

BAYERISCHE AKADEMIE DER WISSENSCHAFTEN
MATHEMATISCH-NATURWISSENSCHAFTLICHE KLASSE

SITZUNGSBERICHTE

JAHRGANG

1966

MÜNCHEN 1967

VERLAG DER BAYERISCHEN AKADEMIE DER WISSENSCHAFTEN

In Kommission bei der C. H. Beck'schen Verlagsbuchhandlung München

Polynome minimaler Steigung und die Approximation durch Polynome

Von Dietbert Ehresmann in München

Vorgelegt von Herrn Georg Aumann am 8. Juli 1966

1. Einleitung

Wir betrachten die folgende *Minimumsaufgabe I*: Fest vorgegeben seien die reelle Zahlenfolge $Y = (Y_0, Y_1, \dots, Y_n)$, $n \geq 1$, wobei wir zunächst $Y_i \neq Y_{i-1}$ voraussetzen für $1 \leq i \leq n$, und das Intervall $[a, b]$ der reellen Achse. Gibt es unter den „Stellenverteilungen“ $X = (X_0, X_1, \dots, X_n)$ mit

$$(I) \quad b \geq X_0 > X_1 > \dots > X_n \geq a$$

eine, für die die n -te Steigung $S(X, Y)$ das Minimum ihres absoluten Betrages annimmt und wieviele? Dabei ist

$$S(X, Y) = \sum_{i=0}^n \frac{Y_i}{\prod_{i \neq j} (X_i - X_j)}.$$

Es zeigt sich, daß die Beantwortung dieser Frage nur dann nicht trivial ist, wenn die vorgegebene reelle Zahlenfolge als „oszillierend“ vorausgesetzt wird:

Definition 1: Die reelle Zahlenfolge $Y = (Y_0, Y_1, \dots, Y_n)$, $n \geq 1$, heißt *oszillierend*, wenn $Y_i \neq Y_{i-1}$ ist für $1 \leq i \leq n$, und wenn $\text{sign}(Y_i - Y_{i-1}) = (-1)^{i-1} \cdot \text{sign}(Y_1 - Y_0)$, d. h. wenn die „größer-als“- und die „kleiner-als“-Relation zwischen benachbarten Elementen abwechseln.

Die Lösung der so eingeschränkten Aufgabe geht Hand in Hand mit der Beantwortung der folgenden *Frage II*: Die oszillierende Folge $Y = (Y_0, Y_1, \dots, Y_n)$ und das Intervall $[a, b]$ der

reellen Achse seien wieder vorgegeben; gibt es unter den Stellenverteilungen $X = (X_0, X_1, \dots, X_n)$ mit

$$(II) \quad b = X_0 > X_1 > \dots > X_n = a$$

solche, für die das Interpolationspolynom $P_{X,Y}$ die Werte $Y_0, Y_1, \dots, Y_n$ als relative Extremwerte annimmt?

Dabei ist $P_{X,Y}$ erklärt durch

$$P_{X,Y}(X_i) = Y_i, \quad i = 0, 1, \dots, n.$$

Das erste Ergebnis dieser Arbeit ist die Äquivalenz der Fragestellungen I und II in dem Sinn, daß beide bei vorgegebener oszillierender Folge Y in *eindeutiger* Weise zu ein und derselben Stellenverteilung X aus $[a, b]$ führen (vgl. Satz 1, 2 und 3). Ein konvergentes Verfahren zur Bestimmung dieser eindeutigen Stellenverteilung X wird angegeben.

Im letzten Abschnitt wenden wir unsere Ergebnisse auf ein *approximationstheoretisches Problem* an: Für den bei der Bestapproximation von in $[-1, 1]$ ϱ -mal beschränkt differenzierbaren Funktionen f durch Polynome von nicht größerem als n -tem Grade auftretenden Fehler $\delta(n, \varrho)$ werden für $n = \varrho - 1$ und $n = \varrho$ „Abschätzungsfaktoren“ $\gamma(n, \varrho)$ bestimmt, mit denen die Ungleichung

$$\delta(n, \varrho) \leq \gamma(n, \varrho) \cdot \sup_{x \in [-1, 1]} |f^{(\varrho)}(x)|, \quad n = \varrho - 1, \varrho,$$

erfüllt ist. Dabei ist $\gamma(\varrho - 1, \varrho) = \frac{1}{2^{\varrho-1} \varrho!}$ scharf. Während $\gamma(\varrho - 1, \varrho)$ bereits auf anderem Wege bestimmt worden war [S. N. Bernstein], ist unser Faktor $\gamma(\varrho, \varrho)$ besser als die bisher für diesen Fall bekannten Abschätzungen.

