JOURNAL OF APPROXIMATION THEORY 1, 340–354 (1968)

Über eine Klasse rationaler Tschebyscheff-Approximationen mit Nebenbedingungen[1]

CARL GEIGER

Institut für Mathematik der Technischen Universität, Clausthal-Zellerfeld, Bundesrepublik Deutschland

1. EINLEITUNG

Bei der Approximation einer auf einem kompakten Intervall stetigen Funktion f durch rationale Funktionen im Sinne der Tschebyscheff-Norm, also bei der Frage nach einer Funktion F_0 aus der Menge V der rationalen Funktionen gegebenen Zähler- und Nennergrades mit der Eigenschaft

$$\max_{x\in[\alpha,\beta]} |f(x)-F_0(x)| = \inf_{F\in V}\left\{\max_{x\in[\alpha,\beta]} |f(x)-F(x)|\right\},$$

ist ein bestimmtes Verhalten der Fehlerfunktion $f-F_0$ von besonderem Interesse: die Fehlerfunktion nimmt ihr Betragsmaximum abwechselnd mit positivem und negativem Vorzeichen an, sie besitzt eine “Alternante”. Dieses Oszillationsverhalten ist insbesondere für die numerische Konstruktion von guten Näherungen von Bedeutung. Wählt man für die approximierenden Funktionen nicht die Menge aller rationalen Funktionen von gegebenem Typ, sondern allgemeiner eine von n Parametern abhängige Funktionenfamilie, so läßt sich, wie Meinardus und Schwedt [*8*] sowie Rice [*10*] gezeigt haben, die beste Approximierende in bestimmten Fällen ebenfalls durch ein Oszillationsverhalten der Fehlerfunktion charakterisieren. Diese Aussagen gelten jedoch schon oft dann nicht mehr, wenn die Ansatzfunktionen zwar rationale Funktionen sind, jedoch zwischen den Koeffizienten bestimmte Bindungen bestehen.

In der vorliegenden Arbeit wird nun diskutiert, wie sich in einem speziellen Fall diese Bindungen auswirken; bezeichnet $P(x)$ ein Polynom von vorgegebenem Grad, so sollen die approximierenden Funktionen gegeben sein durch alle rationalen Ausdrücke der Form

$$P(x)/P(-x). \tag{1.1}$$

Tatsächlich gilt hier, wenn das Approximationsintervall den Punkt $x=0$ enthält, kein Alternantensatz. Doch lassen sich durch Ausnutzen der speziellen

[1] Auszug aus der Dissertation des Verfassers.

Art der Bindungen Aussagen herleiten, die jenen im uneingeschränkten Falle weitgehend analog sind.

Mit der Diskussion dieses speziellen Approximationsproblems ist übrigens eine Reihe verwandter Probleme miterledigt. Schränkt man nämlich die Menge der rationalen Ansatzfunktionen $R(x)$ dadurch ein, daß man das Bestehen einer Beziehung

$$R(l(x)) = L(R(x))$$

fordert, setzt man dabei l und L als linear-gebrochene Funktionen, l überdies als Abbildung des Approximationsintervalls in sich voraus, so sind, falls man überhaupt ein sinnvolles Approximationsproblem erhält, im wesentlichen nur zwei Fälle möglich:

1. die Aufgabe läßt sich zurückführen auf die Approximation mit geraden bzw. ungeraden rationalen Funktionen im Intervall $[-1, +1]$; besitzt die anzunähernde Funktion f selbst diese Symmetrieeigenschaft, so ist das Verhalten der besten Approximierenden bekannt (Invarianzfall, vergl. Meinardus [7]), oder
2. die Aufgabe läßt sich zurückführen auf das im folgenden behandelte Approximationsproblem.

2. Problemstellung. Existenz

Es sei $C[I]$ der Raum der auf dem Intervall $I = [-1, +1]$ stetigen reellwertigen Funktionen. Im folgenden bezeichne stets $f(x)$ die zu approximierende Funktion, $w(x)$ eine fest vorgegebene, in I positive Gewichtsfunktion, und es sei $f, w \in C[I]$. A_n bezeichne die Menge der zulässigen Parametervektoren; dabei gehöre ein $a = (a_0, a_1, \ldots, a_n)$ genau dann zu A_n, wenn die folgenden Bedingungen erfüllt sind

für das a zugeordnete Polynom $P_n(a, x) = \sum_{\nu=0}^{n} a_\nu x^\nu$ gelte

$$P_n(a, -x) > 0 \qquad \text{für } x \in I; \tag{2.1}$$

die Polynome $P_n(a, x)$ und $P_n(a, -x)$ seien teilerfremd. (2.2)

V_n bezeichne die Menge der approximierenden Funktionen

$$F(a, x) = \frac{P_n(a, x)}{P_n(a, -x)}$$

mit $a \in A_n$ (und o.B.d.A. $a_0 = 1$); unter dem Grad von $F(a,x) \in V_n$ werde der Grad von $P_n(a,x)$ verstanden. Offenbar gilt für $F \in V_n$ stets

$$F(a, x) \cdot F(a, -x) = 1 \qquad \text{für } x \in I. \tag{2.3}$$

Die V_n entsprechende Menge aller zulässigen rationalen, nicht der Einschränkung (2.3) unterworfenen Funktionen heiße $R_{n,n}$.

Schließlich sei auf $C[I]$ die Norm

$$\|\varphi\| := \max_{x \in I} (w(x)|\varphi(x)|)$$

eingeführt.

