(B)$, falls für jedes $b \in \mathbb{R}^{n+m}$ ein $x \in \mathfrak{M}(a)$ existiert mit*

$$(f(x) - r(a, x))\langle b, \operatorname{grad} r(a, x)\rangle \leqslant 0.$$

Der Beweis dieses Satzes verläuft analog wie beim vorigen Satz.

SATZ 2.3. *$r(a, \cdot) \in R$ ist genau dann eine Minimallösung zu $(f^+, f^-) \in OU(B)$, falls $0 \in \operatorname{conv}(S')$. Dabei ist $S' := \{(f(x) - r(a, x)) \operatorname{grad} r(a, x) \mid x \in \mathfrak{M}(a)\}$.*

Beweis. Wir definieren wieder H wie in (1.3).

(*a*) $r(a, \cdot)$ sei Minimallösung zu $(f^+, f^-) \in OU(B)$. Da dann $r(a, \cdot)$ auch Minimallösung zu H ist, liegt nach Satz 3.10 in [1] der Nullvektor in $\operatorname{conv}(S'')$ mit

$$S'' := \{(z - r(a, x)) \operatorname{grad} r(a, x) \mid (x, z) \in D[H, r(a, \cdot)]\}.$$

Wir machen nun zwei Fallunterscheidungen.

(1) Es gibt ein $x \in B$ mit $(x, f^+(x))$ und $(x, f^-(x))$ aus $D[H, r(a, \cdot)]$:
Dann ist $f(x) - r(a, x) = 0$, also $0 \in \operatorname{conv}(S')$.

(2) Es gibt kein $x \in B$ mit $(x, f^+(x))$ und $(x, f^-(x))$ aus $D[H, r(a, \cdot)]$:

Nach dem Satz von Carathéodory [3] gibt es k Punkte $(x_1, z_1),..., (x_k, z_k)$ in $D[H, r(a, \cdot)]$ mit

$$0 = \sum_{i=1}^{k} \alpha_i(z_i - r(a, x_i)) \operatorname{grad} r(a, x_i).$$

Dabei ist $k \leqslant n + m + 1$, $\alpha_i > 0$ für $i = 0,..., k$ und $\sum_{i=1}^{k} \alpha_i = 1$. Da $\operatorname{sgn}(z_i - r(a, x_i)) = \operatorname{sgn}(f(x_i) - r(a, x_i)$ ist, gibt es Zahlen $\beta_i > 0$ mit

$$\sum_{i=1}^{k} \alpha_i\beta_i(f(x_i) - r(a, x_i)) \operatorname{grad} r(a, x_i) = 0.$$

Setzen wir $\gamma_j = \alpha_j\beta_j/\sum_{i=1}^{k} \alpha_i\beta_i$ für $j = 1,..., k$, dann ist $\gamma_j > 0$ und $\sum_{j=1}^{k} \gamma_j = 1$.

Außerdem ergibt sich:

$$0 = \sum_{i=1}^{k} \gamma_i(f(x_i) - r(a, x_i)) \operatorname{grad} r(a, x_i),$$

also $0 \in \operatorname{conv}(S')$.

(b) $0 \in \operatorname{conv}(S')$.

(1) $f(x) - r(a, x) = 0$ für ein $x \in \mathfrak{M}(a)$:

Dann ist $r(a, \cdot)$ Minimallösung zu (f^+, f^-).

(2) $f(x) - r(a, x) \neq 0$ für alle $x \in \mathfrak{M}(a)$:

Nach dem Satz von Carathéodory existieren wieder $k \leqslant n + m + 1$ Punkte $x_1, x_2, \ldots, x_k$ in $\mathfrak{M}(a)$ mit $0 = \sum_{i=1}^{k} \alpha_i(f(x_i) - r(a, x_i)) \operatorname{grad} r(a, x_i)$, $\alpha_i > 0$ für $i = 1, \ldots, k$ und $\sum_{i=1}^{k} \alpha_i = 1$. Nun gibt es Zahlen $\beta_i > 0$ und $z_i \in \mathbb{R}$ $(i = 1, \ldots, k)$ mit $z_i - r(a, x_i) = \beta_i(f(x_i) - r(a, x_i))$ und $(x_i, z_i) \in D[H, r(a, \cdot)]$. Also gilt

$$0 = \sum_{i=1}^{k} \alpha_i \frac{(z_i - r(a, x_i))}{\beta_i} \operatorname{grad} r(a, x_i).$$

Setzen wir $\gamma_j := \alpha_j/\beta_j \sum_{i=1}^{k} \alpha_i/\beta_i$, dann ist $\gamma_j > 0$ und $\sum_{j=1}^{k} \gamma_j = 1$. Wieder erhalten wir

$$0 = \sum_{i=1}^{k} \gamma_i(z_i - r(a, x_i)) \operatorname{grad} r(a, x_i),$$

also $0 \in \operatorname{conv}(S'')$. Nach Satz 3.12 in [1] ist dann $r(a, \cdot)$ Minimallösung zu H und damit auch zu (f^+, f^-).

Bemerkung. Gilt für ein $x \in \mathfrak{M}(a)$ die Beziehung $f(x) - r(a, x) = 0$, dann liegt $r(a, x)$ genau in der Mitte zwischen $f^+(x)$ und $f^-(x)$. Einen solchen Punkt nennt Dunham [9] einen "Straddlepunkt".

Wir wollen jetzt noch einige Bezeichnungen einführen.

DEFINITION 2.1. Ein n-dimensionaler Unterraum von $C(B)$, der $g_1, g_2, \ldots, g_n$ als Basis besitzt, heißt ein *Haarscher Unterraum*, wenn für jede Wahl von n verschiedenen Punkten $x_1, x_2, \ldots, x_n$ in B die Determinante der $(n \times n)$-Matrix $(f_j(x_i))$ verschieden von null ist.

Im folgenden sei B das reelle Intervall $[a, b]$. Ist M ein endlichdimensionaler Unterraum von $C[a, b]$, dann setzen wir $\eta(M) :=$ maximale Dimension eines Haarschen Unterraumes von M, $\nu(M) := 1 +$ größte Anzahl von Vorzeichenwechsel, die Elemente von M haben können.

Wir wollen für $B = [a, b]$ die Bestapproximation durch eine Alternante oder einen Straddlepunkt charakterisieren.