Diese Arbeit ist ein Auszug aus der Abhandlung des Verfassers, die auf Grund einer Anregung von G. Aumann entstand und von der Fakultät für Allg. Wissenschaften der Technischen Hochschule München als Dissertation angenommen wurde. Die Beweise, die hier zu einem Teil weggelassen oder nur skizziert sind, sind dort ausführlich dargelegt.

2. Vorbereitungen

Wir wollen zunächst verschiedene Symbole und Bezeichnungen zusammenstellen:

$(X, Y) = (X_0, X_1, \dots, X_n; Y_0, Y_1, \dots, Y_n)$ heißt *Verteilung* der Folge $Y = (Y_0, Y_1, \dots, Y_n)$ aus $[a, b]$, wenn $b \geq X_0 > X_1 > \dots > X_n \geq a$.

$P_{X,Y}$ bezeichnet das zur Verteilung $(X, Y) = (X_0, X_1, \dots, X_n; Y_0, Y_1, \dots, Y_n)$ gehörige Interpolationspolynom, das durch $P_{X,Y}(X_i) = Y_i$, $i = 0, 1, \dots, n$, erklärt ist.

C^n bezeichnet die Menge der im Intervall $[a, b]$ reellwertigen, n -mal beschränkt differenzierbaren Funktionen.

$C_{X,Y}^n := \{f: f \in C^n, f(X_i) = Y_i \text{ für } i = 0, 1, \dots, n \text{ und } \left. \frac{df}{dx} \right|_{x=X_i} = 0 \text{ für } 1 \leq i \leq n-1\}$.

$$V(X) = \begin{vmatrix} 1 & X_n & X_n^2 & \dots & X_n^{n-1} & X_n^n \\ 1 & X_{n-1} & X_{n-1}^2 & \dots & X_{n-1}^{n-1} & X_{n-1}^n \\ \cdot & & & & & \\ \cdot & & & & & \\ 1 & X_1 & X_1^2 & \dots & X_1^{n-1} & X_1^n \\ 1 & X_0 & X_0^2 & \dots & X_0^{n-1} & X_0^n \end{vmatrix}$$

ist die Vandermondesche Determinante von X .

$$W(X, Y) = \begin{vmatrix} 1 & X_n & X_n^2 & \dots & X_n^{n-1} & Y_n \\ 1 & X_{n-1} & X_{n-1}^2 & \dots & X_{n-1}^{n-1} & Y_{n-1} \\ \cdot & & & & & \\ \cdot & & & & & \\ 1 & X_1 & X_1^2 & \dots & X_1^{n-1} & Y_1 \\ 1 & X_0 & X_0^2 & \dots & X_0^{n-1} & Y_0 \end{vmatrix}$$

$S(X, Y) = \frac{W(X, Y)}{V(X)}$ ist die n -te Steigung der Folge Y bezüglich der Stellenverteilung X .

$S[X, f] = \frac{W[X, f]}{V(X)}$ bezeichnet die n -te Steigung der Funktion $f \in C^n$ bezüglich der Stellenverteilung X .

$Z^{[i]}$ geht aus der Folge $Z = (Z_0, Z_1, \dots, Z_n)$ durch Streichen des Elementes Z_i hervor.

$$D_i : = \frac{\partial}{\partial X_i}.$$

$$D_{ij} : = \frac{\partial^2}{\partial X_i \partial X_j}.$$

Das Vorzeichen von $S(X, Y)$ wird fixiert durch

Hilfssatz 1. Ist (X, Y) eine Verteilung der oszillierenden Folge Y , dann ist die Steigung $S(X, Y)$ von Null verschieden und genau dann positiv, wenn $Y_0 > Y_1$ ist.

Der Beweis kann leicht durch vollständige Induktion mit Hilfe der Rekursionsformel

$$(III) \quad S(X, Y) = \frac{1}{X_0 - X_n} (S(X^{[n]}, Y^{[n]}) - S(X^{[0]}, Y^{[0]}))$$

geführt werden. Hilfssatz 1 erlaubt, die beiden Fälle $S > 0$ und $S < 0$ im folgenden einheitlich zu behandeln.

