JOURNAL OF APPROXIMATION THEORY **10**, 41–48 (1974)

On Polya's Algorithm for Regulated Functions

PETER W. LINDSTROM AND BERNARD H. ROSMAN

Department of Mathematics, Boston University, Boston, Massachusetts 02215

Communicated by John R. Rice

Received September 13, 1971

1. INTRODUCTION AND PRELIMINARIES

Let f be a regulated function on $[a, b]$ (i.e., f is a uniform limit of step functions on $[a, b]$). We want to approximate f in the Chebyshev sense by linear combinations of functions which form an extended Chebyshev system of order three on $[a, b]$ [2, Chapter 1]. For a certain subclass of regulated functions we give a new characterization for best approximation and use this to prove a convergence result for Polya's Algorithm. Our convergence result lends support to conjectures of Descloux [1] concerning piecewise analytic functions. The extension of our results to Chebyshev systems in general and to wider subclasses of regulated functions seems more difficult and is under investigation. Define

$$\|f\| = \sup\{|f(x)|;\ a \leqslant x \leqslant b\},$$
$$f_+(x) = \lim_{u\to x} \sup f(u);\ f_-(x) = \lim_{u\to x} \inf f(u).$$

If g is in F, the family of approximating functions, define

$$E_\pm(x; g, f) = f_\pm(x) - g(x),$$
$$e(g) = \max(\| E_+ \|, \| E_- \|).$$

A function g^* in F is said to be a best Chebyshev approximation to f (notation: $g^* \in (b.a.)_f$) if $e(g^*) \leqslant e(g)$ for all g in F. As in [5] we say that $\bar{x} \in (a, b)$ is a straddle point if for some $g \in F$

$$E_+(\bar{x}; g, f) = -E_-(\bar{x}; g, f) = e(g), \tag{1}$$

In addition $\bar{x}$ is a $[-, +]$ point relative to g in F if

(i) $\bar{x}$ is a straddle point,

(ii) $f_-(\bar{x}) = \lim\limits_{\substack{x\to\bar{x}\\ x<\bar{x}}} f(x);\ f_+(\bar{x}) = \lim\limits_{\substack{x\to\bar{x}\\ x>\bar{x}}} f(x).$

It is a $[+, -]$ point if (i) holds and (ii) holds with f_+ and f_- unchanged.

Using this, the equioscillation concept is generalized as follows. We say that the error curves $E_{\pm}(x; g, f)$ alternate n times on $[a, b]$ counting multiplicity if there exists a sequence of $n + 1$ points on $[a, b]$, $x_1 \leqslant x_2 \leqslant \cdots \leqslant x_{n+1}$ such that no point occurs more than twice, and if $x_i = x_{i+1}$ for some i then x_i is a straddle point. As usual we require that at each point of the sequence the error equal $e(g)$ in magnitude and that the sign of the error alternate at consecutive points of the sequence so that a straddle point is either a $[-, +]$ or $[+, -]$ point depending on the sign pattern of the error and if a nonstraddle point x_i is a $+$ point ($E_+(x_i) = e(g)$) then x_{i+1} is a $-$ point ($E_-(x_i) = -e(g)$). Such an approximation g is said to be an equioscillator.

It is shown in [5] that if f has exactly one straddle point in (a, b) then it has one and only one equioscillator (although f may have many best approximations). In the sequel we prove the convergence of Polya's Algorithm to the equioscillator for functions of this type.

2. Characterization Results

The following lemma is a direct consequence of a result in Karlin and Studden [2, p. 24].

Lemma. *If $u_0, \ldots, u_n$ form an extended Chebyshev system of order* 3 *then every nontrivial u-polynomial ($u = \sum_{i=0}^{n} a_i u_i$; a_i real not all zero) has, at most, n zeroes, where zeroes of multiplicity i are counted i times, $i = 1, 2$, and zeroes of multiplicity at least three are counted three times.*

This lemma is used to prove the following result which characterizes the equioscillator in terms of derivatives at the straddle point. We remark that this result is in the same spirit as results in Karlin and Studden [2, pp. 295–298] in connection with generalized Markoff-Bernstein inequalities.