Nun kann das *Approximationsproblem* formuliert werden: zu einer gegebenen Funktion $f \in C[I]$ suche man eine Funktion F_0 der Klasse V_n, die im Sinne der eingeführten Norm f möglichst gut annähert, d.h. für die gilt

$$\|f - F_0\| \leqslant \|f - F\| \qquad \text{für alle } F \in V_n.$$

Eine solche Funktion $F_0 = F(a^0, x)$ heißt *Minimallösung* für f bezüglich V_n, die Größe

$$\rho_n(f) := \inf_{F \in V_n} \|f - F\|$$

heißt *Minimalabweichung*.

SATZ 2.1 (EXISTENZSATZ). *Zu jeder Funktion $f \in C[I]$ und zu jeder Zahl $n \in \{0, 1, 2, \ldots\}$ gibt es eine Minimallösung bezüglich V_n.*

Der Beweis ergibt sich durch einfache Übertragung einer der bekannten Existenzbeweise für uneingeschränkte rationale Approximationen (etwa nach der Idee von Collatz [3], oder nach Rice [9], S. 77).

3. UNTERE SCHRANKEN FÜR $\rho_n(f)$

SATZ 3.1. *Es sei $f \in C[I]$ und $F(a, x) \in V_n$. Ist m ($\leqslant n$) der genaue Grad von $F(a, x)$, so sei zur Abkürzung gesetzt*

$$d(a) = n - \left[\frac{n-m}{2}\right]. \tag{3.1}$$ [2]

(a) *Es gilt stets $\rho_n(f) \geqslant w(0)|f(0) - 1|$.*

(b) *Gibt es zwei Punkte $x_1, x_2 \in I - \{0\}$ mit $x_1 = -x_2$ und*

$$\operatorname{sgn}(f(x_1) - F(a, x_1)) = \operatorname{sgn}(f(x_2) - F(a, x_2)),$$ [3]

so gilt

$$\rho_n(f) \geqslant \min_{i=1,2} (w(x_i)|f(x_i) - F(a, x_i)|).$$

---

[2] $[z]$ bezeichne die größte ganze Zahl $\leqslant z$.

[3] Mit $\operatorname{sgn} z = \begin{cases} 1, & \text{wenn } z > 0 \\ 0, & \text{wenn } z = 0. \\ -1, & \text{wenn } z < 0 \end{cases}$

(c) *Gibt es* $d(a)+1$ *paarweise betragsverschiedene Punkte* $x_i \in I-\{0\}$, *für die in der Anordnung*

$$0 < |x_1| < |x_2| < \ldots < |x_{d(a)+1}|$$

gilt

$$\operatorname{sgn}[x_i(f(x_i)-F(a,x_i))] = -\operatorname{sgn}[x_{i+1}(f(x_{i+1})-F(a,x_{i+1}))],$$

$i=1(1)\,d(a)$, *so ist*

$$\rho_n(f) \geqslant \min_{i=1(1)d(a)+1} (w(x_i)|f(x_i)-F(a,x_i)|).$$

Beweis. Die Aussage (a) ist trivial. Für (b) und (c) wird angenommen, es gebe eine Funktion $F(b,x) \in V_n$ mit

$$\|f-F(b,x)\| < \min_i (w(x_i)|f(x_i)-F(a,x_i)|).$$

Aus der Darstellung

$$F(b,x)-F(a,x) = (f(x)-F(a,x)) - (f(x)-F(b,x))$$

liest man ab

$$\operatorname{sgn}(F(b,x_i)-F(a,x_i)) = \operatorname{sgn}(f(x_i)-F(a,x_i)). \tag{3.2}$$

Mit der Voraussetzung von (b) gilt also

$$\begin{aligned}\operatorname{sgn}(F(b,x_1)-F(a,x_1)) &= \operatorname{sgn}(F(b,x_2)-F(a,x_2)) \\ &= \operatorname{sgn}\left(\frac{1}{F(b,x_1)}-\frac{1}{F(a,x_1)}\right) \\ &= -\operatorname{sgn}(F(b,x_1)-F(a,x_1)).\end{aligned}$$

$F(a,x_i)=F(b,x_i)$ $(i=1,2)$ steht jedoch in Widerspruch zur Annahme. Im Fall (c) gilt wegen (3.2)

$$\operatorname{sgn}(F(b,x_i)-F(a,x_i)) = \epsilon(-1)^i \operatorname{sgn} x_i, \qquad i=1(1)\,d(a)+1$$

mit der festen Zahl $\epsilon=\pm 1$. Der Zähler von $F(b,x)-F(a,x)$ ist ein ungerades Polynom vom Grade $\leqslant 2d(a)-1$, der Nenner ist nach (2.1) positiv in I; also muß $F(b,x)-F(a,x)\equiv 0$ sein in Widerspruch zur Annahme.

Eine nach (a) oder (b) gefundene Schranke für die Minimalabweichung $\rho_n(f)$ weist, wenn sie positiv ist, eine Besonderheit auf: sie ist auch untere Schranke für $\rho_{n+1}(f)$, $\rho_{n+2}(f)$ usf. Welche Konsequenzen dies hat, wird in Abschnitt 5 näher diskutiert werden. Hier soll jedoch noch eine untere Schranke angegeben werden, die von der speziellen Wahl einer Funktion $F(a,x)\in V_n$ unabhängig ist und die nicht schlechter ist als jede nach Satz 3.1 (a) oder (b) konstruierte Schranke.