Ist die Folge Y bzw. die Funktion $f \in C^n$ fest vorgegeben, dann werden wir häufig $S(X, Y)$ bzw. $S[X, f]$ auffassen als Funktion der $n + 1$ Veränderlichen $X_0, X_1, \dots, X_n$. Dabei sind diese auf den durch (I) gekennzeichneten Bereich beschränkt. Es gilt

Hilfssatz 2. Ist $(\bar{X}, Y) = (\bar{X}_0, \bar{X}_1, \dots, \bar{X}_n; Y_0, Y_1, \dots, Y_n)$ eine vorgegebene Verteilung der Folge Y , dann gilt für $i = 0, 1, \dots, n$

$$D_i |S(X, Y)| \Big|_{X=\bar{X}} = (-1)^{i+1} \text{sign}(Y_0 - Y_1) \frac{V(\bar{X}^{[i]})}{V(\bar{X})} \frac{dP_{\bar{X}, Y}}{dx} \Big|_{x=\bar{X}_i}.$$

Beweis. $S[X, P_{\bar{X}, Y}]$ ist eine Konstante, also ist

$$D_i |S(X, Y)| = D_i |S(X, Y)| - D_i |S[X, P_{\bar{X}, Y}]|.$$

Da $W[\bar{X}, P_{\bar{X}, Y}] = W(\bar{X}, Y)$ ist, reduziert sich diese Differenz wegen

$$D_i |S(X, Y)| = \frac{1}{V^2(\bar{X})} (V(X) D_i |W(X, Y)| - |W(X, Y)| D_i V(X))$$

und

$$D_i |S[X, f]| = \frac{1}{V^2(\bar{X})} (V(X) D_i |W[X, f]| - |W[X, f]| D_i V(X))$$

zu der Gleichung

$$\begin{aligned} D_i |S(X, Y)| \Big|_{x=\bar{X}} &= \frac{1}{V(\bar{X})} (D_i |W(X, Y)| - D_i |W[X, P_{\bar{X}, Y}]| \Big|_{x=\bar{X}}) \\ &= \frac{\text{sign}(Y_0 - Y_1)}{V(\bar{X})} (-1)^{i+1} V(\bar{X}^{[i]}) \frac{dP_{\bar{X}, Y}(x)}{dx} \Big|_{x=\bar{X}}; \end{aligned}$$

denn $D_i |W(X, Y)|$ bzw. $D_i |W[X, f]|$ geht aus $|W(X, Y)|$ bzw. $|W[X, f]|$ dadurch hervor, daß die Zeile $(1, X_i, X_i^2, \dots, X_i^{n-1}, Y_i) = w_i(X, Y)$ bzw. $w_i[X, f] = (1, X_i, X_i^2, \dots, X_i^{n-1}, f(X_i))$ ersetzt wird durch

$$D_i w_i(X, Y) = (0, 1, 2X_i, \dots, (n-1)X_i^{n-2}, 0) \text{ bzw.}$$

$$D_i w_i[X, f] = \left(0, 1, 2X_i, \dots, (n-1)X_i^{n-2}, \frac{df}{dx} \Big|_{x=X_i}\right).$$

Hilfssatz 3. Ist die oszillierende Folge $Y = (Y_0, Y_1, \dots, Y_n)$ vorgegeben, und nimmt der absolute Betrag der Steigung $S(X, Y)$, X aus $[a, b]$, für $X = \bar{X}$ ein relatives Minimum an, dann gilt $\bar{X}_n = a$ und $\bar{X}_0 = b$.

Diese Aussage folgt unmittelbar aus den Beziehungen

$$\text{sign } D_0 |S(X, Y)| \Big|_{x=\bar{X}} = -1 \text{ und } \text{sign } D_n |S(X, Y)| \Big|_{x=\bar{X}} = +1,$$

die man leicht mit Hilfe von Hilfssatz 2 erhält.

3. Der Äquivalenzsatz

Satz 1. Ist die oszillierende Folge $Y = (Y_0, Y_1, \dots, Y_n)$ vorgegeben, dann sind die beiden folgenden Aussagen äquivalent:

- (1) Der absolute Betrag der n -ten Steigung $S(X, Y)$, X aus $[a, b]$, nimmt für $X = \bar{X}$ ein relatives Minimum an.
- (2) Das Interpolationspolynom $P_{\bar{X}, V}$, $\bar{X} = (b, \bar{X}_1, \bar{X}_2, \dots, \bar{X}_{n-1}, a)$, gehört zu $C_{\bar{X}, V}^n$.

