JOURNAL OF APPROXIMATION THEORY 53, 335–349 (1988)

The Polya Algorithm on Cylindrical Sets

ROBERT HUOTARI, DAVID LEGG, AND DOUGLAS TOWNSEND

Department of Mathematical Sciences, Indiana University–Purdue University, Fort Wayne, Indiana 46805, U.S.A.

Communicated by E. W. Cheney

Received January 9, 1986

We define the property "E-cylindrical," which relates to a subset of $\mathbb{R}^m$ certain directed cylinders. We investigate some of the consequences of this definition, showing, for example, that polyhedral convex sets and smooth, rotund convex bodies are E-cylindrical. Suppose X is a finite set, F is the set of all real-valued functions on X, $f \in F$, and $K \subset F$ is closed, convex, and E-cylindrical. For $1 < p < \infty$, let f_p be the best l_p-approximation to f by elements of K. We show that $\lim_{p \to \infty} f_p$ exists. We give an example to show that $\{f_p\}$ may fail to converge if X is countably infinite. We discuss the relationship between discrete (l_p) and continuous (L_p) approximation. © 1988 Academic Press, Inc.

1. INTRODUCTION

Let X be a subset of $\mathbb{R}^n$ consisting of m points and let F consist of all real-valued functions on X. For each g in F, define the l_p-norms by

$$\|g\|_p = \left(\sum_{x \in X} |g(x)|^p \right)^{1/p}, \qquad 1 \leqslant p < \infty,$$

and

$$\|g\|_\infty = \max_{x \in X} (|g(x)|).$$

Suppose K is a closed (in the l_2-topology) convex subset of F and let $f \in F$ be fixed. For $1 \leqslant p \leqslant \infty$, a function g in F is a *best l_p-approximation to f by elements of K* if

$$\|f - g\|_p \leqslant \|f - h\|_p, \qquad h \in K.$$

Since the l_p-norm is strictly convex for $1 < p < \infty$, there exists a unique best l_p-approximation, f_p, to f by elements of K.

0021-9045/88 \$3.00

The Polya algorithm is the construction of a best l_∞-approximation as the limit of the f_p as $p \to \infty$. Descloux [1] showed that this limit exists for every f in F when K is a subspace of F. In the present paper, we generalize Descloux's theorem to a certain class of closed convex subsets of F. This class contains all closed convex bodies which are smooth and rotund. It also contains all closed polyhedral convex sets and hence, for example, the set consisting of all nondecreasing functions on X and the set consisting of all convex functions on X.

Next, we give an example to show that in the case of approximation by convex functions, the Polya algorithm does not always converge when X is countably infinite.

Finally, if $f \in C[0, 1]$, we show that, for $1 < p < \infty$, f_p is the limit of a sequence of best discrete convex approximations. In addition to its usefulness in calculation, this fact may be a first step in showing that the Polya algorithm converges when $f \in C[0, 1]$ is being approximated by continuous convex functions.

2. Cylindrical Sets

The property described in this section has been highlighted because it appears to be the most general which will work in our proof of Descloux's theorem. However, it is novel and geometrically compelling, so it may be of independent interest.

Since $(F, \|\cdot\|_p)$ is congruent to $(\mathbb{R}^m, \|\cdot\|_p)$, any discussion of subsets of $\mathbb{R}^m$ is equivalent to a discussion of sets of functions on X, so for our definitions we take the geometric point of view. For any $\mathbf{z}$ in $\mathbb{R}^m$ let $(z_1, ..., z_m)$ be the m-tuple of components of $\mathbf{z}$ and let $\|\mathbf{z}\|_p = (\sum_{i=1}^m |z_i|)^{1/p}$ and $\|\mathbf{z}\|_\infty = \max_{1 \leqslant i \leqslant m} |z_i|$.

If $\mathbf{x}, \mathbf{v} \in \mathbb{R}^m$ and $A \subset \mathbb{R}^m$, let $d(\mathbf{x}, A) = \inf\{\|\mathbf{x} - \mathbf{y}\|_\infty : \mathbf{y} \in A\}$, let $N(A, \delta) = \{\mathbf{z} \in \mathbb{R}^m : d(\mathbf{z}, A) < \delta\}$, and let $L(\mathbf{x}, \mathbf{v})$ be the straight line in $\mathbb{R}^m$ which contains $\mathbf{x}$ and is parallel to the line containing $\mathbf{0}$ and $\mathbf{v}$. (We will, on occasion, abuse the notation by regarding $\mathbf{v}$ as the vector represented by the directed line segment $\mathbf{0v}$.) A subset A of $\mathbb{R}^m$ is said to be $\mathbf{v}$*-cylindrical at* $\mathbf{x}$ if for any $\varepsilon > 0$ there exists $\delta = (\mathbf{x}, \varepsilon) > 0$ such that $d(\mathbf{x}, L(\mathbf{y}, \mathbf{v}) \cap A) < \varepsilon$ whenever $\mathbf{y} \in A$ and $d(\mathbf{y}, L(\mathbf{x}, \mathbf{v})) < \delta$. The set A is said to be $\mathbf{v}$*-cylindrical* if it is $\mathbf{v}$-cylindrical at every $\mathbf{x}$ in $\bar{A}$, the closure of A. Let $E = \{\mathbf{e}_1, ..., \mathbf{e}_m\}$ be the standard basis of $\mathbb{R}^m$. A subset A of $\mathbb{R}^m$ is said to be *E-cylindrical* if A is $\mathbf{e}_i$-cylindrical for $1 \leqslant i \leqslant m$. The set A is said to be *cylindrical in every direction* if A is $\mathbf{v}$-cylindrical for every $\mathbf{v}$ in $\mathbb{R}^m$.