DEFINITION 2.2. $(f^+, f^-, r(a, \cdot))$ besitzt eine *Alternante der Länge* k, falls k Punkte $a \leqslant x_1 < x_2 < \cdots < x_k \leqslant b$ existieren mit

(1) $x_i \in \mathfrak{M}(a)$ für $i = 1, 2, \ldots, k$

(2) $\operatorname{sgn}(f(x_i) - r(a, x_i)) = (-1)^i \epsilon$ für $1, 2, \ldots, k$.

Dabei ist $\epsilon = +1$ oder $\epsilon = -1$ und f wie in (1.4).

Wir brauchen weiterhin einen Hilfssatz von Goldstein, der die Theorie der konvexen Mengen mit der Alternantenbedingung verknüpft.

HILFSSATZ 2.1. *Ist $g_1, g_2, \ldots, g_n$ eine Basis eines Haarschen Unterraums in $C[a, b]$ und liegt 0 in der konvexen Hülle der Vektoren $\lambda_0 \mathfrak{G}(x_0), \ldots, \lambda_n \mathfrak{G}(x_n)$, wobei alle $\lambda_i \neq 0$, $\mathfrak{G}(x) = (g_1(x), \ldots, g_n(x))$ und $a \leqslant x_0 < x_1 < \cdots < x_n \leqslant b$ sind, dann alternieren die λ_i im Vorzeichen.*

Beweis. Vgl. Cheney und Loeb [5].

SATZ 2.4. *Wenn $(f^+, f^-, r(a, \cdot))$ eine Alternante der Länge*

$$1 + \nu(P + r(a, \cdot)\, Q)$$

oder einen Straddlepunkt besitzt, dann ist $r(a, \cdot)$ Mimimallösung zu $(f^+, f^-) \in OU[a, b]$.

Falls $r(a, \cdot)$ Minimallösung ist zu $(f^+, f^-) \in OU[a, b]$, dann besitzt $(f^+, f^-, r(a, \cdot))$ eine Alternante der Länge $1 + \eta(P + r(a, \cdot)\, Q)$ oder einen Straddlepunkt.

Beweis. (*a*) Falls ein Straddlepunkt vorliegt, ist $r(a, \cdot)$ natürlich Minimallösung. Wir wollen also jetzt voraussetzen, $(f^+, f^-, r(a, \cdot))$ besitze eine Alternante der Länge $1 + \nu(P + r(a, \cdot)\, Q)$ und es liege kein Straddlepunkt vor: Wenn wir annehmen, $r(a, \cdot)$ sei keine Minimallösung, dann gibt es nach Satz 2.1 ein $r(b, \cdot) \in R$ mit $(f(x) - r(a, x))(r(b, x) - r(a, x)) > 0$ für alle $x \in \mathfrak{M}(a)$. Da nun $f(x) - r(a, x)$ an $\nu + 1$ aufeinanderfolgenden Stellen in $[a, b]$ abwechselndes Vorzeichen besitzt, gilt dies auch für $r(b, x) - r(a, x)$. Dann muß aber $r(b, x) - r(a, x)$ mindestens ν Vorzeichenwechsel in $[a, b]$ haben. Dies ist aber ein Widerspruch zur Definition von ν.

(*b*) $r(a, \cdot)$ sei Minimallösung zu $(f^+, f^-) \in OU[a, b]$: Nach Definition von η existiert ein Haarscher Unterraum von $P + r(a, \cdot)\, Q$ der Dimension $\eta(P + r(a, \cdot)\, Q)$. Ist $g_1, g_2, \ldots, g_\eta$ eine Basis dieses Unterraums, dann setzen wir $\mathfrak{G}(a, x) = (g_1(x), \ldots, g_\eta(x))$. Nach Satz 2.3 liegt 0 in der konvexen Hülle von$\{(f(x) - r(a, x)) \operatorname{grad} r(a, x) \mid x \in \mathfrak{M}(a)\}$. Da $q(a, x) > 0$ in $[a, b]$ ist, folgt aus (2.2) und dem Satz von Carathéodory:

$$0 \in \operatorname{conv}(\{f(x) - r(a, x)) \mathfrak{G}(a, x) \mid x \in \mathfrak{M}(a)\})$$

Wir machen nun zwei Fallunterscheidungen:

(1) Es gibt ein $x \in \mathfrak{M}(a)$ mit $f(x) - r(a, x) = 0$:

Dann ist x Straddlepunkt und die Behauptung bewiesen.

(2) $f(x) - r(a, x) \neq 0$ für alle $x \in \mathfrak{M}(a)$:

Nach dem Satz von Carathéodory existieren $k + 1$ Punkte $a \leqslant x_0 < x_1 < \cdots < x_k \leqslant b$ in $\mathfrak{M}(a)$ mit

(α) $k \leqslant \eta$

(β) $0 = \sum_{i=0}^{k} \alpha_i (f(x_i) - r(a, x_i)) \mathfrak{G}(a, x_i)$

und $\alpha_i > 0$ für $i = 0, 1, \ldots, k$, $\sum_{i=1}^{k} \alpha_i = 1$.

Da $g_1, \ldots, g_\eta$ einen Haarschen Unterraum aufspannen, muß nun $k \geqslant \eta$ sein. Also ist $k = \eta$ und nach Hilfssatz 2.1 haben die Zahlen $f(x_i) - r(a, x_i)$ alternierendes Vorzeichen, d. h. es gibt eine Alternante der Länge $1 + \eta(P + r(a, \cdot) Q)$.

Bemerkung. Falls bei einer Minimallösung ein Straddlepunkt vorliegt, kann dennoch eine Alternante existieren.

Betrachten wir als Spezialfall für R die Menge $R_m{}^n[a, b]$, d. h. die Menge der rationalen Funktionen p/q, wobei p ein Polynom $\leqslant n$-ten Grades und q ein Polynom $\leqslant m$-ten Grades ist mit $q(x) > 0$ in $[a, b]$, dann ist $P + rQ$ für jedes $r \in R_m{}^n[a, b]$ ein Haarscher Unterraum. In diesem Fall sind aber auch die Zahlen ν und η gleich und es gilt, falls p und q teilerfremd sind:

$$\nu(P + (p/q) Q) = \eta(P + (p/q) Q) = 1 + \max\{n + \partial q, m + \partial p\}$$

(vgl. Cheney [2]). Dabei bezeichnen wir mit ∂p bzw. ∂q den Grad der Polynome p bzw. q.