Beweis. Unmittelbar aus Hilfssatz 2 folgt die Äquivalenz von (2) und der Aussage

$$(1') \quad X_n = a, X_0 = b \text{ und } D_i |S(X, Y)| \Big|_{X=\bar{X}} = 0 \text{ für } 1 \leq i \leq n-1.$$

Es bleibt nur noch zu zeigen, daß aus (2) die in (1) darüberhinaus noch enthaltene Minimumeigenschaft folgt. Dies aber erledigt sich nun in der einfachen Weise, daß sich die Matrix

$$(D_{ij} |S(X, Y)| \Big|_{X=\bar{X}}), \quad 1 \leq i, j \leq n-1,$$

der zweiten partiellen Ableitungen von $|S(X, Y)|$ als Diagonalmatrix mit positiven Elementen in der Hauptdiagonalen ergibt. Man erhält nämlich nach kurzer Rechnung

$$D_{ji} |S(X, Y)| \Big|_{X=\bar{X}} = \frac{1}{V(X)} (D_{ji} |W(X, Y)| - |S(X, Y)| D_{ji} V(X)) \Big|_{X=\bar{X}}$$

und mit $\bar{P} = P_{\bar{X}, V}$

$$D_{ji} |S[X, \bar{P}]| \Big|_{X=\bar{X}} = \frac{1}{V(X)} (D_{ji} |W[X, \bar{P}]| - |S[X, \bar{P}]| D_{ji} V(X)) \Big|_{X=\bar{X}}$$

und hieraus, da $D_{ji} |S[X, \bar{P}]| \Big|_{X=\bar{X}} = 0$,

$$D_{ji} |S(X, Y)| \Big|_{X=\bar{X}} = \frac{1}{V(X)} (D_{ji} |W(X, Y)| - D_{ji} |W[X, \bar{P}]|) \Big|_{X=\bar{X}}.$$

Wie man leicht bestätigt, verschwindet diese Differenz für $i \neq j$, $1 \leq i, j \leq n-1$. Für $i = j$, $1 \leq i \leq n-1$, ergibt sich

$$\text{sign } D_{ii} |S(X, Y)| \Big|_{X=\bar{X}} = +1.$$

Damit ist Satz 1 bewiesen.

4. Die Existenz einer Lösung

Hilfssatz 4. Ist die oszillierende Folge $Y = (Y_0, Y_1, \dots, Y_n)$ vorgegeben, so wächst $|S(X, Y)|$ über alle Schranken, wenn zwei benachbarte Elemente $X_{i-1}, X_i, 1 \leq i \leq n$, der Stellenverteilung X gegen denselben Wert streben und wenn zugleich $X_0 - X_n$ beschränkt bleibt.

Dieser Hilfssatz kann wieder leicht durch vollständige Induktion nach n bewiesen werden.

Satz 2. Zu jeder oszillierenden Folge $Y = (Y_0, Y_1, \dots, Y_n)$ gibt es eine Stellenverteilung $\bar{X}$, für die $|S(X, Y)|$, X aus $[a, b]$, sein Minimum annimmt.

Beweis. Wegen $\inf_{X \text{ aus } [a, b]} |S(X, Y)| \geq 0$ haben wir nur zu zeigen, daß es eine Verteilung $\bar{X}$ aus $[a, b]$ gibt, für die $|S(X, Y)|$ sein Infimum wirklich annimmt. Auf Grund von Hilfssatz 4 ist aber klar, daß man aus jeder „Minimalfolge“ $(^{(\lambda)}X = (^{(\lambda)}X_0, ^{(\lambda)}X_1, \dots, ^{(\lambda)}X_n)$ von Stellenverteilungen aus $[a, b]$, für die

$$\lim_{\lambda \rightarrow \infty} |S(^{(\lambda)}X, Y)| = \inf_{X \in [a, b]} |S(X, Y)|$$

gilt, eine gegen eine Stellenverteilung $\bar{X} = (\bar{X}_0, \bar{X}_1, \dots, \bar{X}_n)$ konvergente Teilfolge auswählen kann. Und da $|S(X, Y)|$ eine stetige Funktion von X ist, ist damit die Behauptung bewiesen.

Dieses Ergebnis bedeutet nach Satz 1 zugleich, daß es zu vorgegebener oszillierender Folge Y immer eine Stellenverteilung $\bar{X}$ gibt, mit der $P_{\bar{X}, Y}$ zu $C_{\bar{X}, Y}^n$ gehört.