To illustrate the geometric origin of our definition of $\mathbf{v}$-cylindrical, we suppose that $B \subset \mathbb{R}^m$ is a convex body (i.e., B is convex and the interior of

B is nonempty), $\mathbf{x} \in B$, $\mathbf{v} \in \mathbb{R}^m$, and $\delta > 0$. Let $N_B(L(\mathbf{x}, \mathbf{v}), \delta)$ denote the cylinder

$$\{L(\mathbf{y}, \mathbf{v}): \mathbf{y} \in B \cap N(L(\mathbf{x}, \mathbf{v}), \delta)\}.$$

A $\mathbf{v}$-*cylinder in* B is a set of the form

$$C(\mathbf{x}, \mathbf{v}, \delta, \alpha, \beta) = N_B(L(\mathbf{x}, \mathbf{v}), \delta) \cap \{\mathbf{z} \in \mathbb{R}^m: \alpha < \mathbf{v} \cdot \mathbf{z} < \beta\},$$

which is contained in B. The convex body B is cylindrical in the direction $\mathbf{v}$ if and only if, for any $\mathbf{x}$ in $\bar{B}$ and $\varepsilon > 0$, there is a $\mathbf{v}$-cylinder, $C = C(\mathbf{x}, \mathbf{v}, \delta, \alpha, \beta)$, in B such that $d(\mathbf{x}, C) < \varepsilon$.

If, in the same context, B is bounded, there is another interpretation: If H is a hyperplane which supports B and is orthogonal to $\mathbf{v}$, and π is the orthogonal projection of $\bar{B}$ onto H, let $d(y) = \sup\{\|x - y\|_\infty: x \in \bar{B} \cap \pi^{-1}(y)\}$, for each y in $\pi(\bar{B})$, and define $c(y)$ similarly, with "sup" replaced by "inf." Then B is $\mathbf{v}$-cylindrical if and only if each of d and c is a continuous real-valued function on $\pi(\bar{B})$.

If $K \subset \mathbb{R}^2$ is convex, then K is cylindrical in every direction. Indeed, if $\mathbf{x} \in K$ and there exists $\mathbf{w}^1$ in the set $\{(y_1, y_2): y_1 < x_1\} \cap K$ such that the slope, m_1, of the line containing $\mathbf{w}^1$ and $\mathbf{x}$ is nonzero, let $\delta_1 = \min(\varepsilon, \varepsilon/|m_1|)$. If there is no such $\mathbf{w}^1$, let $\delta_1 = \varepsilon$. Similarly, let $\delta_2 = \varepsilon$ or, if there exists $\mathbf{w}^2$ in $\{(y_1, y_2): y_1 > x_1\} \cap K$, so that the line containing $\mathbf{w}^2$ and $\mathbf{x}$ has slope $m_2 \neq 0$, let $\delta_2 = \min(\varepsilon, \varepsilon/|m_2|)$. Then $\delta(\mathbf{x}, \varepsilon) = \min(\delta_1, \delta_2)$ satisfies the requirement for K to be $(0, 1)$-cylindrical at $\mathbf{x}$. Since convexity is invariant under rotation, K is cylindrical in every direction.

If $m = 3$, however, there exist convex sets which are not cylindrical in every direction. For example, if $K^* = \bigcup\{\lambda A: 0 \leqslant \lambda \leqslant 1\}$, where $A = \{(x_1, x_2, x_3): (x_1 - 1)^2 + x_2^2 \leqslant 1 \text{ and } x_3 = 1\}$, then K^* fails to be $(0, 0, 1)$-cylindrical. K^* can be smoothed to provide an example of a smooth convex set which is not cylindrical in every direction and two copies of K^* can be pasted base-to-base to provide an absolutely convex counterexample.

(2.1) Lemma. *Suppose* $A_1, A_2 \subset \mathbb{R}^m$ *and each is* $\mathbf{v}$-*cylindrical. Then* $A_1 \cap A_2$ *is* $\mathbf{v}$-*cylindrical.*

Proof. If $\mathbf{x} \in A_1 \cap A_2$ and $\varepsilon > 0$ is given, let $\delta = \min(\delta_1, \delta_2)$, where, for $i = 1, 2$, $d(\mathbf{x}, L(\mathbf{y}, \mathbf{v}) \cap A_i) < \varepsilon$ whenever $\mathbf{y} \in A_i \cap N(L(\mathbf{x}, \mathbf{v}), \delta_i)$. Suppose $\mathbf{y} \in A_1 \cap A_2 \cap N(L(\mathbf{x}, \mathbf{v}), \delta)$. Then there must exist $\mathbf{y}^i$ in $A_i \cap L(\mathbf{y}, \mathbf{v}) \cap N(\{\mathbf{x}\}, \delta)$, $i = 1, 2$. If $\mathbf{y}^1$ is between $\mathbf{y}$ and $\mathbf{y}^2$ on $L(\mathbf{y}, \mathbf{v})$, the convexity of A_2 implies that $\mathbf{y}^1 \in A_2$. On the other hand, if $\mathbf{y}^2$ is between $\mathbf{y}$ and $\mathbf{y}^1$, then $\mathbf{y}^2 \in A_1$. Thus $\mathbf{y}^1 \in A_1 \cap A_2$ or $\mathbf{y}^2 \in A_1 \cap A_2$, so there is a point on $L(\mathbf{y}, \mathbf{v}) \cap A_1 \cap A_2$ sufficiently close to $\mathbf{x}$. This establishes Lemma (2.1).

If $A \subset \mathbb{R}^m$, let A° denote the interior of A and let $\partial A = \bar{A} - A^\circ$, the boundary of A. If $B \subset \mathbb{R}^m$ is a convex body and $\mathbf{x} \in \partial B$, then $\mathbf{x}$ is said to be a *smooth point of* B if there is exactly one hyperplane in $\mathbb{R}^m$ which supports B at $\mathbf{x}$. The set B is said to be *smooth* if every point in ∂B is a smooth point of B. The set B is said to be *rotund* if every point of ∂B is an extreme point of B.

(2.2) THEOREM. *A smooth rotund convex body is cylindrical in every direction.*

Proof. Let B be a smooth rotund convex body in $\mathbb{R}^m$ and let $\mathbf{v} \in \mathbb{R}^m$. If $\mathbf{x} \in B^\circ$, then there exists a $\mathbf{v}$-cylinder in B which contains $\mathbf{x}$ so B is clearly $\mathbf{v}$-cylindrical at $\mathbf{x}$. Suppose $\mathbf{x} \in \partial B$. Let H be the hyperplane which supports B at $\mathbf{x}$. For each $k \geqslant 1$, let $B_k = \{\mathbf{y} \in B\colon d(\mathbf{y}, H) < 1/k\}$. Then B_k is convex and $\mathbf{x}$ is a smooth point of B_k.

