JOURNAL OF APPROXIMATION THEORY 4, 120–136 (1971)

Kolmogoroff's Criterion for Constrained Rational Approximation

L. WUYTACK

Aspirant of the N.F.W.O., University of Leuven, Institute for Applied Mathematics, Heverlee, Belgium

Communicated by G. Meinardus

Received May 20, 1969; revised April 2, 1970

Many problems in approximation theory can be formulated with the help of generalized weight functions. The use of such functions, introduced by Moursund [8] in 1966, proved to be of interest, especially in approximation problems with constraints ([5], [6], [9], [10], [12–15]). Moursund, who considered linear appproximation, studied existence, characterization, and uniqueness of best approximations.

In the present paper we give a generalization of the characterization theorem for the case of generalized rational approximation, using a generalized weight function. In the theory of uniform approximation, without constraints, this theorem is known as Kolmogoroff's criterion ([7, p. 13 and p. 125]). The characterization theorem in Moursund's paper fails to hold in certain special cases. Therefore, the present paper can also be considered as a correction to the theory given in [8].

1. BASIC NOTATIONS AND DEFINITIONS

Suppose X is a compact metric space; the distance between two points, x and y, of X denoted by $d(x, y)$; $C(X)$ is the class of all continuous real valued functions defined on X; f is an element of $C(X)$; P and Q are linear subspaces of $C(X)$ of finite dimension. Let R be the class of generalized rational functions, namely,

$$R = \{p/q : p \in P, q \in Q \text{ and } q > 0 \text{ on } X\}.$$

We suppose that R is not empty.

It was the idea of Moursund to minimize $\sup_{x\in X} | W[x, r(x) - f(x)]|$ instead of $\sup_{x\in X} | r(x) - f(x)|$ (ordinary best approximation). The function $W(x, y)$ has to satisfy certain conditions in order that the approximation problem should make sense ([8, p. 435 and 441]). Thus, we assume that $W(x, y)$ is a function defined for $x \in X$ and $-\infty < y < \infty$, with its range in the extended real line, and having the following properties:

(1a) $\operatorname{sgn} W(x, y) = \operatorname{sgn} y \quad$ for all x, y;

(1b) $W(x, y)$ is monotone nondecreasing in $y \quad$ for all $x \in X$, and if

$$|y_1| < |y_2|, \qquad \operatorname{sgn} y_1 = \operatorname{sgn} y_2 \quad \text{and} \quad |W(x, y_1)| < \infty$$

then

$$|W(x, y_1)| < |W(x, y_2)| \qquad \text{for all} \quad x \in X.$$

(1c) If g is a continuous real valued function defined on X such that $\sup_{x \in X} |W[x, g(x)]| < \infty$, then for every compact subset $Y \subseteq X$ and every $\epsilon > 0$,

$$\sup_{x \in Y} |W[x, g(x)]| < \max\{\sup_{Y} |W[x, g(x) + \epsilon]|, \sup_{Y} |W[x, g(x) - \epsilon]|\}.$$

A function $W(x, y)$ which satisfies the above conditions is called a generalized weight function. Some of the well-known approximation problems which can be stated in terms of generalized weight functions are, uniform approximation with interpolatory constraints (see Example 3); one-sided uniform approximation; and restricted range approximation. If $W(x, y) = y$ we have ordinary uniform approximation.

Let $M(g) = \sup_{x \in X} |W[x, g(x)]|$. An element $r \in R$ is called an approximation to f if $M(r - f) < \infty$. Further, r is called a best approximation to f if it is an approximation to f and if $M(r - f) \leqslant M(r_1 - f)$ for every $r_1 \in R$.

We remark that in order to deal with the question of existence of a best approximation, $W(x, y)$ has to satisfy further restrictions (see [16]). In the case that $W(x, y)$ is continuous, the characterization of a best approximation is similar to that for ordinary best approximation ([11, p. 885] and [2, p. 160]).