Wir erhalten also aus Satz 2.4:

FOLGERUNG 2.1. *$r_0 = p/q$ (p und q teilerfremd) ist genau dann eine Minimallösung zu $(f^+, f^-) \in OU[a, b]$ bezüglich $R_m{}^n[a, b]$, wenn (f^+, f^-, r_0) eine Alternante der Länge* $2 + \max\{n + \partial q, m + \partial p\}$ *besitzt oder ein Straddlepunkt vorliegt.*

3. Eindeutigkeit bei Rationaler Simultanapproximation

Als Spezialisierung von Satz 3.16 aus [1] auf rationale Simultanapproximation erhalten wir:

Satz 3.1. *Ist* $r_0 \in R$ *eine Minimallösung zu* $(f^+, f^-) \in OU(B)$ *und ist* $P + r_0 Q$ *ein Haarscher Unterraum, dann ist* r_0 *eindeutig bestimmt, falls kein Straddlepunkt existiert.*

Wir können diesen Satz etwas verschärfen, falls $B = [a, b]$ ist.

Satz 3.2. *Ist* $r_0 \in R$ *eine Minimallösung zu* $(f^+, f^-) \in OU[a, b]$ *und* $P + r_0 Q$ *ein Haarscher Unterraum, dann ist* r_0 *eindeutig bestimmt, falls eine Alternante der Länge* $1 + \eta(P + r_0 Q)$ *existiert.*

Den Beweis wollen wir unter Anwendung des folgenden Hilfssatzes führen, den wir später noch einmal benötigen.

Hilfssatz 3.1. *Sei* $r(a, \cdot) \in R$ *eine Minimallösung zu* $(f^+, f^-) \in OU[a, b]$, *die eine Alternante der Länge* $1 + \eta(P + r(a, \cdot)\, Q)$ *besitzt, d. h. es existieren Punkte*

$$a \leqslant x_0 < x_1 < \cdots < x_\eta \leqslant b$$

mit

(1) $x_i \in \mathfrak{M}(a)$ *für* $i = 0,..., \eta$,

(2) $\operatorname{sgn}(f(x_i) - r(a, x_i)) = (-1)^i \epsilon$ *für* $i = 0,..., \eta$

($\epsilon = +1$ *oder* $\epsilon = -1$). *Weiterhin sei* $\rho_R(f^+, f^-) > 0$ *und* $P + r(a, \cdot)\, Q$ *ein Haarscher Unterraum. Dann ist* 0 *die einzige Funktion g aus* $P + r(a, \cdot)\, Q$ *mit*

$$(-1)^i \epsilon \cdot g(x_i) \geqslant 0 \qquad \textit{für} \quad i = 0,..., \eta.$$

Beweis. Wir setzen

$$f_1^+(x_i) := f^+(x_i), \qquad \text{falls} \quad (-1)^i \epsilon = 1$$

und

$$f_1^-(x_i) := f^-(x_i), \qquad \text{falls} \quad (-1)^i \epsilon = -1.$$

Auf allen übrigen Punkten von $[a, b]$ definieren wir $f_1^+(x)$ und $f_1^-(x)$ durch $r(a, x)$. Es ist leicht zu verifizieren, daß f_1^+ eine oberhalbstetige und f_1^- eine unterhalbstetige Funktion ist mit $f_1^+(x) \geqslant f_1^-(x)$ für alle $x \in [a, b]$. Nach Konstruktion bilden die Punkte $x_0, x_1,..., x_\eta$ eine Alternante von $(f_1^+, f_1^-, r(a, \cdot))$. Da $P + r(a, \cdot)\, Q$ ein Haarscher Unterraum ist, gilt

$\eta(P + r(a, \cdot)\, Q) = \nu(P + r(a, \cdot)\, Q)$. Deshalb ist nach Satz 2.4 $r(a, \cdot)$ Minimallösung zu (f_1^+, f_1^-). Nach Satz 2.3 ist dann der Nullvektor in der konvexen Hülle der Menge

$$\{(f_1(x_i) - r(a, x_i)) \operatorname{grad} r(a, x_i) \mid i = 0, 1,..., \eta\}$$

enthalten. Da $f_1(x_i) - r(a, x_i) \neq 0$ für $i = 0, 1,..., \eta$ und $\operatorname{sgn}(f_1(x_i) - r(a, x_i)) = \operatorname{sgn}(f(x_i) - r(a, x_i))$, ist 0 in der konvexen Hülle von $\{(-1)^i \epsilon \operatorname{grad} r(a, x_i) \mid i = 0, 1,..., \eta\}$ enthalten, also

$$0 = \sum_{i=0}^{\eta} \alpha_i (-1)^i \epsilon \operatorname{grad} r(a, x_i) \qquad \text{mit}$$
$$\alpha_i \geqslant 0 \qquad \text{für} \quad i = 0, 1,..., \eta \qquad \text{und} \qquad \sum_{i=0}^{\eta} \alpha_i = 1. \tag{3.1}$$

Wegen der Haarschen Bedingung müssen alle $\alpha_i > 0$ sein. Aus (3.1) folgt für jedes $g \in P + r(a, \cdot)\, Q$ mit $g \neq 0$:

$$0 = \sum_{i=0}^{\eta} \alpha_i (-1)^i \epsilon g(x_i) \frac{1}{q(a, x_i)}$$

Nun können wegen der Haarschen Bedingung höchstens $\eta - 1$ der Zahlen $g(x_i)$ verschwinden. Deshalb ist mindestens eine der Zahlen $(-1)^i \epsilon g(x_i)$ positiv und eine negativ.

Beweis von Satz 3.2. (*a*) $\rho_R(f^+, f^-) = 0$: Dann ist $r_0 = f^+ = f^-$.