Suppose $\mathbf{v}$ is not parallel to H. Let $B_k(\mathbf{v}) = \bigcup\{L(\mathbf{y}, \mathbf{v})\colon \mathbf{y} \in B_k\}$. Then $B_k(\mathbf{v})$ is convex and so is the projection, $P_k(\mathbf{v})$, of $B_k(\mathbf{v})$ onto H. We claim that $\mathbf{x}$ is in the relative interior of $P_k(\mathbf{v})$ (i.e., there exists an open set G such that $\mathbf{x} \in G$ and $G \cap H \subset P_k(\mathbf{v})$). Suppose this not the case. Then $L(\mathbf{x}, \mathbf{v}) \cap B_k^\circ = \phi$ so there exists a hyperplane which supports B_k at $\mathbf{x}$ and which contains the line $L(\mathbf{x}, \mathbf{v})$. But this contradicts the fact that $\mathbf{x}$ is a smooth point of B_k and establishes the claim. Let N be a relative neighborhood of $\mathbf{x}$ in $P_k(\mathbf{v})$. Then there exists $\delta > 0$ such that

$$N(L(\mathbf{x}, \mathbf{v}), \delta) \cap P_k(\mathbf{v}) \subset N,$$

whence B is $\mathbf{v}$-cylindrical at $\mathbf{x}$.

Suppose $\mathbf{v}$ is parallel to H. If B is not $\mathbf{v}$-cylindrical at $\mathbf{x}$, then there exist $\varepsilon > 0$ and $\mathbf{y}^k$, $k \geqslant 1$, such that, for each $k \geqslant 1$, $\mathbf{y}^k \in B_k$ and $\|\mathbf{x} - \mathbf{y}^k\|_\infty \geqslant \varepsilon$. Let $\mathbf{y}$ be a limit point of $\{\mathbf{y}^k\}$. Then $\mathbf{y} \neq \mathbf{x}$ but $\mathbf{y} \in H \cap \partial B$, which contradicts the fact that B is rotund. This concludes the proof of Theorem (2.2).

An inspection of the above proof shows that the following statement is also true. If B is a smooth convex body and every hyperplane parallel to $\mathbf{v}$ supports B at at most one point, then B is $\mathbf{v}$-cylindrical.

We now show that any subspace of $\mathbb{R}^m$ is cylindrical in every direction. We will use the following standard result from linear algebra.

(2.3) LEMMA. *If $\mathbf{v}^1, ..., \mathbf{v}^k$ are linearly independent vectors in $\mathbb{R}^n$, then there exists $M = M(\mathbf{v}^1, ..., \mathbf{v}^k) > 0$ such that whenever $a_i \in \mathbb{R}$, $1 \leqslant i \leqslant k$, and $\|\sum_{i=1}^k a_i \mathbf{v}^i\|_\infty < \gamma$, it must be that $|a_i| < M\gamma$ for $1 \leqslant i \leqslant k$.*

(2.4) THEOREM. *Any subspace of $\mathbb{R}^m$ is cylindrical in every direction.*

Proof. Let U be a subspace of $\mathbb{R}^m$ with basis $\{\mathbf{u}^i\colon 1 \leqslant i \leqslant t\}$ and let

$\mathbf{x} = \sum_{i=1}^{t} a_i \mathbf{u}^i$ be any element of U. Since the image of U under a rotation is also a subspace of $\mathbb{R}^m$, we need only show that U is cylindrical in the direction $\mathbf{e}_m = (0, ..., 0, 1)$.

For $1 \leqslant i \leqslant t$, let $\mathbf{v}^i$ be the $(m-1)$-vector, $(u_1^i, ..., u_{m-1}^i)$. Suppose first that $\{\mathbf{v}^i\colon 1 \leqslant i \leqslant t\}$ is linearly independent. Let $P = \sum_{i=1}^{t} \|\mathbf{u}^i\|_\infty$ and let $M = M(\mathbf{v}^1, ..., \mathbf{v}^t)$ be the number guaranteed by Lemma (2.3). Let $\delta = \varepsilon/(PM)$. If $\mathbf{y} = \sum_{i=1}^{t} b_i \mathbf{u}^i$ and $|y_i - x_i| < \delta$ for $1 \leqslant i \leqslant m-1$, then

$$\left| \sum_{j=1}^{t} (b_j - a_j)\, u_i^j \right| < \delta \qquad \text{for} \quad 1 \leqslant i \leqslant m-1,$$

so Lemma (2.3) implies that

$$|b_i - a_i| < \varepsilon/P \qquad \text{for} \quad 1 \leqslant i \leqslant t.$$

Then

$$\begin{aligned} \|\mathbf{x} - \mathbf{y}\|_\infty &= \left\| \sum_{i=1}^{t} (a_i - b_i)\, \mathbf{u}^i \right\|_\infty \\ &\leqslant \sum_{i=1}^{t} |a_i - b_i|\, \|\mathbf{u}^i\|_\infty < \varepsilon \end{aligned}$$

so $\mathbf{y}$ itself is sufficiently close to $\mathbf{x}$.

We now suppose that $\{\mathbf{v}^i\colon 1 \leqslant i \leqslant t\}$ is linearly dependent. In this case any vertical line which intersects U is completely contained in U. Indeed, if $\mathbf{y} \in U$, then $(y_1, ..., y_{m-1})$ has two distinct expansions $a_1\mathbf{v}^1 + \cdots a_t\mathbf{v}^t$ and $b_1\mathbf{v}^1 + \cdots + b_t\mathbf{v}^t$. Let $\mathbf{z} = a_1\mathbf{u}^1 + \cdots + a_t\mathbf{u}^t$ and $\mathbf{w} = b_1\mathbf{u}^1 + \cdots + b_t\mathbf{u}^t$. Since $\{\mathbf{u}^i\}$ is indepedent, $\mathbf{z} \neq \mathbf{w}$. Since $\mathbf{z}$ and $\mathbf{w}$ are both in $L(\mathbf{y}, \mathbf{e}_m)$ and U is a subspace, $L(\mathbf{y}, \mathbf{e}_m) \subset U$. Hence $(y_1, ..., y_{m-1}, x_m) \in U$ and Theorem (2.4) is established.

In proving Descloux's theorem, we will have occasion to use the following property of an E-cylindrical set.