Satz 3.2. *Es sei*

$$\eta^*(x) := w(x)(f(x) - \varphi^*(x)), \qquad \rho^*(f) = \max_{x \in I} |\eta^*(x)| = \|f - \varphi^*\|$$

mit

$$\varphi^*(x) := \frac{1}{w(x)} \left\{ \frac{w(x) f(x) - w(-x) f(-x)}{2} + \left[\left(\frac{w(x) f(x) - w(-x) f(-x)}{2} \right)^2 + w(x) w(-x) \right]^{1/2} \right\}.$$

Dann ist stets

$$\rho_n(f) \geqslant \rho^*(f).$$

Beweis. Die oben definierten Funktionen φ^* und η^* haben, wie man leicht nachrechnet, folgende Eigenschaften:

$$\varphi^*(x) \cdot \varphi^*(-x) = 1 \qquad \text{für } x \in I, \tag{3.3}$$

$$\eta^*(x) = \eta^*(-x) \qquad \text{für } x \in I. \tag{3.4}$$

Nun gilt mit einer beliebigen Funktion $F(a, x) \in V_n$ und einem $\xi \in I$

$$\left.\begin{aligned} w(\xi)(f(\xi) - F(a, \xi)) &= \eta^*(\xi) + (\varphi^*(\xi) - F(a, \xi)) w(\xi) \\ w(-\xi)(f(-\xi) - F(a, -\xi)) &= \eta^*(\xi) - (\varphi^*(\xi) - F(a, \xi)) \cdot \frac{w(-\xi)}{\varphi^*(\xi) F(a, \xi)}; \end{aligned}\right\} \tag{3.5}$$

hierbei wurde (2.3), (3.3) und (3.4) benutzt. Da die Faktoren bei $(\varphi^*(\xi) - F(a, \xi))$ stets positiv sind, folgt aus (3.5)

$$\max_{x = \xi, -\xi} (w(x) |f(x) - F(a, x)|) \geqslant |\eta^*(\xi)|$$

und hieraus schließlich

$$\|f - F\| \geqslant \rho^*(f).$$

Daß $\rho^*(f)$ die beste von n unabhängige untere Schranke für $\rho_n(f)$ ist, ergibt sich aus der späteren Aussage (5.5).

4. Charakterisierung von Minimallösungen

Eine Minimallösung $F(a, x)$ kann auch im vorliegenden Falle durch das Verhalten der Fehlerfunktion $w(f - F)$ in ihren Extremalpunkten charakterisiert werden. Dabei heißt ein Punkt $\xi \in I$ *Extremalpunkt* von $w(f - F)$, wenn gilt

$$w(\xi) |f(\xi) - F(a, \xi)| = \|f - F\|.$$

Die Menge aller Extremalpunkte von $w(x)(f(x) - F(a, x))$ werde mit M_a bezeichnet. Dann gilt

SATZ 4.1. *Es sei* $f(x) \in C[I]$, $F(a,x) \in V_n$, $f(x) \not\equiv F(a,x)$. $F(a,x)$ *ist genau dann Minimallösung für* $f(x)$ *bezüglich* V_n, *wenn eine der folgenden Bedingungen erfüllt ist:*

I. 0 *ist Extremalpunkt von* $w(x)(f(x) - F(a,x))$;

II. *es gibt zwei Extremalpunkte* x_1, x_2 *von* $w(x)(f(x) - F(a,x))$ *mit* $x_1 = -x_2$ *und* $w(x_1)(f(x_1) - F(a,x_1)) = w(x_2)(f(x_2) - F(a,x_2))$;

III. *es gibt* $d(a) + 1$ *Extremalpunkte* $x_1, \ldots, x_{d(a)+1}$ *von* $w(x)(f(x) - F(a,x))$ *mit*

$$0 < |x_1| < |x_2| < \ldots < |x_{d(a)+1}|$$

und

$$\operatorname{sgn}[x_i(f(x_i) - F(a,x_i))] = -\operatorname{sgn}[x_{i+1}(f(x_{i+1}) - F(a,x_{i+1}))], \qquad i = 1(1)\,d(a)$$

($d(a)$ *ist die durch* (3.1) *definierte Zahl* $\leqslant n$).

Dabei tritt genau dann weder I *noch* II *ein, wenn mit der in Abschnitt* 3 *definierten Zahl* $\rho^*(f)$ *gilt*

$$\|f - F\| > \rho^*(f).$$

Beweis. Jede der Bedingungen I, II, III ist *hinreichend*: in jedem der 3 Fälle liefert Satz 3.1, auf Extremalpunkte angewandt,

$$\rho_n(f) \geqslant \|f - F\|.$$

Die Ungleichung besagt, daß F Minimallösung für f bezüglich V_n ist.

Eine der Bedingungen I, II, III ist *notwendig*: es sei also $F(a,x)$ Minimallösung für f bezüglich V_n. Dann gilt nach Satz 3.2

$$\|f - F\| \geqslant \rho^*(f).$$

Fall 1. $\|f - F\| = \rho^* > 0$

$\xi \in I$ sei ein Extremalpunkt der Funktion $\eta^*(x)$:

$$\eta^*(\xi) = \epsilon\rho^* = \epsilon\|f - F\|, \qquad \epsilon = \pm 1.$$

Ist $\xi = 0$, so gilt wegen $\eta^*(0) = w(0)(f(0) - 1)$ die Aussage I. Andernfalls läßt sich das Gleichungssystem (3.5) anschreiben

$$w(\xi)(f(\xi) - F(a,\xi)) = \epsilon\|f - F\| + (\varphi^*(\xi) - F(a,\xi))\,w(\xi)$$

$$w(-\xi)(f(-\xi) - F(a,-\xi)) = \epsilon\|f - F\| - (\varphi^*(\xi) - F(a,\xi))\frac{w(-\xi)}{\varphi^*(\xi)\,F(a,\xi)}$$

und daraus $\varphi^*(\xi) - F(a,\xi) = 0$ ablesen; mit $x_1 = \xi$, $x_2 = -\xi$ gilt also II.