Theorem 1. *Let $f(x)$ be a regulated function on $[a, b]$ with exactly one straddle point at $c \in (a, b)$. Assume that f has continuous one-sided first derivatives in a neighborhood of c. Let u^* be the equioscillator for f, where f is approximated by u-polynomials: $u(x) = \sum_{i=0}^{n} a_i u_i(x)$. Then if $u_0, \ldots, u_n$ form an extended Chebyshev system of order three, u^* is characterized by the following inequalities. If c is a $[-, +]$ point,*

$$\max[u^{*\prime}(c) - f_+'(c), u^{*\prime}(c) - f_-'(c)] \geqslant \max[u'(c) - f_+'(c), u'(c) - f_-'(c)] \geqslant 0; \tag{2}$$

If c is a $[+, -]$ *point*,

$$\max[f_+'(c) - u^{*\prime}(c), f_-'(c) - u^{*\prime}(c)] \geqslant \max[f_+'(c) - u'(c), f_-'(c) - u'(c)] \geqslant 0, \tag{3}$$

for all $u \in (\text{b.a.})_f$. *Furthermore, equality holds in* (2) *or* (3) *if and only if* $u \equiv u^*$.

Proof. We prove only (2) since (3) follows similarly. The right hand inequality of (2) is clear geometrically from the derivative assumptions on f and the fact that $u \in (\text{b.a.})_f$. For the left side, assume that for some u the inequality is false. Then it follows that $u'(c) > u^{*\prime}(c)$. Now from Theorem 1 we have the set of points $a \leqslant x_1 < x_2 < \cdots < x_k < c < x_{k+3} < \cdots < x_{n+2} \leqslant b$ on which the error for u^* equioscillates with magnitude $e(u^*)$. Since $u, u^* \in (\text{b.a.})_f$ and $u(c) = u^*(c)$, we conclude that $u - u^*$ has at least $k + 1$ zeroes in $[a, c]$ and at least $n - k$ zeroes in $[c, b]$. This implies by the lemma that $u \equiv u^*$ which contradicts $u'(c) > u^{*\prime}(c)$. If x_1 is c, the proof is the same.

It remains to examine the case of equality on the left side of (2); that is, $u'(c) = u^{*\prime}(c)$. Let $d = u - u^*$. Then the only case we need consider is where $d(c + \epsilon) < 0$; $d(c - \epsilon) > 0$ for ϵ small and positive. It follows that $d''(c) = 0$. Hence d has a triple zero at c. By counting the remaining zeroes as before we get $d \equiv 0$ or $u \equiv u^*$. This concludes the proof.

The following Corollary is needed below.

COROLLARY 1. *With the hypothesis of Theorem* 1 *we have the following inequalities*:

If c is a $[-, +]$ *point*,

$$\begin{aligned} u^{*\prime}(c) - f_+'(c) &\geqslant u'(c) - f_+'(c) \geqslant 0, \\ u^{*\prime}(c) - f_-'(c) &\geqslant u'(c) - f_-'(c) \geqslant 0. \end{aligned} \tag{4}$$

If c is a $[+, -]$ *point*,

$$\begin{aligned} f_+'(c) - u^{*\prime}(c) &\geqslant f_+'(c) - u'(c) \geqslant 0, \\ f_-'(c) - u^{*\prime}(c) &\geqslant f_-'(c) - u'(c) \geqslant 0, \end{aligned} \tag{5}$$

with equality in (4) *or* (5) *if and only if* $u \equiv u^*$.

Proof. The proof follows from Theorem 1 and its proof and the fact (implied by the theorem) that $u^{*\prime}(c) \geqslant u'(c)$ with equality if and only if $u \equiv u^*$.

The following property of $(\text{b.a.})_f$ is also needed and is of independent interest.

THEOREM 2. *Let u be any relative interior point of* $(\text{b.a.})_f$. *Then the error curve for u, $u(x) - f(x)$ attains its extrema, $\pm E^*$, only at the straddle point c. (We say that u has no critical points other than c, where $|u(c) - f_+(c)| = E^*$; $|u(c) - f_-(c)| = E^*$).*

To prove the theorem, we use the following lemma.