5. Die Eindeutigkeit der Lösung

Satz 3. Zu jeder oszillierenden Folge $Y = (Y_0, Y_1, \dots, Y_n)$ gibt es höchstens eine Verteilung $(\bar{X}, Y)$ von Y derart, daß $\bar{X}_0 = b$ und $\bar{X}_n = a$ ist und daß $P_{\bar{X}, Y} \in C_{\bar{X}, Y}^n$.

Beweis. Wir gehen aus von zwei Stellenverteilungen $^{(\lambda)}X = = ^{(\lambda)}X_0, ^{(\lambda)}X_1, \dots, ^{(\lambda)}X_n$, $\lambda = 1, 2$, mit $^{(\lambda)}X_0 = b$ und $^{(\lambda)}X_n = = a$, für die $P_\lambda = P_{(\lambda)X,Y} \in C^n_{(\lambda)X,Y}$, und fragen nach derjenigen stetigen Abbildung $x_2 = \varphi(x_1)$ der reellen Achse auf sich, die die folgenden Eigenschaften hat:

- (1) $x_2 = \varphi(x_1)$ ist monoton und $\varphi(a) = a$.
- (2) Für alle reellen x_1 gilt $P_1(x_1) = P_2(\varphi(x_1))$.

Da $P_1(x)$ und $P_2(x)$ die gleichen relativen Extremwerte in derselben Reihenfolge annehmen, existiert die gesuchte Abbildungsfunktion sicher und ist durch die Eigenschaften (1) und (2) eindeutig charakterisiert.

Insbesondere ordnet φ jedem Punkt $^{(1)}X_i$ der Stellenverteilung $^{(1)}X$ den Punkt $^{(2)}X_i$ der Stellenverteilung $^{(2)}X$ zu, $i = 0, 1, \dots, n$. Unser Ziel ist zu zeigen, daß die Funktion $\varphi(x)$ unter den gestellten Bedingungen nur die Identität sein kann.

a) Für alle reellen x_1 ist φ holomorph und umkehrbar eindeutig. Kritisch sind dabei nur die Punkte $^{(1)}X_i$, $1 \leq i \leq n-1$, der Stellenverteilung $^{(1)}X$. Für diese Punkte aber folgen beide Eigenschaften auf Grund der Monotonieforderung (1) für φ .

Wir betrachten nun die analytische Fortsetzung Φ der Funktion φ ins Komplexe und schreiben z_λ statt x_λ , $\lambda = 1, 2$. Dabei verstehen wir unter $z_2 = \Phi(z_1)$ eine Lösung der Gleichung

$$(IV) \quad P_1(z_1) = P_2(z_2),$$

deren Einengung auf die reelle Achse mit $x_2 = \varphi(x_1)$ übereinstimmt. Aus dem Newton-Diagramm der Gleichung, die aus (IV) durch die Transformation $z'_\lambda = \frac{1}{z_\lambda}$, $\lambda = 1, 2$, hervorgeht, entnimmt man ferner, daß Φ im Punkte $z_1 = \infty$ einen Pol erster Ordnung hat. Damit ist also Φ in einem Gebiet G^* , das die reelle Achse und den unendlich fernen Punkt in seinem Innern enthält, eindeutig. Wir wählen nun ein Teilgebiet G von G^* derart, daß G die Punkte $^{(1)}X_i$, $0 \leq i \leq n$, und den Punkt ∞ in seinem Innern enthält, daß auch das Bild $\Phi(G)$ die Punkte $^{(2)}X_i$, $0 \leq i \leq n$, im Innern enthält und daß sein Komplement einfach zusammenhängend ist.

b) Mit H_2 bezeichnen wir die der w_2 -Ebene überlagerte Riemannsche Fläche der Umkehrfunktion von $w_2 = P_2(z_2)$.

Das Funktionenpaar

$$z_2 = \Phi(z_1), \quad w_2 = P_1(z_1)$$

bildet G auf einen Teil H_2' von H_2 ab, der *alle* Windungspunkte von H_2 enthält. Im Komplement $H_2'' = H_2 - H_2'$ ist die Umkehrfunktion eindeutig. Läßt man z_1 auch noch das Komplement von G durchlaufen, so beschreibt das Funktionenpaar

$$z_2 = \zeta(P_1(z_1)), \quad w_2 = P_1(z_1)$$

das Komplement H_2'' . Daher ist $z_2 = \zeta_2(P_1(z_1))$ die analytische Fortsetzung von $z_2 = \Phi(z_1)$. $\zeta_2(P_1(z_1))$ ist aber eindeutig und regulär. Also ist $z_2 = \Phi(z_1)$ in der ganzen z_1 -Ebene eindeutig und regulär fortsetzbar. Als ganze Funktion, die im Punkte $z_1 = \infty$ einen Pol erster Ordnung hat, ist Φ linear, und wegen $\Phi(a) = a$ und $\Phi(b) = b$ folgt $P_1 \equiv P_2$. Damit ist Satz 3 bewiesen.