(2.5) Lemma. *Suppose B is an E-cylindrical subset of $\mathbb{R}^m$. For any* $\mathbf{x}$ *in B, $1 \leqslant k \leqslant m$, and $\varepsilon > 0$, there exists $\delta = \delta(\mathbf{x}, \varepsilon) > 0$ such that, if $\mathbf{y} \in B$ and* $\max\{|x_i - y_i|\colon 1 \leqslant i \leqslant k\} < \delta$, *then there exists* $\mathbf{z}$ *in B such that $z_i = y_i$, $1 \leqslant i \leqslant k$, and $\|\mathbf{x} - \mathbf{z}\|_\infty < \varepsilon$.*

Proof. For each j with $k < j \leqslant m$, let V_j be the subspace of $\mathbb{R}^m$ satisfying the equations

$$x_{k+1} = \cdots = x_{j-1} = x_{j+1} = \cdots = x_m = 0.$$

Then B_j, the orthogonal projection of B onto V_j, is cylindrical in the direction $\mathbf{e}_j$, so there exists z_j such that $(y_1, ..., y_k, z_j) \in B_j$ and

$$\max\{|x_j - z_j|, |x_i - y_i|, 1 \leqslant i \leqslant k\} < \varepsilon.$$

Thus the coordinates, z_j, can be chosen independently, so the vector $(y_1, ..., y_k, z_{k+1}, ..., z_m)$ satisfies the conclusion of Lemma (2.5).

By translation invariance, any linear manifold in $\mathbb{R}^m$ is cylindrical in every direction. If f is a linear functional on the real linear space $\mathbb{R}^m$, then $\{x: f(x) < \alpha)$ and $\{x: f(x) \leqslant \alpha\}$ are *half-spaces.* If K is a half-space and $\mathbf{x}$ is in the interior of K, the criterion for K to be cylindrical in every in every direction is clearly satisfied. Since the boundary of a half-space is a linear manifold, Theorem (2.4) shows that any half-space is cylindrical in every direction, and, by Lemma (2.1), any polyhedral convex set (the intersection of a finite number of half-spaces) is cylindrical in every direction.

Several sets of interest in Approximation Theory are polyhedral convex sets. For example, if $K \subset \mathbb{R}^m$ can be completely described using inequalities relating the coordinates of points in K, then K is a polyhedral convex set. We wish to describe in more detail two such sets. To do so we will revert to the point of view of functions on X.

Let $\mathscr{P}$ be any partial order on X. We say that $g: X \to \mathbb{R}$ is *$\mathscr{P}$-nondecreasing* if $g(x) \leqslant g(y)$ whenever $(x, y) \in \mathscr{P}$. (Note that any lattice, $\mathscr{L}$, of subsets of X can be partially ordered by containment, which induces a partial order, $\mathscr{P}(\mathscr{L})$, on X. Thus the set of all $\mathscr{L}$-measurable functions is the same as the set of all $\mathscr{P}(\mathscr{L})$-nondecreasing functions.) Since the set, K, of all $\mathscr{P}$-nondecreasing functions is exactly the intersection of the half-spaces $\{g: g(x) \leqslant g(y), (x, y) \in \mathscr{P}\}$, K is a polyhedral convex set.

We say that a function $g: X \to \mathbb{R}$ is *convex* if and only if

$$g\left(\sum_{i=1}^{k} \lambda_i x_i\right) \leqslant \sum_{i=1}^{k} \lambda_i g(x_i)$$

whenever, $x_1, ..., x_k$ and $\sum \lambda_i x_i$ are in X, $\lambda_1, ..., \lambda_k \geqslant 0$, and $\sum \lambda_i = 1$. The set of all convex functions on X is a polyhedral convex set.

3. Descloux's Theorem on Cylindrical Sets

Suppose $K \subset F$ is closed, convex, and E-cylindrical, $f \in F$, and f_p is the best l_p-approximation to f by elements of K, $1 < p < \infty$.

(3.1) Theorem. *For each $x \in X$, $\lim_{p \to \infty} f_p(x)$ exists.*

Before we prove Theorem 3.1, we construct the *strict approximation* f_∞ to f. We use critical sets similar to those described in Descloux [1].

Let W_1 be the set of all best l_∞-approximations to f by elements of K. There exists a nonempty subset, C_1, of X such that, if $x \in C_1$ then all elements, h, of W_1 have the same value at x and

$$|h(x) - f(x)| = d(f, K) = d(f, W_1).$$

Let $r_1 = d(f, W_1)$ and, for each x in C_1, let $z_1(x)$ be the value assumed by every element of W_1 at x.

We now introduce some convenient notation. If $Y \subseteq X$ and g is any real-valued function on X, let

$$\|g\|_Y = \sup_{x \in Y} |f(x)|.$$

If in addition $W \subseteq K$, let

$$d_Y(f, W) = \inf_{k \in W} \|f - k\|_Y.$$

Now let $D_1 = X/C_1$ and let $r_2 = d_{D_1}(f, W_1)$. Note that $r_2 < r_1$. If

$$W_2 = \{k \in W_1 : \|f - k\|_{D_1} = r_2\}$$

then as above there is a nonempty set $C_2 \subseteq D_1$ such that if $x \in C_2$ then all elements h of W_2 have the same value at x, and

$$|h(x) - f(x)| = r_2.$$

For $x \in C_2$, let $z_2(x)$ be the value assumed by every element of W_2.

We may continue to solve the restricted problems and define critical sets C_j, values $z_j(x)$, and numbers r_j until X is exhausted.

The strict approximation f_∞ to f is defined by

$$f_\infty(x) = z_j(x) \qquad \text{for} \quad x \in C_j.$$

Proof of Theorem 3.1. We claim that $\lim_{p \to \infty} f_p(x) = f_\infty(x)$ for each $x \in X$.

The proof is by induction. For each x in C_1, $\lim_{p \to \infty} f_p(x) = f_\infty(x)$, since, if not, f_∞ would be a better l_p-aproximation to f by elements of K than is f_p, for sufficiently large p.