LEMMA 1. *Let X be the compact set $[a, b] - [c - \delta, c + \delta]$, where $0 < \delta < \min(c - a, b - c)$. Let $c_0 = [f_+(c) + f_-(c)]/2$ and define*

$$S(c_0) = \left\{ a \cdot g = \sum_{i=1}^{n} a_i g_i \mid a \cdot g(c) = c_0 \right\}.$$

Then f has a unique b.a. *from $S(c_0)$, where f is any regulated function with no straddle points on X with respect to approximation by linear combinations of $g_1, \ldots, g_n$.*

Proof. First assume $c_0 = 0$. $S(0)$ is a linear space of dimension $n - 1$ and is a Chebyshev set on X, for if $a \cdot g$ has $n - 1$ zeroes on X, it has n zeroes on $[a, b]$; hence $a \cdot g \equiv 0$ since $g_1, \ldots, g_n$ form a Chebyshev set on $[a, b]$. It follows by results in [5] that f has a unique b.a. from $S(0)$. If $c_0 \neq 0$, then $S(c_0)$ is an affine set of dimension $n - 1$ and can be written as $S(c_0) = S(0) + a_0 \cdot g$, where $a_0 \cdot g$ is a fixed approximant such that $a_0 \cdot g(c) = c_0$ [2]. Define $\bar{f}(x) = f(x) - c_0$. Then we have

$$\begin{aligned} \min_{u \in S(c_0)} \| f - u \| &= \min_{\bar{u} \in S(0)} \| \bar{f} + c_0 - (\bar{u} + a_0 \cdot g) \| \\ &= \min_{\bar{u} \in S(0)} \| \bar{f} + c_0 - a_0 \cdot g - \bar{u} \|, \end{aligned}$$

which is attained by a unique element of $S(0)$. Hence a unique b.a. to f on X from the class $S(c_0)$ exists.

Proof of Theorem 2. It follows by a short convexity argument that all relative interior elements of $(\text{b.a.})_f$ coincide at their critical points. Call these points $c, x_1, \ldots, x_k$. Clearly $k \leqslant n - 2$. The proof will be completed by constructing a b.a. which has only c as a critical point. Assume for concreteness that c is a $[-, +]$ point. By Corollary 1 we can choose a relative interior point $\tilde{u} = \tilde{a} \cdot g$ near u^* such that $\tilde{u}'(x) - f'(x)$ is positive in small one-sided neighborhoods of c.

Now let $B = \{a \cdot g \mid \| a \cdot g \| \leqslant 2 \| f \| \}$ and let $v = b \cdot g \in B$. Then $\| v' \|$ is uniformly bounded and it follows that there exists λ such that for any $v \in B$:

$$\begin{aligned} \lambda v'(c) + \tilde{u}'(c) - f_+'(c) > 0; \\ \lambda v'(c) + \tilde{u}'(c) - f_-'(c) > 0. \end{aligned} \tag{6}$$

From (6) it follows that $|\tilde{u}(x) - f(x) + \lambda v(x)|$ attains the value E^* in $[c - \delta, c + \delta]$ only at the straddle point c, where $\delta > 0$ is chosen sufficiently small so that $x_1, ..., x_k \notin [c - \delta, c + \delta]$. Let $X = [a, b] - [c - \delta, c + \delta]$. By Lemma 1, f has a unique b.a. on X from $S(c_0)$, call it $u_0 = b_0 \cdot g$. The uniqueness implies that u_0 could not be a relative interior point of $(\text{b.a.})_f$, since all these points yield the same error norm on X. Now take $\tilde{b} \cdot g = \tilde{a} \cdot g + \alpha(b_0 - \tilde{a}) \cdot g$ for $\alpha > 0$ sufficiently small such that $\alpha \leqslant \lambda$ and $\alpha(b_0 - \tilde{a}) \in B$. Hence in $[c - \delta, c + \delta]$, $|f(x) - \tilde{b} \cdot g(x)|$ attains E^* only at c and in X we have

$$|\tilde{b} \cdot g(x) - f(x)| \leqslant (1 - \alpha) |\tilde{a} \cdot g(x) - f(x)| + \alpha |b_0 \cdot g(x) - f(x)| < E^*.$$

Hence we have a b.a. with no critical points other than c so that all relative interior points of $(\text{b.a.})_f$ have critical points only at c.