(*b*) $\rho_R(f^+, f^-) > 0$: Ist $r = p/q$ eine weitere Minimallösung zu (f^+, f^-), dann gehört die Funktion $q(r - r_0)$ zu $P + r_0 Q$ und es gilt:

$$q(r - r_0) = q[(f - r_0) - (f - r)].$$

Also gilt für alle Alternantenpunkte $x_0, ..., x_\eta$ von (f^+, f^-, r_0) wegen

$$|(f - r_0)(x_i)| \geqslant |(f - r)(x_i)|$$

die Beziehung

$$(-1)^i \epsilon q(x_i)(r - r_0)(x_i) = (-1)^i \epsilon q(x_i)[(f - r_0)(x_i) - (f - r)(x_i)] \geqslant 0.$$

Nach dem vorhergehenden Hilfssatz gilt dann:

$$r = r_0 .$$

Bevor wir ein Analogon zum strengen Eindeutigkeitssatz von Cheney [2, 3] angeben, wollen wir einen Hilfssatz von Cheney [3] anführen.

HILFSSATZ 3.2. *Sei* $r_0 = p_0/q_0$ *ein Element von* R *(*R *über* $[a, b]$ *betrachtet) mit* $\dim(P + r_0 Q) = \dim(P) + \dim(Q) - 1$.

Wenn nun für $p \in P$ *und* $q \in Q$ *gilt*

(1) $\| p \| + \| q \| = \| p_0 \| + \| q_0 \|$,

(2) $p = r_0 q$,

(3) $q(x) \geqslant 0$ *in* $[a, b]$,

dann ist $p = p_0$ *und* $q = q_0$.

Beweis. Vgl. Cheney [3, S. 165].

SATZ 3.3 (Strenger Eindeutigkeitssatz). *Sei* $r_0 = p_0/q_0 \in R$ *eine Minimallösung zu* $(f^+, f^-) \in OU[a, b]$ *mit einer Alternante der Länge* $1 + \eta(P + r_0 Q)$. *Wenn jetzt* $\eta(P + r_0 Q) = \dim(P) + \dim(Q) - 1$ *ist, dann existiert eine Konstante* $\gamma > 0$, *so daß für alle* $r \in R$ *gilt*:

$$\Delta(r) \geqslant \Delta(r_0) + \gamma \| r - r_0 \|.$$

Beweis. Wir machen zwei Fallunterscheidungen.

(*a*) $\rho_R(f^+, f^-) = 0$: Dann gilt

$$\begin{aligned} \Delta(r) = \|(f^+ - r, f^- - r)\| &\geqslant \|(r - r_0, r - r_0)\| - \|(f^+ - r_0, f^- - r_0)\| \\ &= \| r - r_0 \|. \end{aligned}$$

Also ist $\gamma = 1$.

(*b*) $\rho_R(f^+, f^-) > 0$: Zu jedem $r \in R$ mit $r \neq r_0$ definieren wir:

$$\gamma(r) := (\Delta(r) - \Delta(r_0))/\| r - r_0 \|. \tag{3.2}$$

Wir müssen nun zeigen, daß 0 kein Häufungspunkt von $\gamma(r)$ ist. Dazu nehmen wir indirekt an, es gebe eine Folge $\{r_k\}$ in R, so daß

(1) $r_k \neq r_0$ für $k = 1, 2, \ldots$,

(2) $\gamma(r_k) \to 0$ für $k \to \infty$.

Weiter sei $r_k = p_k/q_k$ mit $p_k \in P$, $q_k \in Q$ und $q_k(x) > 0$ in $[a, b]$. Wir können annehmen, daß

$$\| p_k \| + \| q_k \| = 1 \qquad \text{für} \quad k = 0, 1, 2, \ldots.$$

Setzen wir

$$p_k = \sum_{\nu=1}^{n} \alpha_\nu^{(k)} u_\nu, \qquad q_k = \sum_{\nu=1}^{m} \beta_\nu^{(k)} u_{n+\nu}$$

$$\mu_1 = \min_{\substack{\alpha_\nu \\ \Sigma_1^n |\alpha_\nu| = 1}} \left\| \sum_{\nu=1}^{n} \alpha_\nu u_\nu \right\|, \qquad \mu_2 = \min_{\substack{\beta_\nu \\ \Sigma_1^m |\beta_\nu| = 1}} \left\| \sum_{\nu=1}^{m} \beta_\nu u_{n+\nu} \right\|,$$

dann ist $\mu_1 > 0$ und $\mu_2 > 0$ und es gilt:

$$|\alpha_\nu^{(k)}| \leqslant (1/\mu_1)\|p_k\| \leqslant 1/\mu_1 \qquad \text{für} \quad \nu = 1, 2, \ldots, n \quad \text{und}$$

$$|\beta_\nu^{(k)}| \leqslant (1/\mu_2)\|q_k\| \leqslant 1/\mu_2 \qquad \text{für} \quad \nu = 1, 2, \ldots, m$$

Da also $|\alpha_\nu^{(k)}|$ für $\nu = 1, \ldots, n$ und $k = 1, 2, \ldots,$ und $|\beta_\nu^{(k)}|$ für $\nu = 1, \ldots, m$ und $k = 1, 2, \ldots,$ beschränkt sind, existiert also eine konvergente Teilfolge von $\{p_k\}$ und eine von $\{q_k\}$.

Wir nehmen o.B.d.A. an:

$$p_k \to p \in P \qquad \text{für} \quad k \to \infty$$

$$q_k \to q \in Q \qquad \text{für} \quad k \to \infty.$$

Sei nun $\gamma(r_k) \leqslant \frac{1}{2}$, dann gilt wegen (3.2):

$$\begin{aligned} \| r_k - r_0 \| &= \|(r_k - r_0, r_k - r_0)\| \\ &\leqslant \|(f^+ - r_0, f^- - r_0)\| + \|(f^+ - r_k, f^- - r_k)\| \\ &\leqslant 2\|(f^+ - r_0, f^- - r_0)\| + \tfrac{1}{2}\| r_k - r_0 \| \end{aligned}$$

oder

$$\| r_k - r_0 \| \leqslant 4\|(f^+ - r_0, f^- - r_0)\|.$$

Also ist $\| r_k - r_0 \|$ hiermit beschränkt für $k = 1, 2, \ldots.$ Wir wollen jetzt zeigen, daß $p = r_0 q$ ist: Sei $a \leqslant x_0 < x_1 < \cdots < x_\eta \leqslant b$ Alternante zu (f^+, f^-, r_0), also $x_i \in \mathfrak{M}(r_0)$ und $\operatorname{sgn}(f(x_i) - r_0(x_i)) = (-1)^i \epsilon$ für $i = 0, 1, \ldots, \eta$ ($\epsilon = +1$ oder $\epsilon = -1$). Dann gilt für $i = 0, 1, \ldots, \eta$:

$$\begin{aligned} \gamma(r_k)\| r_k - r_0 \| &= \Delta(r_k) - \Delta(r_0), \\ &\geqslant (-1)^i \epsilon (f - r_k)(x_i) + Z(x_i) - (-1)^i \epsilon (f - r_0)(x_i) - Z(x_i) \\ &= (-1)^i \epsilon (r_0 - r_k)(x_i) \\ &= (-1)^i \epsilon \frac{r_0(x_i)\, q_k(x_i) - p_k(x_i)}{q_k(x_i)}. \end{aligned} \tag{3.3}$$