Folgerung. Zu jeder vorgegebenen oszillierenden Folge $Y = (Y_0, Y_1, \dots, Y_n)$ gibt es genau eine Stellenverteilung $\bar{X} = (\bar{X}_0, \bar{X}_1, \dots, \bar{X}_n)$ derart, daß

1. $P_{\bar{X}, Y} \in C_{\bar{X}, Y}^n$
2. $|S(X, Y)|$, X aus $[a, b]$, für $X = \bar{X}$ sein absolutes und zugleich relatives Minimum annimmt.

Bemerkung zu den Beweismitteln. Während wir alle übrigen Sätze dieser Arbeit mit Mitteln der reellen Analysis beweisen, haben wir bei Satz 3 die Theorie der Funktionen einer komplexen Veränderlichen zu Hilfe genommen. In dem für unsere Anwendungen wichtigen Spezialfall, daß Y eine alternierende Folge betragsgleicher Elemente ist, läßt sich aber ein Beweis geben, der sich nur auf die reelle Analysis stützt (vgl. die Bemerkung im Anschluß an Satz 4).

6. Zur Konstruktion der Lösung

Zu einer vorgegebenen oszillierenden Folge kann man die nach dem Vorangehenden eindeutig bestimmte Stellenverteilung $\bar{X}$ folgendermaßen approximieren: Man gehe von einer Verteilung (X, Y) von Y aus $[a, b]$ aus, bei der $X_0 = b$, $X_n = a$, die aber sonst beliebig ist. Nun kann man durch eine kleine Verschiebung jeweils nur eines der Punkte X_i , $1 \leq i \leq n-1$, den absoluten Betrag $S(X, Y)$ verkleinern, wenn man X_i in der durch Hilfsatz 2 vorgeschriebenen Richtung verschiebt. Man hat dazu im wesentlichen nur den Wert der Ableitung von $P_{X,Y}$ an der Stelle $x = X_i$ zu bestimmen.

Dieses Verfahren konvergiert, da $|S(X, Y)|$ nur ein einziges relatives Minimum besitzt.

Bemerkung. Die obigen Untersuchungen und Resultate lassen sich ausdehnen auf „schwach oszillierende“ Folgen, welche als Grenzfälle von oszillierenden Folgen definiert werden können [3].

7. Anwendungen auf Abschätzungsfaktoren in Approximationsproblemen

Es sei jetzt $[a, b] = [-1, 1]$.

A) Auf dem Vektorraum C der im Intervall $[-1, 1]$ der reellen Achse reellwertigen, beschränkten Funktionen betrachten wir die Norm

$$p[f] = \sup_{|x| \leq 1} |f|$$

und die Pseudonorm

$$q_r[f] = \begin{cases} p[f^{(r)}], & \text{falls } f \text{ } r\text{-mal beschränkt} \\ \text{differenzierbar ist, } \infty & \text{sonst.} \end{cases}$$

Dabei bezeichnet $f^{(r)}$ die r -te Ableitung von f nach x in $[-1, 1]$. Wenn nicht anders vermerkt, ist stets $r \geq 1$ vorausgesetzt.

Wir wollen nun das *Approximationsproblem* $\mathfrak{A}(n, r)$ erklären: Die zu approximierenden Elemente sind die Elemente des Teilraumes $C^r = \{g : g \in C^r \text{ und } q_r[g] < \infty\}$ von C , also die r -mal

differenzierbaren Funktionen mit beschränkter r -ter Ableitung; der Teilraum der approximierenden Elemente ist der Nullraum $A^{[n+1]}$ von q_{n+1} ; er besteht genau aus den Polynomen höchstens n -ten Grades.

Nach dem Satz von A. Haar gibt es zu jedem $g \in C^r$ ein Element aus $A^{[n+1]}$, das wir mit $P_n[g]$ bezeichnen wollen, für das gilt:

$$p[g - P_n[g]] < p[g - a]$$

für alle $a \neq P_n[g]$, $a \in A^{[n+1]}$.