3. Polya's Algorithm

The following two lemmas are used to prove the main theorem of this paper on the convergence of Polya's Algorithm.

LEMMA 2. *Let B be a compact set in R^n. If m^+, m^- are quantities such that*

(i) $m^+ < \inf\{E_+'(c, d) \mid d \in B\}$,

(ii) $m^- < \inf\{E_-'(c, d) \mid d \in B\}$,

there exists $\epsilon > 0$ such that for all $d \in B$ we have

$$\text{on} \quad [c, c + \epsilon], \qquad E_+(x, d) > E_+(c, d) + m^+(x - c), \tag{7}$$

$$\text{on} \quad [c - \epsilon, c], \qquad E_-(x, d) < E_-(c, d) + m^-(x - c). \tag{8}$$

Proof. We prove only (7), since (8) follows in the same way. Let $\alpha(d) = \sup\{x \mid x \in [c, b] \text{ and } E_+'(x_-d) > m^+\}$. By the continuity of $E_+(x, d)$, $\alpha(d) > c$. Now, $\alpha(d)$ is a lower semicontinuous function of d. For if $\alpha(d_0) > \beta$, then on $[c, \beta]$, $E_+'(x, d_0) > m^+$. But $E_+'(x, d)$ is a continuous function of x and d; hence if there exists $x_n \in [c, \beta]$ and d_n converging to d_0 such that $E_+'(x_n, d_n) \leqslant m^+$, it follows that there is an $x^* \in [c, \beta]$ such that $E_+'(x^*, d_0) \leqslant m^+$. Hence $\{d \in R^n \mid \alpha(d) > \beta\}$ is an open set and $\alpha(d)$ is lower semicontinuous. Thus $\alpha(d)$ assumes a minimum greater than c on B. An application of the mean value theorem then gives (7). It is clear that the Lemma also holds with inf replaced by sup and the inequality signs reversed.

LEMMA 3. *Define* $f(\delta) = a(1 + \delta)^p + b(1 - \delta)^p$, *where* $a > 0$, $b > 0$. *Then* $\lim_{p\to\infty} \min_{\delta \in [-1/2, 1/2]} f(\delta) = 2(ab)^{1/2}$.

Proof. Let $c = b/a$ and consider $F(\delta) = (1 + \delta)^p + c(1 - \delta)^p$. We have $F'(\delta) = p[(1 + \delta)^{p-1} - c(1 - \delta)^{p-1}]$ and $F''(\delta) = p(p - 1)[(1 + \delta)^{p-2} + c(1 - \delta)^{p-2}] > 0$ for $\delta \in [-1/2, 1/2]$. A short calculation shows that the unique minimum of $F(\delta)$ in $[-1/2, 1/2]$ is at the point $\delta = (c^{1/(p-1)} - 1)/(1 + c^{1/(p-1)})$ for p sufficiently large. Substituting this value for δ into $F(\delta)$ and letting $v = c^{1/(p-1)}$, we get that the minimum value is $(2v/(1 + v))^p + c(2/(1 + v))^p$. Taking the limit as $p \to \infty$ and using L'Hopital's rule we get $2c^{1/2}$ which gives the result for $f(\delta)$ after multiplying by a.

THEOREM 3. *If $p_n \to \infty$ then $\| a_{p_n} \cdot g - a^* \cdot g \| \to 0$, where $a_{p_n} \cdot g$ is the best L_{p_n} approximation to f and $a^* \cdot g$ is the equioscillator.*

Proof. Without loss of generality, we take $E^* = 1$, for otherwise we may divide f by E^*. (The case $E^* = 0$ is trivial.) Assume the theorem is false. Then we may assume that $\{a_p\}$ (we drop the subscript for convenience) converges to $a_0 \neq a^*$. By the Polya algorithm $a_0 \in (\text{b.a.})_f$. Near a^* we may choose a relative interior point $\bar{a}$ of $(\text{b.a.})_f$ which by Theorem 2 has the property that when $x \neq c$, $|\bar{a} \cdot g(x) - f(x)| < 1$.