Durch Grenzübergang $k \to \infty$ folgt für $i = 0, 1, \ldots, \eta$:

$$(-1)^i \epsilon (p - r_0 q)(x_i) \geqslant 0.$$

Aus Hilfssatz 3.1 folgt nun $p = r_0 q$ und aus Hilfssatz 3.2 $p = p_0$ und $q = q_0$.

Da $q_0(x) > 0$ für $x \in [a, b]$, existiert ein $\delta > 0$ mit $q_0(x) \geqslant 2\delta$ in $[a, b]$.

Wir können also *o*.B.d.A. annehmen, daß $q_k(x) \geqslant \delta$ ist in $[a, b]$ für $k = 1, 2, \ldots$. Nun sei

$$c := \inf_{\substack{g \in P + r_0 Q \\ \|g\| = 1}} \max_{0 \leqslant i \leqslant \eta} (-1)^i \, \epsilon g(x_i).$$

Da das Infimum angenommen wird, ist nach Hilfssatz 3.1 $c > 0$. Aus der Definition von c und Ungleichung (3.3) folgt, daß ein Alternantenpunkt $x_j (j = j(k))$ existiert mit:

$$\begin{aligned} \gamma(r_k) \| r_0 - r_k \| &\geqslant (-1)^j \, \epsilon (r_0 q_k - p_k)(x_j)/q_k(x_j) \\ &\geqslant (-1)^j \, \epsilon (r_0 q_k - p_k)(x_j) \\ &\geqslant c \cdot \| r_0 q_k - p_k \| \\ &\geqslant c\delta \, \| r_0 - r_k \|. \end{aligned}$$

Da $r_k \neq r_0$ ist für alle k und $\gamma(r_k) \to 0$ für $k \to \infty$, führt dies zu einem Widerspruch.

FOLGERUNG 3.1. *Ist $r_0 = p_0/q_0 \in R_m^n[a, b]$ Minimallösung zu $(f^+, f^-) \in OU[a, b]$ mit einer Alternante der Länge $n + m + 2$, p_0 und q_0 teilerfremd und* $\min(n - \partial p_0, m - \partial q_0) = 0$, *dann gibt es ein $\gamma > 0$, so daß für alle $r \in R_m^n[a, b]$ gilt:*

$$\Delta(r) \geqslant \Delta(r_0) + \gamma \cdot \| r - r_0 \|.$$

4. STETIGKEITSEIGENSCHAFTEN

Wie in der Theorie der Tschebyscheffapproximation einer einzigen Funktion führen wir nun einen Bestapproximationsoperator T ein, der dem Paar $(f^+, f^-) \in OU[a, b]$ die Menge aller Bestapproximationen aus R zuordnet:

$$T(f^+, f^-) := \{ r \in R \mid \Delta(r) = \rho_R(f^+, f^-) \}.$$

Diese Menge kann leer sein. Im Falle $R = R_m^n[a, b]$ ist $T(f^+, f^-)$ nicht leer, wie man durch Übertragung des Existenzsatzes für die Approximation einer Funktion erkennt. Jedoch kann es auch in diesem Fall mehrere Minimallösungen geben. Ein einfaches Beispiel dafür ist folgendes:

In $[0, 1]$ sollen die Funktion $f^+(x) = 1 - x$ und $f^-(x) \equiv 0$ durch Geraden approximiert werden. Dann ist $\{1/2 - cx \mid 0 \leqslant c \leqslant 1\}$ die Menge der Minimallösungen (Abb. 1).

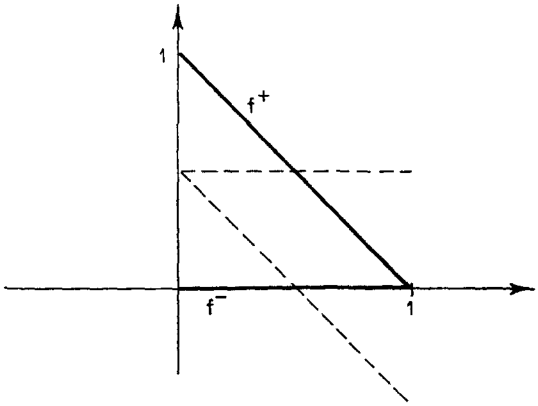

Abb. 1

Maehly und Witzgall [11] haben als erste eine hinreichende Bedingung für die Stetigkeit von T bei der Approximation einer Funktion durch Elemente aus $R_m{}^n[a, b]$ angegeben. Werner [14] zeigte, daß sie auch notwendig ist. Später hat Cheney [2] bestätigt, daß diese Bedingung auch für verallgemeinerte rationale Approximation hinreichend ist. Zusammen mit Loeb [6] zeigte er ihre Notwendigkeit.

Im folgenden wollen wir diese Bedingung verallgemeinern auf die Simultanapproximation und führen dazu einen verallgemeinerten Begriff der Normalität ein.

DEFINITION 4.1. $(f^+, f^-) \in OU[a,b]$ heißt normal, wenn es ein $r_0 \in T(f^+, f^-)$ gibt mit den Eigenschaften:

(1) $\eta(P + r_0 Q) = \dim(P) + \dim(Q) - 1$

(2) r_0 wird durch eine Alternante charakterisiert.

Aus dem Eindeutigkeitssatz 3.2 folgt sofort

SATZ 4.1. *Ist (f^+, f^-) normal, dann besteht $T(f^+, f^-)$ aus genau einem Element.*

Falls $f^+ = f^-$ ist, dann ist diese Definition der Normalität mit derjenigen bei einer einzigen zu approximierenden Funktion identisch. In diesem Fall ist die Eigenschaft (2) immer gesichert.