Fall 2. $\|f-F\| > \rho^*$

Dann ist 0 sicher nicht Extremalpunkt:

$$w(0)|f(0)-1| = |\eta^*(0)| \leqslant \rho^* < \|f-F\|.$$

Und gäbe es ein $\xi \in I$ mit

$$w(\xi)(f(\xi)-F(a,\xi)) = w(-\xi)(f(-\xi)-F(a,-\xi)),$$

so folgte aus (3.5)

$$w(\xi)|f(\xi)-F(a,\xi)| = |\eta^*(\xi)| \leqslant \rho^* < \|f-F\|,$$

d.h. auch II kann nicht eintreten.

Es ist nun noch die Gültigkeit von III nachzuweisen. Da für die Punkte der Menge M_a der Extremalpunkte, wie eben gezeigt, weder I noch II gelten kann, gibt es Zahlen $\xi_0, \xi_1, \ldots, \xi_p$ $(1 \leqslant p < \infty)$, δ, ϵ mit

$$0 = \xi_0 < \xi_1 < \ldots < \xi_p = 1, \qquad \delta > 0,\ \epsilon = 1 \text{ oder } -1,$$

so daß für alle x mit $\xi_i \leqslant |x| \leqslant \xi_{i+1}$ gilt

$$\left.\begin{array}{c} -\|f-F\|+\delta < w(x)(f(x)-F(a,x))\cdot \operatorname{sgn} x \leqslant \|f-F\|, \\ \text{wenn } \epsilon(-1)^i = 1 \\ -\|f-F\| \leqslant w(x)(f(x)-F(a,x))\cdot \operatorname{sgn} x < \|f-F\|-\delta, \\ \text{wenn } \epsilon(-1)^i = -1\,; \end{array}\right\} \tag{4.1}$$

dabei soll in jedem Doppelintervall $[-\xi_{i+1}, -\xi_i] \cup [\xi_i, \xi_{i+1}]$ das Gleichheitszeichen mindestens einmal angenommen werden. Es wird nun die Annahme

$$p \leqslant d(a) \tag{4.2}$$

getroffen. Der dem bekannten Beweis für den Alternantensatz (siehe Achieser [*1*], S. 55) zugrundeliegenden Idee folgend, betrachtet man das Polynom

$$\phi(x) = \epsilon x \prod_{\nu=1}^{p-1} (\xi_\nu^2 - x^2).$$

Dieses Polynom, das also höchstens vom Grad $2d(a)-1 \leqslant n+m$ sein kann, besitzt die Eigenschaft

$$\operatorname{sgn}(x\phi(x)) = \epsilon(-1)^i \qquad \text{für } \xi_i < |x| < \xi_{i+1}, \tag{4.3}$$

$i = 0(1)p-1$. Ist $P(x)$ das Zählerpolynom der Minimallösung $F(a,x)$ ($P(x)$, $P(-x)$ teilerfremd, Grad von $P(x) = m \leqslant n$), so läßt $\phi(x)$ die Darstellung

$$\phi(x) = Q(x)\cdot P(-x) - Q(-x)\cdot P(x)$$

zu; dabei läßt sich wegen (4.2) $Q(x)$ als Polynom vom Grade $\leqslant n$ (mit dem Koeffizientenvektor b) wählen.[4] Mit diesem Polynom $Q(x)$ sei nun

$$F_t = F(a+tb, x) = \frac{P(x)+tQ(x)}{P(-x)+tQ(-x)};$$

für ein $t_0 > 0$ ist $F_t \in V_n$ für $t \in [0, t_0]$.

Damit ist dann

$$w(x)(f(x) - F(a+tb, x))\cdot \operatorname{sgn} x$$
$$= w(x)(f(x) - F(a,x))\cdot \operatorname{sgn} x - t\cdot \frac{w(x)\cdot Q(x)\cdot \operatorname{sgn} x}{P(-x)(P(-x)+tQ(-x))},$$

und es ist

$$\sigma = \max_{\substack{x\in I\\ t\in[0,t_0]}} \frac{w(x)|\phi(x)|}{P(-x)(P(-x)+tQ(-x))}$$

eine endliche Zahl. Wegen (4.1) und (4.3) gilt nun für alle x mit $\xi_i < |x| < \xi_{i+1}$ und alle t mit $0 < t \leqslant \min(t_0, (\delta/2\sigma))$

$$-\|f-F\| + \frac{\delta}{2} < w(x)(f(x) - F(a+tb, x))\cdot \operatorname{sgn} x < \|f-F\|,$$
$$\text{wenn } \epsilon(-1)^i = 1,$$
$$-\|f-F\| < w(x)(f(x) - F(a+tb, x))\cdot \operatorname{sgn} x < \|f-F\| - \frac{\delta}{2},$$
$$\text{wenn } \epsilon(-1)^i = -1,$$

also zusammen mit (4.1)

$$w(x)|f(x) - F(a+tb, x)| < \|f-F\| \qquad \text{für alle } x \in I.$$

Dies stellt einen Widerspruch zur Minimaleigenschaft von $F(a,x)$ dar. Statt (4.2) muß also gelten

$$p \geqslant d(a) + 1.$$

$d(a)+1$ Punkte $x_i \in [-\xi_{i+1}, -\xi_i] \cup [\xi_i, \xi_{i+1}]$ $(i = 1(1)\, d(a)+1)$, für die in (4.1) das Gleichheitszeichen steht, besitzen alle in III geforderten Eigenschaften.