Now by Lemma 2 it follows that there is a compact ball B_0 about a_0 such that for all $d \in B_0$

$$\begin{aligned} E_+(x, d) &< R_0(x - c) + E_+(c, d), \qquad \text{on} \quad [c, c + \epsilon] \\ E_-(x, d) &> L_0(x - c) + E_-(c, d), \qquad \text{on} \quad [c - \epsilon, c] \end{aligned}$$

where R_0, L_0 are chosen to satisfy (i) and (ii) in Lemma 2 (with inf replaced by sup). Note that (7) and (8) hold for any smaller (positive) ϵ and for any ball contained in B_0. Set $E_+(c, d) = 1 + \delta_0$; $E_-(c, d) = -1 + \delta_0$, where $|\delta_0| = 0$ if $d \in (\text{b.a.})_g$ and otherwise $|\delta_0|$ is small (for balls B_0 close to a_0). We have

$$\begin{aligned} &\int_a^b | d \cdot g(x) - f(x)|^p \, dx \\ &\quad \geqslant \int_{c-\epsilon}^{c+\epsilon} | d \cdot g(x) - f(x)|^p \, dx \\ &\quad > \int_{c-\epsilon}^{c} (L_0(x - c) + 1 + \delta_0)^p \, dx + \int_c^{c+\epsilon} (1 - \delta_0 - R_0(x - c))^p \, dx \\ &\quad = \frac{1}{p+1} \Big[\frac{1}{L_0} (1 + \delta_0)^{p+1} + \frac{1}{R_0} (1 - \delta_0)^{p+1} - \frac{1}{L_0} (1 + \delta_0 - L_0\epsilon)^{p+1} \\ &\qquad - \frac{1}{R_0} (1 - \delta_0 - R_0\epsilon)^{p+1} \Big]. \end{aligned}$$

Now pick B_0 about a_0 such that $| 1 + \delta_0 - L_0\epsilon | < \eta$ and $| 1 - \delta_0 - R_0\epsilon | < \eta$, where $\eta < 1$. We then have

$$(p + 1)\int_a^b | d \cdot g(x) - f(x)|^p \, dx > \frac{1}{L_0}(1 + \delta_0)^{p+1} + \frac{1}{R_0}(1 - \delta_0)^{p+1} - 2\eta^{p+1}.$$

Now using Lemma 3, given $\epsilon' > 0$ and small η, there exists $P(\epsilon')$, such that for $p \geqslant P(\epsilon')$ we have

$$(p + 1)\int_a^b | d \cdot g(x) - f(x)|^p \, dx > \frac{2}{L_0 R_0} - \epsilon' - \eta^{p+1}. \tag{*}$$

Next we perform a similar calculation for points $\bar{\alpha}$ in a ball $\bar{B}$ about $\bar{a}$. We have again by use of Lemma 2

$$\begin{aligned} E_+(x, \bar{d}) &> \bar{R}(x - c) + E_+(c, \bar{d}), \qquad \text{on} \quad [c, c + \epsilon] \\ E_-(x, \bar{d}) &< \bar{L}(x - c) + E_-(c, \bar{d}), \qquad \text{on} \quad [c - \epsilon, c], \end{aligned}$$

where $\bar{R}$, $\bar{L}$ satisfy (i) and (ii) in the lemma. Without loss of generality we may assume that the interval $I = [c - \epsilon, c + \epsilon]$ coincides with the interval of the previous paragraph. We get after integration as before with $E_+(c, \bar{d}) = 1 + \bar{\delta}$; $E_-(c, \bar{d}) = -1 + \bar{\delta}$;