SATZ 4.2. *Ist $f^+ = f^- \in R$ oder (f^+, f^-) normal, dann gibt es eine Umgebung U von (f^+, f^-) in $OU[a, b]$, so daß $T(g^+, g^-) \neq \varnothing$ für alle $(g^+, g^-) \in U$.*

Es existiert weiterhin eine Zahl $\beta > 0$, so daß für $r_0 \in T(f^+, f^-)$ und $r_g \in T(g^+, g^-)$ gilt:

$$\| r_0 - r_g \| \leqslant \beta \|(f^+, f^-) - (g^+, g^-)\|.$$

Beweis. Wir wollen zuerst den zweiten Teil des Satzes beweisen. Dazu machen wir zwei Fallunterscheidungen.

(a) $f^+ = f^- \in R$: Dann ist $r_0 = f^+ = f^-$ und es gilt:

$$\begin{aligned} \| r_0 - r_g \| &\leqslant \|(r_g, r_g) - (g^+, g^-)\| + \|(g^+, g^-) - (f^+, f^-)\| \\ &\leqslant 2 \|(f^+, f^-) - (g^+, g^-)\|. \end{aligned}$$

Wir können also $\beta = 2$ wählen.

(b) (f^+, f^-) sei normal: Die Suche nach einer Minimallösung zu (g^+, g^-) kann auf solche $r \in R$ beschränkt werden mit

$$\|(g^+, g^-) - (r, r)\| \leqslant \|(g^+, g^-) - (r_0, r_0)\|.$$

Wegen des strengen Eindeutigkeitssatzes gilt für solche r:

$$\begin{aligned} \gamma \| r - r_0 \| &\leqslant \| (f^+, f^-) - (r, r)\| - \| (f^+, f^-) - (r_0, r_0)\| \\ &\leqslant \|(f^+, f^-) - (g^+, g^-)\| + \|(g^+, g^-) - (r, r)\| \\ &\quad - \|(f^+, f^-) - (r_0, r_0)\| \\ &\leqslant \|(f^+, f^-) - (g^+, g^-)\| + \|(g^+, g^-) - (r_0, r_0)\| \\ &\quad - \|(f^+, f^-) - (r_0, r_0)\| \\ &\leqslant 2 \|(f^+, f^-) - (g^+, g^-)\|. \end{aligned}$$

Also gilt für $r_g \in T(g^+, g^-)$:

$$\| r_g - r_0 \| \leqslant 2\gamma^{-1} \|(f^+, f^-) - (gf, g^-)\|.$$

Damit können wir $\beta := 2\gamma^{-1}$ setzen.

Um den ersten Teil des Satzes zu beweisen, sei $r_0 = p_0/q_0$ und $\| p_0 \| + \| q_0 \| = 1$. Die Zahl $\epsilon_1 := (1/2) \min_{x \in [a,b]} q_0(x)$ ist größer als null. Wir wählen nun $\epsilon_2 > 0$ so, daß aus

$$\left.\begin{aligned} &r = p/q \in R, \\ &\| p \| + \| q \| = 1, \\ &\| r - r_0 \| < \epsilon_2, \end{aligned}\right\} \Rightarrow \| q - q_0 \| < \epsilon_1 .$$

Eine solche Zahl ϵ_2 existiert, denn nehmen wir an, es gebe eine Folge $\{r_k\} = \{p_k/q_k\}$ in R mit

$$\| q_k \| + \| p_k \| = 1, \qquad \| r_k - r_0 \| \xrightarrow{k \to \infty} 0 \qquad \text{und} \qquad \| q_k - q_0 \| \geqslant \epsilon_1 ,$$

dann folgt o.B.d.A. aus dem Satz von Bolzano–Weierstraß (wie im Beweis von Satz 3.3): es existieren $q \in Q$ und $p \in P$ mit

(1) $q(x) \geqslant 0$ in $[a, b]$

(2) $\| p - p_k \| \to 0$ für $k \to \infty$

(3) $\| q - q_k \| \to 0$ für $k \to \infty$.

Aus $\| r_k - r_0 \| \to 0$ für $k \to \infty$ folgt dann wegen

$$\| p_k - r_0 q_k \| = \| q_k(r_k - r_0) \| \leqslant \| r_k - r_0 \|$$

auch $\| p_k - r_0 q_k \| \to 0$ für $k \to \infty$.

Damit gilt wegen

$$\| p - r_0 q \| \leqslant \| p - p_k \| + \| p_k - r_0 q_k \| + \| r_0(q_k - q) \|$$

die Beziehung $p = r_0 q$. Nach Hilfssatz 3.2 ist dann $q = q_0$ und $p = p_0$, im Widerspruch zu $\| q - q_0 \| \geqslant \epsilon_1$.

Um den Beweis zu vollenden, sei nun

$$U := \left\{ (g^+, g^-) \in OU[a, b] \mid \| (f^+, f^-) - (g^+, g^-) \| < \frac{\epsilon_2}{\beta} \right\}$$

Dann erfüllt eine Minimallösung r_g zu $(g^+, g^-) \in U$, falls sie existiert, wegen (a) und (b) die Bedingung

$$\| r_g - r_0 \| < \epsilon_2 .$$

Wenn wir nun $r_g = p/q$ durch $\| p \| + \| q \| = 1$ normieren, folgt also:

$$\| q - q_0 \| < \epsilon_1 .$$

Deshalb können wir uns bei der Suche nach einer Bestapproximation zu $(g^+, g^-) \in U$ auf

$$\left\{ r = \frac{p}{q} \in R \mid \| p \| + \| q \| = 1, q(x) \geqslant \epsilon_1 \text{ in } [a, b] \right\}$$

beschränken. Diese Menge ist kompakt in $C[a, b]$ und enthält damit eine Minimallösung zu $(g^+, g^-) \in U$.

Es ist noch offen, ob die Bedingung in diesem Satz auch notwendig für die Stetigkeit des Bestapproximationsoperators ist.