---

[4] Beweis etwa: $P(x)$, $P(-x)$ teilerfremd

$$\Rightarrow \exists \text{ Polynome } R(x), S(x)\colon 1 = R(x)P(-x) - S(x)P(x); \;|\cdot\phi(x)$$

Division mit Rest $R(x)\cdot\phi(x) = T(x)P(x) + U(x)$ mit $\operatorname{Grad} U(x) \leqslant \operatorname{Grad} P(x) = m \leqslant n$; der Teil $T(x)\cdot P(x)$ wird mit in den zweiten Summanden hineingenommen

$$\phi(x) = U(x)P(-x) - V(x)P(x)$$

$\operatorname{Grad}\phi \leqslant m+n$, $\operatorname{Grad} P(x) = m$, $\operatorname{Grad} U \leqslant n \Rightarrow \operatorname{Grad} V \leqslant n$. Die Behauptung folgt aus

$$\phi(x) - \phi(-x) = 2\phi(x) = [U(x) + V(-x)]P(-x) - [U(-x) + V(x)]P(x).$$

Übrigens kann man bei der obigen Division $U(0) = 1$ und damit $Q(0) = 1$ erreichen.

Bemerkungen. Sind die Extremalpunkte mit den in III angegebenen Eigenschaften alle positiv (oder alle negativ), so liegt eine *Alternante* im üblichen Sinne vor. Dies gibt zu folgender Definition Anlaß:

Eine Menge von $k+1$ Punkten $x_1, \ldots, x_{k+1} \in I$ mit

$$0 < |x_1| < |x_2| < \ldots < |x_{k+1}|$$

und

$$w(x_i)(f(x_i) - F(a, x_i)) \cdot \operatorname{sgn} x_i = \epsilon(-1)^i \|f - F\|, \tag{4.4}$$

$\epsilon = 1$ oder -1, $i = 1(1)k+1$, heißt *A-Menge* der Länge $k+1$.

Eine A-Menge ausreichender Länge charakterisiert hier nach Satz 4.1 jedoch nur dann eine Funktion $F \in V_n$ als Minimallösung, wenn $\|f - F\| > \rho^*(f)$ ist (vergl. jedoch Satz 5.1).

Ohne Beweis sei noch eine Eindeutigkeitsaussage angegeben:

SATZ 4.2. *Existiert zu $f \in C[I]$ und $F(a,x) \in V_n$ eine A-Menge der Länge $d(a)+1$, d.h. ist $F(a,x)$ durch Satz* 4.1.III *als Minimallösung charakterisiert, so ist $F(a,x)$ die einzige Minimallösung für f bezüglich V_n.*

Schließlich sei noch bemerkt, daß die Fälle I, II, III von Satz 4.1 durchaus zugleich auftreten können.

5. DIE KLASSE E. $\rho_n(f)$ UND $E_{n,n}(f)$. RELATIVE APPROXIMATIONEN

Die Funktionalgleichung (2.3) und die Bedingung $F(a, 0) = 1$ sondern die Funktionen der Klasse V_n aus allen in I stetigen rationalen Funktionen von Zähler- und Nennergrad n aus. Es ist naheliegend, zu vermuten, daß diejenigen Funktionen f, die selbst dieser Funktionalgleichung mit $f(0) = 1$ genügen, in Hinblick auf die Approximation bezüglich V_n in bestimmter Weise ausgezeichnet sind.

Es werden deshalb alle Funktionen $f \in C[I]$, die den Bedingungen

$$f(x) f(-x) = 1 \qquad \text{für alle } x \in I; \qquad f(0) = +1 \tag{5.1}$$

genügen, zur *Klasse E* zusammengefaßt. Es gilt $V_n \subset E$.

Setzt man zur Abkürzung

$$\eta(x) = w(x)(f(x) - F(a, x)), \tag{5.2}$$

so gilt für $f, F \in E$ die später benötigte Beziehung

$$\eta(-x) = -\frac{w(-x)}{w(x) f(x) F(a, x)} \eta(x). \tag{5.3}$$

SATZ 5.1. (a) *Die in Satz* 3.2 *angegebene untere Schranke $\rho^*(f)$ ist genau dann gleich* 0, *wenn f zur Klasse E gehört.*

(b) *Ist $f \in E$, so ist $F (F \not\equiv f)$ genau dann die eindeutig bestimmte Minimallösung für f bezüglich V_n, wenn eine A-Menge der Länge $d(a)+1$ existiert.*

Bemerkung. Die Umkehrung der Aussage (b) kann falsch sein. Zum Beispiel wird die Minimallösung für $f = e^x + c$ (c Konstante > 0) für alle n durch eine A-Menge charakterisiert, obwohl diese Funktion nicht zu E gehört.

Beweis. (a) Ist $f \in E$, so rechnet man für die in Satz 3.2 definierte Funktion φ^* leicht nach, daß sie in I — unabhängig von der Gewichtsfunktion $w(x)$! — mit f übereinstimmt: also ist $\eta^* \equiv 0$. Umgekehrt folgt aus $\rho^*(f) = 0$ sofort $f \equiv \varphi^*$ und damit wegen (3.3) auch $f \in E$.