Wir sehen $q_r[g]$ als „zugänglich“ an und fragen nach einem möglichst guten nur von n und r abhängendem Faktor $\gamma = \gamma(n, r)$, mit dem die Ungleichung

$$p[g - P_n[g]] \leq \gamma q_r[g]$$

für alle $g \in C^r$ erfüllt ist.

Die Differenz $g - P_n[g]$ nimmt den Wert $\delta := \delta_n[g] = p[g - P_n[g]]$ an mindestens $n + 2$ Stellen $X_0, X_1, \dots, X_{n+1}$ in $[-1, 1]$ mit wechselndem Vorzeichen wirklich an; und diese Eigenschaft charakterisiert $P_n[g]$. Die Folge $Y(\delta) = \pm(\delta, -\delta, \dots, (-1)^{n+1}\delta)$ ist eine oszillierende Folge. Wir sagen in diesem Fall kurz, daß „ g die Verteilung $(X, Y(\delta))$ zuläßt“. Es gilt

Hilfssatz 5. Ist $\gamma = \gamma(n, r)$ eine nur von n und r abhängige Funktion mit der Eigenschaft, daß die Ungleichung $\delta_n[g] \leq \gamma q_r[g]$ für alle Funktionen $g \in C^r$ erfüllt ist, die eine Verteilung $(X, Y(\delta_n[g]))$ zulassen mit $X_0 = 1, X_{n+1} = -1$, dann ist die Ungleichung auch für alle übrigen Funktionen aus C^r erfüllt.

Beweisskizze. Läßt $g(x)$ die Verteilung $(X, Y(\delta_n[g]))$ mit $X = (X_0, X_1, \dots, X_{n+1})$ aus $[-1, 1]$ zu, dann betrachte man etwa den Koordinatenwechsel $x \rightarrow \xi$

$$\xi = \left(x - \frac{X_0 + X_{n+1}}{2} \right) \left(\frac{2}{X_0 - X_{n+1}} \right),$$

der die Übergänge $g(x) \rightarrow h(\xi) = g\left(\frac{1}{2}(\xi(X_0 - X_{n+1}) + X_0 + X_{n+1})\right)$ und $P_n[g] \rightarrow \bar{P}_n = P_n[h]$ nach sich zieht.

B) Die Stellenverteilung $X = (X_0, X_1, \dots, X_k)$ mit $X_i = \cos \frac{i\pi}{k}$, $i = 0, 1, \dots, k$, heißt die *k-te Tschebyscheff-Verteilung*. Die X_i sind die relativen Extremstellen des Tschebyscheff-Polynoms vom Grade k . Die relativen Extremwerte dieser Polynome haben alle den gleichen absoluten Betrag. Es folgt also aus den Sätzen 1, 2 und 3 sofort

Satz 4. Die *k-te Tschebyscheff-Verteilung* minimisiert den Betrag der zur oszillierenden Folge $Y(\delta) = \pm(\delta, -\delta, \pm \dots, (-1)^k \delta)$ gehörenden *k-ten Steigung*, und es gibt keine andere Stellenverteilung aus $[-1, 1]$, für die der absolute Betrag der Steigung $S(X, Y(\delta))$ ein relatives Minimum annimmt.

Bemerkung. In diesem Spezialfall der oszillierenden „Tschebyscheff-Folgen“ $Y(\delta)$ kann man die Eindeutigkeit der den Betrag von $S(X, Y(\delta))$ minimisierenden Stellenverteilung auch mit Hilfe des Theorems von P. L. Tschebyscheff, also Satz 4 ausschließlich mit den Mitteln der reellen Analysis beweisen.

C) Wir wenden uns jetzt dem *Approximationsproblem* $\mathfrak{A}(r-1, r)$ zu. Nach dem Mittelwertsatz der Differentialrechnung gilt für $g \in C^r$ $q_r[g] \geq r! |S[X, g]|$ für irgendeine $(r+1)$ -stellige Verteilung X aus $[-1, 1]$. Ferner gilt $S[X, g] = S(X, Y(\delta_{r-1}[g]))$, falls g die Verteilung $(X, Y(\delta_{r-1}[g]))$ zuläßt. Da für die Tschebyscheff-Verteilung $\bar{X}$ die Gleichung $|S(\bar{X}, Y(\delta))| = 2^{r-1} \delta$ gilt, folgt nun aus Satz 6 und Hilfssatz 5 für alle $g \in C^r$ die Ungleichung

$$(V) \quad \delta_{r-1}[g] \leq \frac{1}{2^{r-1} r!} q_r[g].$$

Es ist also $\gamma = \frac{1}{2^{r-1} r!}$ ein nur von r abhängiger Abschätzungsfaktor. Für das Approximationsproblem $\mathfrak{A}(r-1, r)$ ist γ der Vergleichsfaktor; denn für g mit $g(x) = x^r$ steht in (V) das Gleichheitszeichen. Wir können also den folgenden Satz aussprechen:

Satz 5 [S. N. Bernstein]. Der größte bei der Approximation einer in $[-1, 1]$ reellwertigen, r -mal differenzierbaren Funktion f mit beschränkter r -ter Ableitung, $r \geq 1$, durch Polynome höchstens vom Grade $r-1$ auftretende absolute Fehler $\delta_{r-1}[f]$ genügt der ohne weitere Annahmen über f keiner Verschärfung mehr fähigen Ungleichung

$$\delta_{r-1}[f] \leq \frac{1}{2^{r-1} r!} q_r[f].$$

Bemerkung. Beim Beweis dieses Satzes sind wir wesentlich anders vorgegangen als S. N. Bernstein. S. N. Bernstein hat die Klasse derjenigen Funktionen f aus C^r betrachtet, deren r -te Ableitung $f^{(r)}$ die Bedingung

$$p[f^{(r)}] \leq M$$

bei festem $M > 0$ erfüllt, und den Satz bewiesen, daß aus

$$|f^{(r)}(x)| \leq g^{(r)}(x), \quad x \in [-1, 1],$$

stets

$$\delta_{r-1}[f] \leq \delta_{r-1}[g]$$

folgt. Mit $g(x) = \frac{M}{r!} x^r$ erhält man hieraus Satz 5. Wir haben demgegenüber die Teilmenge C_δ^r derjenigen Funktionen aus C^r betrachtet, für die $\delta_{r-1}[f]$ gleich dem festen Wert δ ist, und haben die Frage gestellt, welche Verteilung $(\bar{X}, Y(\delta))$ eine Funktion f aus C_δ^r zulassen muß, damit der absolute Betrag von $S(X, Y(\delta))$ für $X = \bar{X}$ minimal wird (Satz 4).

D) Unser Vorgehen hat dabei den Vorteil, daß wir auch noch für das *Approximationsproblem* $\mathfrak{A}(r, r)$ einen guten Abschätzungsfaktor erhalten. Diesen wollen wir nun herleiten. Jetzt haben wir durch Polynome vom Grade $n = r$ zu approximieren. Mit $\delta = \delta_r[g]$ gilt $2^r \delta = |S(\bar{X}, Y(\delta))|$. Wegen der Rekursion (III) und des Mittelwertsatzes gilt ferner

$$S[X, g] \leq \max (|S[X^{[n+1]}, g]|, |S[X^{[0]}, g]|) \leq \frac{q_r[g]}{r!},$$

so daß nun für alle $g \in C^r$ folgt $\delta_r[g] \leq \frac{1}{2^r r!} q_r[g]$.

Es gilt also

Satz 6: Der größte bei der Approximation einer in $[-1, 1]$ reellwertigen, r -mal differenzierbaren Funktion f mit beschränkter r -ter Ableitung durch Polynome vom Grade r auftretende absolute Fehler $\delta = \delta_r[f]$ genügt der Ungleichung

$$\delta \leq \frac{1}{2^r r!} q_r[f].$$

Man erkennt an der Funktion $g = |x| - \frac{1}{2}$, daß die Ungleichung des Satzes 8 wenigstens im Falle $n = r = 1$ scharf ist.

Die bisher beste Abschätzung für diesen Fall war (vgl. [4])

$$\delta \leq \frac{1}{2^{r-1} r!} q_r[f].$$

Literatur

- [1] N. I. Achieser: Vorlesungen über Approximationstheorie, Berlin, 1953.
- [2] G. Aumann: Über den Vergleichsfaktor bei linearen Approximationsproblemen, Numerische Mathematik 5, 1963, S. 68–72.
- [3] D. Ehresmann: Polynome minimaler Steigung, Diss., TH München 1965.
- [4] G. Meinardus: Approximation von Funktionen und ihre numerische Behandlung, Berlin, 1964.
- [5] I. P. Natanson: Konstruktive Funktionentheorie. Deutsche Übersetzung von K. Bögel, Berlin, 1955.
- [6] N. E. Nörlund: Vorlesungen über Differenzenrechnung, New York, 1954.