By way of induction, suppose that $\lim_{p \to \infty} f_p(x) = f_\infty(x)$ for each x in $C_1 \cup \cdots \cup C_k$. We wish to show that $\lim_{p \to \infty} f_p = f_\infty$ on C_{k+1}. Suppose not. Then there is a point c_0 in C_{k+1}, an $\varepsilon > 0$, and a sequence $p_j \to \infty$ such that $|f_{p_j}(c_0) - f_\infty(c_0)| \geqslant \varepsilon$ for all p_j. Pick a subsequence $\{q_j\}$ of $\{p_j\}$ such that $f_{q_j}(x)$ converges at each x in X. Let $h(x) = \lim_{j \to \infty} f_{q_j}(x)$. Since K is closed, $h \in K$, and, by the induction hypothesis, $h = f_\infty$ on $C_1 \cup \cdots \cup C_k$. Let $D_k = X - (C_1 \cup \cdots \cup C_k)$ and let $r = \max\{|h(x) - f(x)|: x \in D_k\}$. Since $h(c_0) \neq f_\infty(c_0)$, there must exist a point c^* in D_k such that

$$|h(c^*) - f(c^*)| = r > r_{k+1}.$$

Let $\eta = r - r_{k+1}$. By Lemma (2.5), for sufficiently large q_j there exists a function $f^*_{q_j}$ in K such that

$$f^*_{q_j} = f_{q_j}(x), \qquad x \in C_1 \cup \cdots \cup C_k$$

and

$$|f^*_{q_j}(x) - f_\infty(x)| < \eta/4, \qquad x \in D_k.$$

It follows that

$$\begin{aligned} |f^*_{q_j}(x) - f(x)| &\leqslant |f^*_{q_j}(x) - f_\infty(x)| + |f_\infty(x) - f(x)| \\ &< r_{k+1} + \eta/4, \qquad x \in D_k \end{aligned}$$

and

$$\begin{aligned} |f_{q_j}(c^*) - f(c^*)| &\geqslant |f(c^*) - h(c^*)| - |h(c^*) - f_{q_j}(c^*)| \\ &> r - \eta/2 = r_{k+1} + \eta/2 \end{aligned}$$

for sufficiently large q_j. Hence

$$\begin{aligned} &\|f^*_{q_j} - f\|^{q_j}_{q_j} - \|f_{q_j} - f\|^{q_j}_{q_j} \\ &\quad = \sum_{x \in D_k} |f^*_{q_j}(x) - f(x)|^{q_j} - \sum_{x \in D_k} |f_{q_j}(x) - f(x)|^{q_j} \\ &\quad < m(r_{k+1} + \eta/4)^{q_j} - (r_{k+1} + \eta/2)^{q_j} \\ &\quad < 0 \end{aligned}$$

for sufficiently large q_j, which is a contradiction. Hence $\lim_{p \to \infty} f_p(x) = f_\infty(x)$ for each x in C_{k+1} and Theorem (3.1) is established.

We remark that the proof of Theorem (3.1) is still valid in the context of weighted l_p-approximation, i.e., where w is a positive weight function and the l_p-norm is defined by

$$\|g\|_{p,w} = \left(\sum_{x \in X} |g(x)|^p w(x) \right)^{1/p}.$$

4. An Example

In this section we describe an example which shows that Theorem (3.1) cannot be extended to a countably infinte domain. Descloux [1] showed that this theorem could not be extended to the case of approximation on [0, 1] by straight lines passing through the origin. In our example the approximating set consists of all convex functions and there are no constraints.

Fix δ in (0, 0, 1). We will define sequences $\{a_j\}$ and $\{b_j\}$ so that $0 < a_j < b_j < a_{j+1} < 1$ for all j and $a_j \to 1$ as $j \to \infty$; and sequences $\{p_j\}$ and $\{q_j\}$ so that p_j and q_j are positive integers and $p_j < q_j < p_{j+1}$. For the sequences $\{a_j\}$ and $\{b_j\}$ so defined we will define the function f by

$$\begin{aligned} f(x) &= 0, \quad && \text{if} \quad x = 0, x = b_1 = 1 - \delta/8, \text{ or } x = b_j, j = 2, 3, 4, ..., \\ &= 1, && \text{if} \quad x = 2, \\ &= 2, && \text{if} \quad x = a_1 = 1 - \delta/4 \text{ or } x = a_j, j = 2, 3, 4, ..., \end{aligned} \tag{4.1}$$

and we define the following weight function

$$\begin{aligned} w(x) &= 1 - \delta/4, \quad && \text{if} \quad x = 0, \\ &= b_j - a_j, && \text{if} \quad x = a_j, j = 1, 2, 3, ..., \\ &= a_{j+1} - b_j, && \text{if} \quad x = b_j, j = 1, 2, 3, ..., \\ &= 1, && \text{if} \quad x = 2. \end{aligned} \tag{4.2}$$

The example is motivated by the following observations. If f were defined only for $x \leqslant b_n$ and $x = 2$, then, since $f(a_n) = 2$ and $f(b_n) = 0$, it is easy to show that $f_p(x) \to 1$, as $p \to \infty$, for all x, where f_p is the best weighted p-approximant to f by convex functions. On the other hand, if f were defined only for $x \leqslant a_{n+1}$ and $x = 2$, then, since $f(b_n) = 0$ and $f(a_{n+1}) = 2$, it is possible to show that $f_p(x)$ tends to 1 for $x \leqslant b_n$ and to points on the line segment joining $(b_n, 1)$ to $(2, 2 - \varepsilon)$ for $x > b_n$ (for an appropriately chosen $\varepsilon > 0$). This convex function minimizes the distance $|f(a_{n+1}) - f_p(a_{n+1})|$ without letting $|f(2) - f_p(2)|$ exceed that distance.

Before defining a_j and b_j for $j \geqslant 2$, we show that regardless of how they are defined

$$\lim_{p \to \infty} f_p(a_1) = \lim_{p \to \infty} f_p(b_1) = \lim_{p \to \infty} f_p(0) = 1. \tag{4.3}$$