It is important to notice that property (1b) of $W(x, y)$ does not imply property (1c). As a simple example consider $X = [0, 1]$ and

$$\begin{aligned} &W(x, y) = y \quad \text{if } y \leqslant 0, \qquad x \in [0, 1], \\ &W(x, y) = y/2 \quad \text{if } y > 0, \qquad x = 0, \\ &W(x, y) = (2 - 2x + xy)/2 \quad \text{if } y > 0, \qquad x \in (0, 1]. \end{aligned}$$

This $W(x, y)$ satisfies (1a) and (1b), but not (1c). To see this, take $g = 1$ and $Y = [0, 1/2]$. Then $\sup_{x \in Y} |W[x, g(x)]| = 1$ but also, if $|\epsilon| < 1/2$,

$$\sup_{x \in Y} |W[x, g(x) \pm \epsilon]| = \sup_{x \in Y} \left| 1 + x \cdot \left(\frac{-1 \pm \epsilon}{2}\right) \right| = 1.$$

In ordinary uniform approximation an important role is played by the extremals of the error curve $e(x) = r(x) - f(x)$. A more general concept of extremal point is used in the theory of uniform approximation with generalized weight functions ([8, p. 443]). Let

$$U(t, \delta) = \{x : x \in X \text{ and } d(t, x) < \delta\},$$
$$U'(t, \delta) = \{x : x \in X \text{ and } d(t, x) \leqslant \delta\}.$$

Suppose $e = r - f$ and $M(e) = E < \infty$.

(2a) A point $t \in X$ is called a zero extremal with respect to r and f if for every $\epsilon > 0$ and $\delta > 0$ there exist points x_1, $x_2 \in U(t, \delta)$ such that $W[x_1, e(x_1) + \epsilon] > E$ and $W[x_2, e(x_2) - \epsilon] < -E$.

(2b) A point $t \in X$ is called a plus extremal with respect to r and f if the following two conditions are satisfied:

(i) For each $\epsilon_1 > 0$ and $\delta_1 > 0$ there exists a point $x_1 \in U(t, \delta_1)$ such that $W[x_1, e(x_1) + \epsilon_1] > E$.

(ii) There exist $\epsilon_2 > 0$ and $\delta_2 > 0$ such that $x_2 \in U(t, \delta_2)$ implies $W[x_2, e(x_2) - \epsilon_2] > -E$.

(2c) A point $t \in X$ is called a minus extremal with respect to r and f if the following two conditions are satisfied:

(i) For each $\epsilon_1 > 0$ and $\delta_1 > 0$ there exists a point $x_1 \in U(t, \delta_1)$ such that $W[x_1, e(x_1) - \epsilon_1] < -E$.

(ii) There exist $\epsilon_2 > 0$ and $\delta_2 > 0$ such that $x_2 \in U(t, \delta_2)$ implies $W[x_2, e(x_2) + \epsilon_2] < E$.

We denote by $C(r)$ the set of zero, plus and minus extremals with respect to r and f. The points of $C(r)$ are called critical points with respect to r and f. Observe that $t \notin C(r)$ if there exist $\epsilon > 0$, $\delta > 0$ such that $x \in U(t, \delta)$ implies $|W[x, e(x) \pm \epsilon]| \leqslant E$.

2. Properties of the Set of Critical Points

In the proof of the characterization theorem an important role is played by properties of the critical points. Some of these properties are mentioned below.

LEMMA 1. *If $g \in C(X)$, $t \in X$ and for all $\epsilon > 0$, $\delta > 0$ there exists a point $x \in U(t, \delta)$ such that $W[x, g(x) + \epsilon] > 0$, then $g(t) \geqslant 0$.*

Proof. Suppose $g(t) = b < 0$. Because of the continuity of g there exists $\delta_0 > 0$ such that $x \in U(t, \delta_0)$ implies

$$g(x) + \epsilon_0 < 0 \qquad (\epsilon_0 = -b/2).$$

Using (1a), we get: there exist $\delta_0 > 0$ and $\epsilon_0 > 0$ such that $x \in U(t, \delta_0)$ implies $W[x, g(x) + \epsilon_0] < 0$. This is a contradiction; thus $g(t) < 0$ is not possible.