(b) Die Behauptung folgt aus (a) und der letzten Aussage von Satz 4.1 sowie aus Satz 4.2.

Liegt eine Funktion f vor, die nicht zur Klasse E gehört, so läßt die dann von f verschiedene Hilfsfunktion φ^* — die nach (3.3) ja zu E gehört — folgende Interpretation zu: φ^* ist Minimallösung für *f bezüglich E.* Um dies zu zeigen, braucht man nur zu bemerken, daß sich der Beweis von Satz 3.2 nicht ändert, wenn man F als beliebige Funktion $\in E$ (statt wie dort $\in V_n$) einführt.

Man kann φ^* dazu benutzen, eine Einschließung für $\rho_n(f)$ herzuleiten. Ist nämlich F^* die Minimallösung für φ^* bezüglich V_n, so gilt wegen

$$\rho_n(f) \leqslant \|f - F^*\| \leqslant \|f - \varphi^*\| + \|\varphi^* - F^*\| = \rho^*(f) + \rho_n(\varphi^*)$$

$$\rho^*(f) \leqslant \rho_n(f) \leqslant \rho^*(f) + \rho_n(\varphi^*). \tag{5.4}$$

Aus dem nächsten Satz kann gefolgert werden, daß jede Funktion der Klasse E beliebig genau durch Funktionen $\in V_n$ approximiert werden kann, wenn man nur n hinreichend groß wählt. Mit (5.4) liefert dies

$$\lim_{n\to\infty} \rho_n(f) = \rho^*(f). \tag{5.5}$$

Um schließlich eine Aussage über $\rho_n(f)$ für $f \in E$ zu gewinnen, wird diese Größe mit der Minimalabweichung $E_{n,n}(f)$ für die Approximation mit uneingeschränkten rationalen Funktionen verglichen.

SATZ 5.2. *Es bezeichne $\rho_n(f)$ die Minimalabweichung für $f \in E$ bezüglich V_n, $E_{n,n}(f)$ diejenige bezüglich der Familie $R_{n,n}$ der rationalen Funktionen mit Zählergrad $\leqslant n$, Nennergrad $\leqslant n$, mit derselben positiven Gewichtsfunktion $w(x)$. Dann gilt mit den Konstanten*

$$\alpha = \max_{x\in I}\left(\frac{w(x)}{w(-x)} f^2(x)\right), \qquad \beta = \max_{x\in I} \frac{f(x)}{w(-x)}$$

für alle n mit $E_{n,n}(f) < 1/\beta$ *die Ungleichung*

$$E_{n,n}(f) \leqslant \rho_n(f) \leqslant \frac{\alpha E_{n,n}(f)}{1 - \beta E_{n,n}(f)}. \tag{5.6}$$

Die Konstante α *nimmt den kleinstmöglichen Wert* 1 *genau dann an, wenn mit einer geraden Funktion* $v(x)$ *gilt*

$$w(x) = \frac{v(x)}{f(x)}.$$

Beweis. Zu der Minimallösung F_0 für f bzgl. V_n gehöre die A-Menge

$$M = \{x_i\}_1^{d(a)+1}, \qquad \text{d.h. es gelte mit } \eta(x) = w(x)(f(x) - F(a,x))$$

$$0 < |x_1| < |x_2| < \ldots < |x_{d(a)+1}|, \qquad |\eta(x_i)| = \|f - F\| = \rho_n(f)$$

und

$$\operatorname{sgn} x_i \, \eta(x_i) = -\operatorname{sgn} x_{i+1} \, \eta(x_{i+1}), \qquad i = 1(1)\, d(a). \tag{5.7}$$

Es wird nun die Menge M um die Punkte $-x_i$ ergänzt und der Größe nach angeordnet; man erhält so die Menge $\bar{M} = \{\xi_i\}_1^{2d(a)+2}$ mit

$$\xi_i = \begin{cases} -|x_{d(a)+2-i}| & \text{für } i = 1(1)\, d(a) + 1 \\ |x_{i-d(a)-1}| & \text{für } i = d(a) + 2(1)\, 2d(a) + 2 \end{cases}$$

Aus (5.3) liest man ab

$$\operatorname{sgn} \eta(\pm |x|) = \pm \operatorname{sgn} x \eta(x)$$

und hiermit folgt aus (5.7) für $i = 1(1)\, 2d(a) + 1$

$$\operatorname{sgn} \eta(\xi_i) = -\operatorname{sgn} \eta(\xi_{i+1}). \tag{5.8}$$

Ist $m \leqslant n$ der genaue Zähler- und Nennergrad von F_0, so besteht $\bar{M}$ wegen

$$d(a) = n - \left[\frac{n-m}{2}\right] \geqslant \frac{n+m}{2}$$

aus mindestens $n + m + 2$ Punkten. Damit gestattet (5.8) die Anwendung des Satzes von de la Vallée-Poussin aus der Theorie der rationalen Approximationen (vergl. etwa Meinardus [*6*], Satz 93.3; Werner [*11*], Satz 10.3) und man erhält eine untere Schranke für die Minimalabweichung $E_{n,n}(f)$

$$E_{n,n}(f) \geqslant \min_{i=1(1)\,2d(a)+2} |\eta(\xi_i)|. \tag{5.9}$$

Ist ξ_i ein Punkt von M, so ist $|\eta(\xi_i)| = \rho_n(f)$. Andernfalls gibt es einen Punkt $x_k \in M$, so daß $\xi_i = -x_k$ ist und man kann $|\eta(\xi_i)|$ unter Benutzung von (5.3) abschätzen

$$|\eta(\xi_i)| = |\eta(-x_k)| = \frac{w(-x_k)\,|\eta(x_k)|}{f(x_k)\,[w(x_k)\,f(x_k) - \eta(x_k)]} \geqslant \frac{\rho_n(f)}{\alpha + \beta \cdot \rho_n(f)}$$

mit den im Satz angegebenen Zahlen α, β. Aus (5.9) folgt damit

$$E_{n,n}(f) \geqslant \frac{\rho_n(f)}{\alpha + \beta \cdot \rho_n(f)}$$

und hieraus die Behauptung (5.6).