To see this, suppose, for example, that $\liminf_{p \to \infty} f_p(a_1) < 1$. Then there is an $\varepsilon > 0$ and a sequence $p_j \uparrow \infty$ so that $f_{p_j}(a_1) < 1 - \varepsilon$ for all j. If we let $f^*_{p_j}(x) \equiv 1$, then

$$\|f - f_{p_j}\|^{p_j}_{p_j,w} - \|f - f^*_{p_j}\|^{p_j}_{p_j,w} > (1+\varepsilon)^{p_j} w(a_1) - 1 \cdot \sum_{x<1} w(x) > 0 \tag{4.4}$$

if p_j is sufficiently large, a contradiction. On the other hand, if $\limsup_{p\to\infty} f_p(a_1) > 1$, then there is an $\varepsilon > 0$ and a sequence $p_j \to \infty$ so that $f_{p_j}(a_1) > 1 + \varepsilon$, for all j. If in addition, $\limsup_{j\to\infty} f_{p_j}(0) > 1 + \varepsilon/2$, then there are arbitrarily large p_j's so that $f_{p_j}(0) > 1 + \varepsilon/4$. Since $f(0) = 0$, by an argument similar to that in (4.4), we have that $f^*_{p_j}(x) \equiv 1$ is a better p_j-approximant to f than is f_{p_j}, a contradiction. But, if in addition, $\limsup_{j\to\infty} f_{p_j}(0) \leqslant 1 + \varepsilon/2$, then for sufficiently large p_j we have $f_{p_j}(0) < 1 + (3\varepsilon)/4$. It follows from convexity that $f_{p_j}(b_1) > 1 + \varepsilon$. Since $f(b_1) = 0$, we would again have that $f^*_{p_j}(x) \equiv 1$ is a better p_j-approximant to f than is f_{p_j}, for large p_j, a contradiction. Thus, $\lim_{p\to\infty} f_p(a_1) = 1$. The proofs that $\lim_{p\to\infty} f_p(b_1) = \lim_{p\to\infty} f_p(0) = 1$ are similar.

We have defined $a_1 = 1 - \delta/4$ and $b_1 = 1 - \delta/8$. We now define $p_1 = 1$ and $q_1 = 2$. To define the sequences $\{a_j\}$ and $\{b_j\}$ inductively, we assume that a_j, b_j, p_j, and q_j have been defined for $j = 1, 2, ..., n-1$. Using techniques similar to those above, we can show that regardless of how a_j and b_j are defined for $j \geqslant n$, we have $\lim_{p\to\infty} f_p(a_{n-1}) = \lim_{p\to\infty} f_p(b_{n-1}) = 1$. Thus, we choose $p_n > q_n$ so large that regardless of how a_j and b_j are defined for $j \geqslant n$, we have

> the line through $(a_{n-1}, f_{p_n}(a_{n-1}))$ and $(b_{n-1}, f_{p_n}(b_{n-1}))$ has slope with absolute value less than $\delta/[4(2 - a_{n-1})]$ and $|f_{p_n}(a_{n-1}) - 1| < \delta/4$. (4.5)

We also choose $a_n > b_{n-1}$ so that $a_n < 1$ and

$$2^{p_n}(1 - a_n) < \delta^{p_n} - (\delta/2)^{p_n}. \tag{4.6}$$

We then define $f^*_{q_n}(x)$ by

$$\begin{aligned} f^*_{q_n}(x) &= f_{q_n}(x), \quad x \leqslant b_{n-1} \\ &= f_{q_n}(b_{n-1}) + (x - b_{n-1})(2 - \delta - f_{q_n}(b_{n-1}))/(2 - b_{n-1}), \qquad x \geqslant a_n. \end{aligned}$$

Since $\lim_{p\to\infty} f_p(a_{n-1}) = \lim_{p\to\infty} f_p(b_{n-1}) = 1$, regardless of how a_j and b_j are defined for $j \geqslant n$, we may choose q_n so large that

$$\begin{aligned} &\text{(a)} \quad f^*_{q_n}(x) \text{ is convex,} \\ &\text{(b)} \quad |f_{q_n}(b_{n-1}) - 1| < \delta/4. \end{aligned} \tag{4.7}$$

Since $b_{n-1} > a_{n-1} \geqslant 1-\delta/4$, we have by (4.7)(b) that

$$2 - f_{q_n}(b_{n-1}) - (a_n - b_n)(2 - 2\delta - f_{q_n}(b_{n-1}))/(2 - b_{n-1}) > 1 - \delta/2.$$

Thus, it is clear that we can choose q_n so large that (4.7) and the following inequality both hold:

$$\begin{aligned} A = \{ &|2 - f_{q_n}(b_{n-1}) - (a_n - b_{n-1})(2 - 2\delta - f_{q_n}(b_{n-1}))/(2 - b_{n-1})|^{q_n} \\ &- |2 - f_{q_n}(b_{n-1}) - (a_n - b_n)(2 - \delta - f_{q_n}(b_{n-1}))/(2 - b_{n-1})|^{q_n} \} \\ &\times \tfrac{1}{2}(1 - a_n) - (1 - \delta)^{q_n} > 0. \end{aligned} \tag{4.8}$$

Finally, we choose $b_n > \frac{1}{2}(a_n + 1)$ so large that $b_n < 1$ and

$$2^{q_n}(1 - b_n) < A, \tag{4.9}$$

where A is the quantity in (4.8).

With $f(x)$ and $w(x)$ defined as in (4.1) and (4.2) for the sequences $\{a_j\}$ and $\{b_j\}$ defined inductively above, we will show that

$$f_{p_{j-1}}(2) \leqslant 1 + \delta \qquad \text{for all } j, \tag{4.10}$$

and

$$f_{q_{j-1}}(2) \geqslant 2 - 2\delta \qquad \text{for all } j. \tag{4.11}$$

It follows from (4.10) and (4.11) that $\lim_{p \to \infty} f_p(2)$ does not exist.

If (4.10) is not true, then for some j,

$$f_{p_j}(2) > 1 + \delta. \tag{4.12}$$

Now define $f^*_{p_j}$ by

$$\begin{aligned} f^*_{p_j}(x) &= f_{p_j}(x), & x &< a_j, \\ &= f_{p_j}(a_{j-1}) + (x - a_{j-1})(f_{p_j}(b_{j-1}) - f_{p_j}(a_{j-1}))/(b_{j-1} - a_{j-1}), & x &\geqslant a_j. \end{aligned}$$

Clearly $f^*_{p_j}$ is convex, and by (4.5) we conclude that $|f^*_{p_j}(2) - 1| < \delta/2$. Thus, from (4.5), (4.6), and (4.12) we have

$$\begin{aligned} &\|f - f_{p_j}\|^{p_j}_{p_j,w} - \|f - f^*_{p_j}\|^{p_j}_{p_j,w} \\ &\quad > |f(2) - f_{p_j}(2)|^{p_j} - |f(2) - f^*_{p_j}(2)|^{p_j} - \sum_{2 > x \geqslant a_j} |f(x) - f^*_{p_j}(x)|^{p_j} w(x) \\ &\quad > \delta^{p_j} - (\delta/2)^{p_j} - 2^{p_j}\Bigl(\sum_{2 > x \geqslant a_j} w(x)\Bigr) \\ &\quad = \delta^{p_j} - (\delta/2)^{p_j} - 2^{p_j}(1 - a_j) > 0, \end{aligned}$$

a contradiction which proves (4.10).