5. Eine Verallgemeinerung eines Verfahrens von Cheney–Loeb

Cheney und Loeb [4] leiteten ein Verfahren zur Lösung der rationalen Tschebyscheffapproximation einer einzigen Funktion her, das nicht die Existenz einer Alternante wie der Remez–Algorithmus benutzt. Da bei der

Simultanapproximation zweier Funktionen f^+ und f^- über dem Intervall $[a, b]$ nicht in jedem Fall eine Alternante existiert, erscheint es sinnvoll zu versuchen, ein Verfahren zu verallgemeinern, das nicht auf einer Alternante beruht.

Es sei wieder $(f^+, f^-) \in OU(B)$ und $P, Q,$ und R wie in 2. Wir wollen also $\Delta(a) = \max_{x\in B} \{|f(x) - r(a, x)| + Z(x)\}$ bezüglich R minimieren (f und Z wie in 1.4 bzw. 1.5). Bei der Suche nach einer Minimallösung können wir uns auf solche $a \in \mathbb{R}^{n+m}$ mit $\| a \| \leqslant 1$ beschränken. Dabei soll

$$\| a \| := \max_{1\leqslant i\leqslant n+m} | a_i | \quad \text{sein.}$$

Bei dem folgenden Iterationsverfahren sei der Ausgangsvektor a^1 mit $\| a^1 \| \leqslant 1$ so gewählt, daß

$$q(a^1, x) > 0 \qquad \text{für alle } x \in B.$$

Wir nehmen nun an, wir haben a^k mit $\| a^k \| \leqslant 1$ und $q(a^k, x) > 0$ für alle $x \in B$ bestimmt. Dann definieren wir für $a \in \mathbb{R}^{n+m}$ die Hilfsfunktion

$$\delta_k(a) := \sup_{x\in B} \{| f(x)\, q(a, x) - p(a, x)| + (Z(x) - \Delta(a^k))\, q(a, x)\}. \qquad (5.1)$$

$\delta_k(a)$ ist eine in $\mathbb{R}^{n+m}$ konvexe, also auch stetige Funktion (vgl. Collatz-Wetterling [7, S. 77]).
Wir bestimmen nun einen Vektor $b^{k+1} \in \mathbb{R}^{n+m}$ so, daß er die Funktion $\delta_k(a)$ im Einheitswürfel $\| a \| \leqslant 1$ minimiert. Dies ist ein Problem der konvexen Optimierung. Wir behaupten:

(α) Ist $q(b^{k+1}, x_0) = 0$ für ein $x_0 \in B$ oder $\delta_k(b^{k+1}) = 0$, dann ist $(a^k, \cdot)$ Minimallösung zu $(f^+, f^-) \in OU(B)$.

Beweis. Ist $q(b^{k+1}, x_0) = 0$, so folgt aus (5.1):

$$\delta_k(b^{k+1}) \geqslant | p(b^{k+1}, x_0)| \geqslant 0.$$

Wegen

$$\delta_k(a^k) = \sup_{x\in B} \{q(a^k, x)[| f(x) - r(a^k, x)| + Z(x) - \Delta(a^k)]\} = 0$$

ist $\delta_k(b^{k+1}) = 0$, und wir brauchen nur noch für diesen Fall zu zeigen, daß $r(a^k, \cdot)$ Minimallösung ist: Nehmen wir indirekt an, es gebe eine bessere Approximation als $r(a^k, \cdot)$, z.B. $r(b, \cdot)$. Dann ist $\Delta(b) = \Delta(a^k) - \epsilon_1$ mit $\epsilon_1 > 0$. Weiterhin gilt:

$$\begin{aligned}\delta_k(b) &= \sup_{x\in B} \{q(b, x)[|f(x) - r(b, x)| + Z(x) - \Delta(a^k)]\} \\ &= \sup_{x\in B} \{q(b, x)[| f(x) - r(b, x)| + Z(x) - \Delta(b) - \epsilon_1]\}.\end{aligned}$$

Da es ein $\epsilon_2 > 0$ gibt mit $q(b, x) \geqslant \epsilon_2$ für alle $x \in B$, folgt:

$$\delta_k(b) \leqslant -\epsilon_1\epsilon_2 < 0.$$

Dies ist aber ein Widerspruch zu $\delta_k(b^{k+1}) = 0$.

Liegt also die Situation (α) vor, dann setzen wir

$$a^{k+j} := a^k \qquad \text{für} \quad j = 1, 2, \ldots,$$

und das Verfahren ist beendet. Andernfalls definieren wir

$$a^{k+1} := b^{k+1}$$

und wiederholen das Verfahren mit der Hilfsfunktion $\delta_{k+1}(a)$ usw.

Für dieses Verfahren gilt dann folgender Satz:

SATZ 5.1. *$\Delta(a^k)$ konvergiert monoton fallend gegen die Minimalabweichung $\rho_R(f^+, f^-)$. Die Konvergenz ist mindestens linear, d.h.*

$$\Delta(a^{k+1}) - \rho_R(f^+, f^-) \leqslant \gamma(\Delta(a^k) - \rho_R(f^+, f^-)) \qquad \textit{mit} \quad 0 \leqslant \gamma < 1,$$

falls eine Minimallösung existiert.

Beweis. Ist $q(a^{k+1}, x) > 0$ für alle $x \in B$ und $\delta_k(a^{k+1}) < 0$, dann gilt für alle $x \in B$:

$$|f(x) - r(a^{k+1}, x)| + Z(x) < \Delta(a^k)$$

und daher

$$\Delta(a^{k+1}) < \Delta(a^k).$$

Die Folge $\{\Delta(a^k)\}$ ist also monoton fallend und hat einen Grenzwert L. Wir müssen nun zeigen, daß $L = \rho_R(f^+, f^-)$: Dazu nehmen wir indirekt an, L sei nicht gleich $\rho_R(f^+, f^-)$. Dann existiert eine Funktion $r(a', \cdot) \in R$ mit $\Delta(a') < L$ und $\| a' \| \leqslant 1$. Es gilt nun:

$$\delta_k(a^{k+1}) \leqslant \delta_k(a') = \sup_{x \in B} \{q(a', x)[|f(x) - r(a', x)| + Z(x) - \Delta(a^k)]\}.$$