Schließlich folgt aus $\alpha = 1$ für die Funktion $v(x) = w(x) f(x)$ sofort $v(x) \leqslant v(-x)$ für $x \in I$, d.h. $v(x)$ muß gerade sein.

Vom Standpunkt der Anwendung ist zu berücksichtigen, daß bei der Approximation einer Funktion f bezüglich V_n nur halbsoviele Parameter zu bestimmen sind wie bei der Approximation bezüglich $R_{n,n}$; der im Falle $f \in E$ durch Satz 5.2 abgeschätzte, durch die Beschränkung auf V_n entstehende Genauigkeitsverlust erscheint diesem Vorteil gegenüber gering. Oft wird schon $\rho_{n+1}(f) < E_{n,n}(f)$ sein.

Besonders auffällig ist der sehr geringe Unterschied zwischen $\rho_n(f)$ und $E_{n,n}(f)$ im Falle insbesondere der *relativen Approximation*, d.h. der Approximation einer positiven Funktion f bezüglich V_n mit dem Gewicht $w(x) = 1/f(x)$. Einen Hinweis auf diese Sonderstellung der relativen Approximation liefert auch die Beobachtung, daß die Fehlerfunktion $\eta^{\text{rel}}(x)$ nach (5.3) der Beziehung

$$\eta^{\text{rel}}(-x) = -\frac{1}{1 - \eta^{\text{rel}}(x)} \cdot \eta^{\text{rel}}(x)$$

genügt, also für gute Approximationen eine nahezu ungerade Funktion ist. Man kann nun zeigen: durch "kleine" Abänderung von η^{rel} kann man stets eine ungerade Funktion $\bar{\eta}$ erhalten, die als Fehlerfunktion eines bestimmten anderen Approximationsproblems auftritt; bei diesem zugeordneten Problem handelt es sich um die Annäherung einer ungeraden Funktion $\bar{f}$ durch ungerade rationale Funktionen mit geradem Gewicht $\bar{w}$ (Näheres siehe [*4*], Seite 30 ff.). Übrigens erlaubt dieser Sachverhalt eine günstige Berechnung guter relativer Approximationen für Funktionen $f \in E$ bezüglich $R_{n,n}$ auch ohne die in dieser Arbeit behandelte Theorie (vergl. Cody und Ralston [*2*]).

6. Normalität. Konstruktion von Minimallösungen. Beispiele.

Auch bei dem hier behandelten Approximationsproblem kann die Erscheinung auftreten, daß sich trotz Erhöhung der Parameterzahl im Ansatz für $F(a,x)$ keine bessere Näherung für f ergibt. Man definiert: eine Funktion $f \in C[I]$ heißt *n-normal*, wenn keine Minimallösung bezüglich V_n zugleich Minimallösung bezüglich $V_{n'}$, $n' \neq n$, ist (vergl. Meinardus [*6*], S. 156). Liegt eine Minimallösung bezüglich V_n vor, so kann man dem Verlauf der Fehlerfunktion entnehmen, ob f n-normal ist. Man beweist nämlich leicht: *f ist genau dann n-normal, wenn gilt: ist $F \in V_n$ Minimallösung für f bezüglich V_n,* *so ist F vom genauen Grad n, es gibt eine A-Menge der Länge $n+1$, jedoch keine der Länge $n+2$, und es liegt keiner der Fälle* I, II *von Satz* 4.1 *vor.* Eine notwendige und hinreichende Bedingung für die Normalität, die nicht die Lösung eines Approximationsproblems voraussetzt, ist nicht bekannt; doch können die für den uneingeschränkten rationalen Fall gefundenen Normalitätskriterien von Loeb [*5*] und von Meinardus und Schwedt [*8*] auf das vorliegende Problem übertragen werden (siehe [*4*], Sätze 8.2 und 8.3).

Die Normalität ist eine für die *numerische Konstruktion guter Näherungen* wichtige Eigenschaft. Dies soll hier nicht näher erläutert werden; es sei nur erwähnt, daß man sich in den meisten Fällen auch helfen kann, wenn mit einer Anomalität gerechnet werden muß. Der wesentliche Teil der numerischen Rechnung besteht in der Anwendung eines "linearisierten Remez-Algorithmus", der sich vom uneingeschränkten rationalen Falle her (vergl. etwa [*6*], S. 160 ff.; [*11*], S. 173 ff.) ohne Schwierigkeiten übertragen läßt (siehe auch [*4*]). Auch hier kann das Verfahren versagen, doch erhält man meist in wenigen Iterationsschritten sehr gute Näherungen.