If (4.11) is not true, then for some j, $f_{q_j}(2) < 2 - 2\delta$. Since the equation of the line through $(b_{j-1}, f_{q_j}(b_{j-1}))$ and $(2, 2-2\delta)$ is $y = f_{q_j}(b_{j-1}) + (x - b_{j-1})(2 - 2\delta - f_{q_j}(b_{j-1}))/(2 - b_{j-1})$, we have by convexity that

$$f_{q_j}(a_j) < f_{q_j}(b_{j-1}) + (a_j - b_{j-1})(2 - 2\delta - f_{q_j}(b_{j-1}))/(2 - b_{j-1}). \quad (4.13)$$

By (4.7)(a), $f^*_{q_j}$ is convex and, using (4.13), (4.8), and (4.9),

$$\begin{aligned}
&\|f - f_{q_j}\|^{q_j}_{q_j,w} - \|f - f^*_{q_j}\|^{q_j}_{q_j,w}\\
&\quad \geqslant \{|f(a_j) - f_{q_j}(a_j)|^{q_j} - |f(a_j) - f^*_{q_j}(a_j)|^{q_j}\}\, w(a_j)\\
&\quad\quad - \sum_{a_j < x < 1} |f(x) - f_{q_j}(x)|^{q_j}\, w(x) - |f(2) - f^*_{q_j}(2)|^{q_j}\\
&\quad \geqslant \{|2 - f_{q_j}(b_{j-1}) - (a_j - b_{j-1})(2 - 2\delta - f_{q_j}(b_{j-1}))/(2 - b_{j-1})|^{q_j}\\
&\quad\quad - |2 - f_{q_j}(b_{j-1}) - (a_j - b_{j-1})(2 - \delta - f_{q_j}(b_{j-1}))/(2 - b_{j-1})|^{q_j}\}\\
&\quad\quad \times (b_j - a_j) - 2^{q_j}(1 - b_j) - |1 - (2 - \delta)|^{q_j} > 0.
\end{aligned}$$

This is a contradiction which proves (4.11).

5. Continuous Convex Approximations

Let f in $C[0, 1]$ and p in $(1, \infty)$ be fixed. Let f_p be the best L_p-approximation to f by continuous convex functions on $[0, 1]$. In this section we show that f_p is the limit of a sequence of convex functions associated with the solutions to certain discrete convex approximation problems.

Let $\{X_n\colon n \geqslant 1\}$ be a sequence of partitions of $[0, 1]$ such that $\lim_{n\to\infty} \delta_n = 0$, where δ_n is the mesh of X_n, $n \geqslant 1$. Define $w_n\colon X_n \to \mathbb{R}$ by

$$\begin{aligned}
w_n(x) &= 0, & x &= x_0,\\
&= x - \max\{y \in X_n\colon y < x\}, & x &\neq x_0,
\end{aligned}$$

let $f^n = f|X_n$, and let f^n_p be the best l_p-approximation to f^n by convex functions on X_n. For any function $g\colon X_n \to \mathbb{R}$, let $\bar{g}$ be the piecewise linear function on $[0, 1]$ which agrees with g on X_n and is linear on each interval in $[0, 1] - X_n$.

(5.1) Lemma. *If $n > 1$, then* $\|f^n_p\|_\infty \leqslant 6\|f^n\|_\infty$.

Proof. Suppose $\|f^n\|_\infty = M$. Since $\max(g, -M)$ is convex, it is necessary that $g(x) \geqslant -M$ for all x in X_n. If there exists y in $X_n \cap (0, 1)$

such that $g(y) > 6M$ then $g(x) > 6M$ either for all x in X_n greater than y or for all x in X_n less than y. Since the two cases can be dealt with similarly, we treat only the first.

We will derive a contradiction by constructing a convex function which is strictly $\|\cdot\|_{p,w_n}$-closer to f than is g. Let $\beta = \sup\{t < 1: \bar{g}(t) = M\}$ and define $\bar{h}$: $[0, 1] \to \mathbb{R}$ by

$$\begin{aligned} \bar{h}(t) &= \bar{g}(t), && t \in [0, \beta], \\ &= \bar{g}(\beta) + \bar{g}'_-(\beta)(t-\beta), && t \in [\beta, 1], \end{aligned}$$

where $\bar{g}'_-(\beta)$ is the left derivative of $\bar{g}$ at β. Let $h = \bar{h} | X_n$. If $g(y) \neq h(y)$ for some y in $X_n \cap (\beta, 1)$, h is clearly a better $\|\cdot\|_{p,w_n}$-approximation to f than is g, a contradiction. Thus, $\bar{g}$ must be linear on $(\beta, 1)$, with slope $\bar{g}'_-(\beta)$.

Choose z in X_n so that $(z, 1)$ is the maximal open interval on which $\bar{g}$ has slope $\bar{g}'_-(\beta)$. Choose m so that $\bar{g}'_-(z) < m < \bar{g}'_+(z)$ and so that the line L with slope m containing the point $(z, g(z))$ satisfies the condition $L(1) > \max(g(1) - M, 6M)$. Define K: $X_n \to \mathbb{R}$ by

$$\begin{aligned} k(x) &= g(x), && x \in [0, z] \cap X_n, \\ &= L(x), && x \in [z, 1] \cap X_n. \end{aligned}$$

Let σ and τ satisfy $\bar{k}(\sigma) = M$ and $\bar{k}(\tau) = 4M$.