LEMMA 2. *If $h \in C(X)$, $t \in X$ and for all $\epsilon > 0$, $\delta > 0$ there exists a point $x \in U(t, \delta)$ such that $W[x, h(x) - \epsilon] < 0$ then $h(t) \leqslant 0$.*

The proof of this lemma is similar to that of Lemma 1. Using Lemma 1 and 2, and definitions (2a), (2b), and (2c) we get

THEOREM 1. *If r is an approximation to f and $t \in C(r)$ then: if t is a zero extremal, $r(t) = f(t)$; if t is a plus extremal, $r(t) - f(t) \geqslant 0$; and if t is a minus extremal, $r(t) - f(t) \leqslant 0$.*

For illustrations of this theorem see the examples given later. We remark that $r(t) = f(t)$ is possible if t is a plus (or minus) extremal and that $| W[t, r(t) - f(t)]| < E = M(r - f)$ is possible if $t \in C(r)$ (see [17]).

LEMMA 3. *Suppose $u \in C(X)$, $M(u) = E < \infty$, $\epsilon > 0$, $\delta > 0$, $t \in X$ and $V = U'(t, \delta/2)$.* If $| W[x, u(x) \pm 2\epsilon]| \leqslant E$ for $x \in U(t, \delta)$, *then*

$$\sup_V | W[x, u(x) \pm \epsilon]| < E.$$

Proof. For $x \in U(t, \delta)$ we have

$$-E \leqslant W[x, u(x) \pm 2\epsilon] \leqslant E. \tag{2.1}$$

If $u(x) - 2\epsilon \leqslant h(x) \leqslant u(x) + 2\epsilon$ for all $x \in U(t, \delta)$ then, using property (1b), (2.1) implies

$$| W[x, h(x)] | \leqslant E \qquad \text{for all} \quad x \in U(t, \delta). \tag{2.2}$$

Put $A = X \setminus U(t, \delta)$; then A and V are disjoint closed subsets of the compact metric space X. Using Ursysohn's theorem (see [1, p. 74]) there exists an element $v \in C(X)$ such that

$$v(x) = 0 \quad \text{if } x \in A; \qquad v(x) = 1 \quad \text{if} \quad x \in V;$$
$$v(x) \in [0, 1] \quad \text{if } x \in X \setminus (A \cup V).$$

Put $h^+ = u + \epsilon v$ and $h^- = u - \epsilon v$. Then we have for all $x \in U(t, \delta)$:

$$u(x) \leqslant h^+(x) \leqslant u(x) + \epsilon \qquad \text{and} \qquad u(x) - \epsilon \leqslant h^-(x) \leqslant u(x).$$

Because of (2.2) we get then, for all $x \in U(t, \delta)$,

$$| W[x, h^+(x)]| \leqslant E \qquad \text{and} \qquad | W[x, h^-(x)]| \leqslant E.$$

Using the fact that $u(x) = h^+(x) = h^-(x)$ for $x \in A$, we get

$$M(h^+) < \infty \qquad \text{and} \qquad M(h^-) < \infty.$$

Now property (1c) may be applied with V as Y; and h^+ as g. We obtain

$$\sup_V | W[x, h^+(x)]| < \max\{\sup_V | W[x, h^+(x) + \epsilon], \sup_V | W[x, h^+(x) - \epsilon]|\}.$$

Using (2.2) and the definition of h^+ we get

$$\sup_V | W[x, u(x) + \epsilon]| < E.$$

In the same way, using h^- instead of h^+, we obtain

$$\sup_V | W[x, u(x) - \epsilon]| < E.$$

With the help of Lemma 3 we are now able to prove a theorem which is similar to one given by Moursund (see [8, theorem 4, p. 442]).

THEOREM 2. *If $M(r - f) = E < \infty$ and $t \in X \setminus C(r)$ then there exist $\epsilon > 0$, $\delta > 0$ and $E_0 > 0$ such that $x \in U(t, \delta)$ implies*

$$| W[x, r(x) - f(x) \pm \epsilon]| \leqslant E_0 < E.$$

Proof. Since t is not a critical point, there exist $\epsilon_1 > 0$ and $\delta_1 > 0$ such that $x \in U(t, \delta_1)$ implies

$$| W[x, e(x) \pm \epsilon_1]| \leqslant E \qquad (e(x) = r(x) - f(x)),$$

that is

$$| W[x, e(x) \pm 2\epsilon]| \leqslant E \qquad (\epsilon = \epsilon_1/2).$$

Putting $\delta = \delta_1/2$, $V = U'(t, \delta)$, and $E_0 = \sup_V | W[x, e(x) \pm \epsilon]|$, we get, using Lemma 3, $E_0 < E$ which concludes the proof.