$$\begin{aligned} (p + 1)&\int_a^b | \bar{d} \cdot g(x) - f(x)|^p \, dx \\ &\leqslant \frac{1}{\bar{L}}(1 + \bar{\delta})^{p+1} + \frac{1}{\bar{R}}(1 - \bar{\delta})^{p+1} - \frac{1}{\bar{L}}(1 + \bar{\delta} - \bar{L}\epsilon)^{p+1} \\ &\quad - \frac{1}{\bar{R}}(1 - \bar{\delta} - \bar{R}\epsilon)^{p+1} + \int_{[a,b]-I} | \bar{d} \cdot g(x) - f(x)|^p \, dx \\ &\leqslant \frac{1}{\bar{L}}(1 + \bar{\delta})^{p+1} + \frac{1}{\bar{R}}(1 - \bar{\delta})^{p+1} + \int_{[a,b]-I} | \bar{d} \cdot g(x) - f(x)|^p \, dx. \end{aligned}$$

Now let $M(\bar{d}) = \max_{[a,b]-I} | \bar{d} \cdot g(x) - f(x)|$. By choosing $\bar{B}$ small enough and using Theorem 2, we can find $\eta^* < 1$ such that $M(\bar{d}) \leqslant \eta^* < 1$, for all $\bar{d} \in \bar{B}$. Hence

$$\begin{aligned} (p + 1)&\int_a^b | \bar{d} \cdot g(x) - f(x)|^p \, dx \\ &\leqslant \frac{1}{\bar{L}}(1 + \bar{\delta})^{p+1} + \frac{1}{\bar{R}}(1 - \bar{\delta})^{p+1} + (b - a)(\eta^*)^p. \end{aligned}$$

Using Lemma 3 and continuity, it follows that there exists $\bar{\delta}$ and $P(\epsilon'')$ such that for all $p \geqslant P(\epsilon'')$ we have

$$\frac{1}{\bar{L}}(1+\bar{\delta})^{p+1}+\frac{1}{\bar{R}}(1-\bar{\delta})^{p+1}+(b-a)(\eta^*)^p$$
$$\leqslant \frac{2}{(\bar{L}\bar{R})^{1/2}}+\epsilon''+(b-a)(\eta^*)^p.$$

Hence by associating $\bar{\delta}$ with an element of $\bar{B}$, it follows that for $p \geqslant P(\epsilon'')$ there exists $\bar{a}_p \in \bar{B}$ such that

$$(p+1)\,\|\bar{a}_p\cdot g-f\|^p \leqslant \frac{2}{(\bar{L}\bar{R})^{1/2}}+\epsilon''+(b-a)(\eta^*)^p. \qquad (**)$$

To finish the proof we compare (*) with (**) using the fact from Corollary 1 that $\bar{L}\bar{R} > L_0R_0$. Specifically, choose ϵ', ϵ'' such that $2/(\bar{L}\bar{R})^{1/2}+\epsilon'' < 2/(L_0R_0)^{1/2}-\epsilon'$ and choose $p \geqslant \max[p(\epsilon'), p(\epsilon'')]$ such that

$$\epsilon''+(b-a)(\eta^*)^p+\epsilon'+2\eta^{p+1} < \frac{2}{(L_0R_0)^{1/2}}-\frac{2}{(\bar{L}\bar{R})^{1/2}}.$$

This means we have found an $\bar{a}_p$ for p in this range such that $\|\bar{a}_p\cdot g-f\|_p \leqslant \|a_p\cdot g-f\|_p$ which is a contradiction by uniqueness of best L_p approximation. This proves the theorem.

Finally we remark that if f has no straddle points the best approximation is unique [5], and hence Polya's Algorithm converges to the equioscillator.

References

1. J. Descloux, Approximation stricte, "Funktionalanalysis, Approximationstheorie, Numerische Mathematik" (Oberwolfach, 1965), pp. 18–23. Birkhauser, Basel, 1967.
2. S. Karlin and W. Studden, "Tchebycheff Systems with Applications in Analysis and Statistics," Interscience, New York, 1966.
3. J. Rice, "The Approximation of Functions," Vol. 1, Addison-Wesley, Reading, MA, 1964.
4. R. Rockafellar, "Convex Analysis," Princeton, Princeton, New Jersey, 1970.
5. B. H. Rosman and P. D. Rosenbaum, Chebyshev approximation of regulated functions by unisolvent families of functions, *J. Approximation Theory*, **4** (1971), 113–120.