If $p = \infty$, clearly k is $\|\cdot\|_{p,w_n}$-closer to f than is g. Suppose $1 \leqslant p < \infty$. Given $\delta \in \mathbb{R}$, let $\Psi_\delta(x) = |x|^p - |x-\delta|^p$. Then, for $x > \delta/2 > 0$,

$$\Psi'_\delta(x) > 0. \tag{5.2}$$

Let $A = \{x \in X_n : x \geqslant \tau\}$ and $B = \{x \in X_n : z < x \leqslant \sigma$ and $|f(x) - k(x)| > \frac{1}{2}|g(x) - k(x)|\}$. Then, for $x \in A$,

$$\begin{aligned} &|f(x) - g(x)|^p - |f(x) - k(x)|^p \\ &\quad > [(\bar{g}(\tau) - \bar{k}(\tau)) + (k(x) - f(x))]^p - [k(x) - f(x)]^p \\ &\quad > [(\bar{g}(\tau) - \bar{k}(\tau)) + 2M]^p - [2M]^p, \end{aligned} \tag{5.3}$$

where the last inequality follows from (5.2). If $x \in B$, then, considering $-\Psi$ in (5.2),

$$\begin{aligned} &|f(x) - g(x)|^p - |f(x) - k(x)|^p \\ &\quad = |(f(x) - k(x)) - (g(x) - k(x))|^p - |f(x) - k(x)|^p \\ &\quad \geqslant [2M - (g(x) - k(x))]^p - [2M]^p \\ &\quad \geqslant [2M - (\bar{g}(\tau) - \bar{k}(\tau))]^p - [2M]^p. \end{aligned} \tag{5.4}$$

If $x \in X_n - (A \cup B)$, then $|f(x) - k(x)| \leqslant |f(x) - g(x)|$. Combining (5.3) and (5.4) gives

$$\begin{aligned}
&\|f - g\|_{p,w_n}^p - \|f - k\|_{p,w_n}^p \\
&\quad > \sum_{A \cup B} [|f(x) - g(x)|^p - |f(x) - k(x)|^p]\, w_n(x) \\
&\quad > (1 - \tau)[((\bar{g}(\tau) - \bar{k}(\tau)) + 2M)^p - (2M)^p] \\
&\qquad + (\sigma - z)[(2M - (\bar{g}(\tau) - \bar{k}(\tau)))^p - (2M)^p].
\end{aligned}$$

Let $\delta = \bar{g}(\tau) - \bar{k}(\tau)$. Using (5.2) and the fact that $(\sigma - z) < (1 - \tau)$, we have that

$$\begin{aligned}
&\|f - g\|_{p,w_n}^p - \|f - k\|_{p,w_n}^p \\
&\quad > (1 - \tau)[(2M + \delta)^p - (2M)^p - (2M)^p + (2M - \delta)^p] \\
&\quad = (1 - \tau)[\Psi_\delta(2M + \delta) - \Psi_\delta(2M)] > 0.
\end{aligned}$$

Thus k is a better $\|\cdot\|_{p,w_n}$-approximation to f than is g, a contradiction which establishes Lemma (5.1).

The above proof can be modified to establish the upper bound of $5\|f^n\|_\infty$. We believe that even this estimate is not sharp.

Let d be the distance from f to the set of convex functions on $[0, 1]$, i.e., $d = [\int_0^1 |f - f_p|^p]^{1/p}$. Similarly, for $n > 1$, let $d_n = [\sum_{x \in X_n} |f^n(x) - f_p^n(x)|^p w_n(x)]^{1/p}$, the weighted distance from f^n to the set of convex functions on X_n.

(5.5) Theorem. *The sequence* $\{f_p^n : n > 1\}$ *converges pointwise on* $(0, 1)$ *and uniformly on closed subsets of* $(0, 1)$ *to* f_p.

Proof. Clearly $\|f^n\|_\infty \leqslant \|f\|_\infty$ for all $n > 1$, so Lemma (5.1) implies that $\|f_p^n\|_\infty \leqslant 6\|f\|_\infty$. Thus, by Theorem 10.9 in [4], every subsequence of $\{f_p^n\}$ contains a subsequence which converges uniformly on compact subsets of $(0, 1)$. By Theorem IV.6.7 in [2], the set $\{f_p^n\}$ is equicontinuous on any interval of the form $[a, b]$, where $0 < a < b < 1$.

Let $\varepsilon > 0$ be given and choose N_1 so that, for $n \geqslant N_1$,

$$\int_{y_n}^{z_n} |f - f_p^n|^p \geqslant \int_0^1 |f - f_p^n|^p - \varepsilon/2,$$

where $y_n = \min(X_n \cap (0, 1))$ and $z_n = \max(X_n \cap (0, 1))$. Then there exists $N_2 \geqslant N_1$ such that, for $n \geqslant N_2$,

$$
\begin{aligned}
d^p &= \int_0^1 |f - f_p|^p \\
&\geqslant \sum_{x \in X_n} |f(x) - f_p(x)|^p w_n(x) - \varepsilon \\
&\geqslant (d_n)^p - \varepsilon
\end{aligned}
$$

and

$$
\begin{aligned}
(d_n)^p &= \sum_{x \in X_n} |f(x) - \bar{f}_p^n(x)|^p w_n(x) \\
&\geqslant \sum_{x \in X_n \cap (0,1)} |f(x) - \bar{f}_p^n(x)|^p w_n(x) \\
&\geqslant \int_{y_n}^{z_n} |f - \bar{f}_p^n|^p - \varepsilon/2 \\
&\geqslant \int_0^1 |f - \bar{f}_p^n|^p - \varepsilon \\
&\geqslant d^p - \varepsilon.
\end{aligned}
$$

It follows that $\lim_{n \to \infty} d_n = d$.

For $1 < p < \infty$, the following property holds: given $\varepsilon > 0$, there exists $\delta > 0$ such that, if g is convex and $\|f - g\|_p \leqslant d + \delta$, then $\|g - f_p\|_p \leqslant \varepsilon$. This property and the previous calculations imply that $\bar{f}_p^n \to f_p$ in L_p as $n \to \infty$. Thus, by Lemma 4 in [3], $\bar{f}_p^n \to f_p$ pointwise on $(0, 1)$ and uniformly on closed subsets of $(0, 1)$.

References

1. J. Descloux, Approximations in L^p and Chebyshev approximations, *J. Soc. Ind. Appl. Math.* **11** (1963), 1017–1026.
2. N. Dunford and J. Schwartz, "Linear Operators," Part I, Interscience, New York, 1958.
3. R. Huotari and A. Meyerowitz, Best L_1-approximation by convex functions, submitted for publication.
4. R. Rockafellar, "Convex, Analysis," Princeton, Univ. Press, Princeton, NJ, 1970.