Beispiel 1. $f = e^x + 0.0005, \quad w = 1$
f ist n-normal für alle $n = 1, 2, 3, \ldots$, insbesondere ist

$$\rho_n(f) > \rho^*(f) = 0.0005.$$

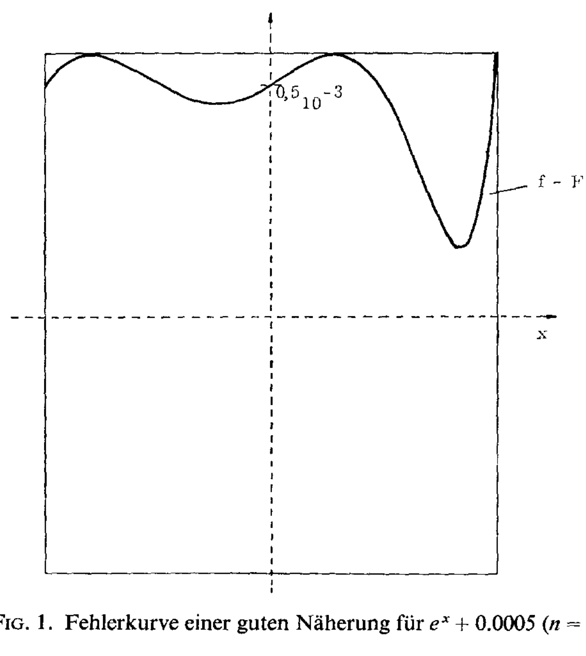

FIG. 1. Fehlerkurve einer guten Näherung für $e^x + 0.0005$ ($n = 2$).

Die Näherung

$$F(a,x) = \frac{1 + 0.499847038x + 0.081669218x^2}{1 - 0.499847038x + 0.081669218x^2}$$

liefert mit Satz 3.1(c) die Einschließung

$$0.56944_{10}{}^{-3} \leqslant \rho_2(e^x + 0.0005) \leqslant 0.56946_{10}{}^{-3}.$$

Beispiel 2. $f = e^{x(3-4x^2)}$, $\quad w = 1$

Es ist $f \in E$, jedoch ist f für kein n n-normal: die Minimallösung bezüglich V_{3k} ist auch Minimallösung bezüglich V_{3k+1}, V_{3k+2}. Für

$$F(a,x) = \frac{1 + 1.404087x - 1.872116x^3}{1 - 1.404087x + 1.872116x^3}$$

erhält man

$$0.41321_{10}{}^{-1} \leqslant \rho_3(e^{x(3-4x^2)}) \leqslant 0.41322_{10}{}^{-1}.$$

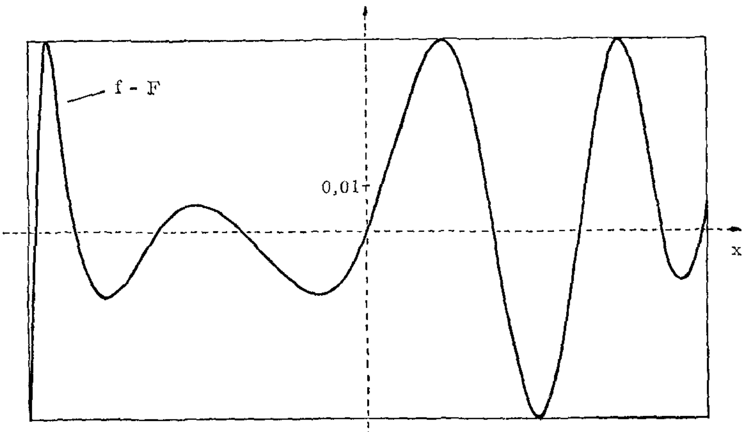

FIG. 2. Fehlerkurve einer guten Näherung für $e^{x(3-4x^2)}$ $(n = 3)$.

LITERATUR

1. N. I. ACHIESER, "Theory of Approximation." Ungar, New York, 1956.
2. W. J. CODY AND A. RALSTON, A note on computing approximations to the exponential function. *Commun. ACM* **10** (1967), 53–55.
3. L. COLLATZ, Tschebyscheffsche Annäherung mit rationalen Funktionen. *Abhandl. Math. Sem. Univ. Hamburg* **24** (1960), 70–78.
4. C. GEIGER, Tschebyscheffsche Approximationen mit rationalen Funktionen unter gewissen Nebenbedingungen. Dissertation, Technischen Universität Clausthal, 1967.
5. H. L. LOEB, Rational Chebyshev Approximation. *Notices Am. Math. Soc.* **11** (1964), 335.
6. G. MEINARDUS, Approximation von Funktionen und ihre numerische Behandlung. Springer, Berlin, 1964.

7. G. MEINARDUS, Invarianz bei rationalen Approximationen. *Computing* **2** (1966), 115–118.
8. G. MEINARDUS UND D. SCHWEDT, Nicht-lineare Approximationen. *Arch. Rational Mech. Anal.* **17** (1964), 297–326.
9. J. R. RICE, "The Approximation of Functions," Vol. 1. Addison-Wesley, Reading, Mass., 1964.
10. J. R. RICE, The characterization of best nonlinear Tchebycheff approximations. *Trans. Am. Math. Soc.* **96** (1960), 322–340.
11. H. WERNER, Vorlesung über Approximationstheorie. Springer, Berlin, 1966.
12. H. WERNER, Tschebyscheff-Approximation im Bereich der rationalen Funktionen bei Vorliegen einer guten Ausgangsnäherung. *Arch. Rational Mech. Anal.* **10** (1962), 205–219.
13. H. WERNER, Die konstruktive Ermittlung der Tschebyscheff-Approximierenden im Bereich der rationalen Funktionen. *Arch. Rational Mech. Anal.* **11** (1962), 368–384.