Wegen

$$|f(x) - r(a', x)| + Z(x) \leqslant \Delta(a') < \Delta(a^k)$$

folgt also

$$\begin{aligned} \delta_k(a^{k+1}) \leqslant \delta_k(a') &\leqslant \sup_{x \in B} \{q(a', x)[\Delta(a') - \Delta(a^k)]\} \\ &= \alpha[\Delta(a') - \Delta(a^k)], \end{aligned}$$

wobei

$$\alpha := \min_{x \in B} q(a', x) > 0.$$

Ebenso erhält man:

$$\begin{aligned} 0 \geqslant \delta_k(a^{k+1}) &\geqslant \beta \sup_{x \in B} \{|f(x) - r(a^{k+1}, x)| + Z(x) - \Delta(a^k)\} \\ &= \beta[\Delta(a^{k+1}) - \Delta(a^k)], \end{aligned}$$

wobei

$$\beta := \sup_{\|a\| \leqslant 1} \max_{x \in B} q(a, x) > 0.$$

Dann gilt:

$$\begin{aligned} \Delta(a^{k+1}) - \Delta(a^k) &\leqslant (1/\beta)\, \delta_k(a^{k+1}) \\ &\leqslant (\alpha/\beta)[\Delta(a') - \Delta(a^k)]. \end{aligned}$$

Für $k \to \infty$ folgt daraus:

$$0 \leqslant (\alpha/\beta)[\Delta(a') - L].$$

Dies ist aber ein Widerspruch zu $\Delta(a') < L$. Also ist $L = \rho_R(f^+, f^-)$.

Ist $r(b, \cdot)$ mit $\| b \| \leqslant 1$ Minimallösung zu (f^+, f^-), dann folgt die lineare Konvergenz aus:

$$\begin{aligned} \Delta(a^{k+1}) - \Delta(a^k) &= \Delta(a^{k+1}) - \rho_R(f^+, f^-) + \rho_R(f^+, f^-) - \Delta(a^k) \\ &\leqslant (\alpha_1/\beta)[\rho_R(f^+, f^-) - \Delta(a^k)], \end{aligned}$$

wobei

$$\alpha_1 := \min_{x \in B} q(b, x).$$

Durch Umschreiben erhält man nämlich:

$$\Delta(a^{k+1}) - \rho_R(f^+, f^-) \leqslant (1 - \alpha_1/\beta)[\Delta(a^k) - \rho_R(f^+, f^-)]. \tag{5.2}$$

Dabei ist $0 \leqslant 1 - \alpha_1/\beta < 1$.

Bemerkung. Da die Vektoren a^k beschränkt sind, existiert eine Teilfolge, die gegen einen Vektor $b \in \mathbb{R}^{n+m}$ konvergiert. Bei rationaler Polynomapproximation über dem Intervall $[a, b]$ können wir zeigen, daß $r(b, \cdot)$ Minimallösung der vorgeschriebenen Form ist, wenn man zuvor gemeinsame Faktoren in Zähler und Nenner kürzt.

Falls $(f^+, f^-) \in OU[a, b]$ normal ist, dann kann man über die Konvergenz der Funktionen $r(a^k, \cdot)$ folgendes aussagen.

SATZ 5.2. *Ist* $(f^+, f^-) \in OU[a, b]$ *normal, dann konvergieren die Funktionen* $r(a^k, \cdot)$ *aus dem oben beschriebenen Verfahren linear gegen die Minimallösung* r_0 *von* (f^+, f^-):

$$\| r(a^k, \cdot) - r_0 \| \leqslant A\theta^{k-1} \qquad mit \quad 0 \leqslant \theta < 1.$$

Beweis. Aus dem strengen Eindeutigkeitssatz und aus (5.2) folgt:

$$\begin{aligned} \| r(a^k, \cdot) - r_0 \| &\leqslant \gamma^{-1}[\Delta(a^k) - \rho_R(f^+, f^-)] \\ &\leqslant \gamma^{-1} \left(1 - \frac{\alpha_1}{\beta}\right)^{k-1} [\Delta(a^1) - \rho_R(f^+, f^-)]. \end{aligned}$$

Mit $\theta := (1 - \alpha_1/\beta)$ und $A := \gamma^{-1}[\Delta(a^1) - \rho_R(f^+, f^-)]$ ergibt sich die Behauptung.

LITERATUR

1. H. P. BLATT, Nicht-lineare gleichmäßige Simultanapproximation, *J. Approximation Theory* **8** (1973).
2. E. W. CHENEY, Approximation by generalized rational functions, in "Approximation of functions." (H. L. Garabedian, Ed.) pp. 101–110, Elsevier, Amsterdam, 1965.
3. E. W. CHENEY, "Introduction to Approximation Theory," McGraw-Hill, New York, 1966.
4. E. W. CHENEY AND H. L. LOEB, On rational Chebyshev approximation, *Numer. Math.* **4** (1962), 124–127.
5. E. W. CHENEY AND H. L. LOEB, Generalized rational approximation, *J. Siam Numer. Anal. Ser. B* **1** (1964), 11–25.
6. E. W. CHENEY AND H. L. LOEB, On the continuity of rational approximation operators, *Arch. Rational Mech. Anal.* **21** (1966), 391–401.
7. L. COLLATZ AND W. WETTERLING, "Optimierungsaufgaben," Springer-Verlag, Berlin, 1966.
8. J. B. DIAZ AND H. W. MCLAUGHLIN, Simultaneous approximation of a set of bounded real functions, *Math. Comp.* **23** (1969), 583–593.
9. C. B. DUNHAM, Simultaneous Chebyshev approximation of functions on an interval, *Proc. Amer. Math. Soc.* **18** (1967), 472–477.
10. M. GOLOMB, On the uniformly best approximation of functions given by incomplete data, M.R.C. Technical Summary Report 121, Dec. 1959, The University of Wisconsin, Madison, Wisconsin.
11. H. J. MAEHLY AND CH. WITZGALL, Tschebyscheff-Approximationen in kleinen Intervallen. II, *Numer. Math.* **2** (1960), 293–307.
12. G. MEINARDUS AND D. SCHWEDT, Nicht-lineare Approximationen, *Arch. Rational Mech. Anal.* **17** (1964), 297–326.
13. E. RÉMÈS, Sur la détermination des polynomes d'approximation de degré donné, *Comm. de la Soc. Math. de Kharkof, Sér. 4* **10** (1934), 41–63.
14. H. WERNER, On the rational Tschebyscheff operator, *Math. Z.* **86** (1964), 317–326.