Making use of Theorem 2, it is now possible to obtain an important property of the set $C(r)$, expressed in the following:

THEOREM 3. *The set $C(r)$ of critical points of an approximation r to f is closed.*

The proof can be carried out as in [8, Theorem 6, p. 443].

The set of minus extremals (or plus extremals) need not be closed. We show this by the following example. Suppose $X = [-1, 1]$ and consider the following generalized weight function:

$$W(x, y) = n \cdot y \quad \text{if } x = 1/n, \qquad n = 1, 2,...;$$
$$W(x, y) = y \quad \text{if } x \neq 0 \quad \text{and} \quad x \neq 1/n, \qquad n = 1, 2,...;$$
$$W(0, y) = 2 + y \quad \text{if } y > 0; \qquad W(0, y) = 0 \quad \text{if } y = 0;$$
$$W(0, y) = -2 + y \quad \text{if } y < 0.$$

Let $f(x) = x + 1$ on X, and let us approximate f by a function which is constant on X, using the above generalized weight function $W(x, y)$. A best approximation for f is $r = 1$, with $M(r - f) = 1$. The point $x = 0$ is a zero extremal and the points $x_n = 1/n$, for $n = 1, 2,...$, are all minus extremals. Thus, the set of minus extremals is not closed in X.

3. SUFFICIENT AND NECESSARY CONDITIONS FOR BEST APPROXIMATION

As in the theory of ordinary uniform approximation an important role is played by the linear subspaces $P + rQ$ of $C(X)$, where $r \in R$. Such a linear subspace consists of the elements of $C(X)$ of the form $p + rq$, where $p \in P$ and $q \in Q$. In the following theorems we set, as before, $e = r - f$.

THEOREM 4. *Suppose $r \in R$ is an approximation to f with $M(e) = E > 0$. A sufficient condition for r to be a best approximation is that there exists no element $v \in P + rQ$ with $v \not\equiv 0$ and*

$$\begin{array}{ll} v(t) \leqslant 0 & \textit{for every plus extremal } t \in C(r) \textit{ with } W[t, e(t)] < E; \\ v(t) < 0 & \textit{for every plus extremal } t \in C(r) \textit{ with } W[t, e(t)] = E; \\ v(t) \geqslant 0 & \textit{for every minus extremal } t \in C(r) \textit{ with } W[t, e(t)] > -E; \\ v(t) > 0 & \textit{for every minus extremal } t \in C(r) \textit{ with } W[t, e(t)] = -E; \\ v(t) = 0 & \textit{for every zero extremal } t \in C(r). \end{array} \tag{3.1}$$

Proof. Suppose r is not a best approximation to f. Then there exists a better one in R, say, $r_1 = p_1/q_1$, so that if $e_1 = r_1 - f$ then

$$M(e_1) < E. \tag{3.2}$$

Put $v = q_1(r_1 - r)$; then $v \in P + rQ$. We show that v satisfies all the conditions (3.1).

(a) $v(t) \leqslant 0$ for every plus extremal $t \in C(r)$ with $W[t, e(t)] < E$. Suppose $v(t) > 0$, i.e., $r_1(t) - r(t) = a > 0$. By continuity of r_1 and r in X, there exists $\delta > 0$ such that $x \in U(t, \delta)$ implies $\epsilon < r_1(x) - r(x)$, where $\epsilon = a/2$. Then $r_1(x) - f(x) > r(x) - f(x) + \epsilon$ for all $x \in U(t, \delta)$. Because of the monotonicity property (1b), this implies:

$$W[x, r(x) - f(x) + \epsilon] \leqslant W[x, r_1(x) - f(x)] \qquad \text{for all} \quad x \in U(t, \delta). \tag{3.3}$$

Using definition (2b) for a plus extremal, we get that there exists a point $x_1 \in U(t, \delta)$ with $W[x_1, e(x_1) + \epsilon] > E$. This result, together with (3.3) contradict (3.2). Consequently, we must have $v(t) \leqslant 0$.

(b) $v(t) < 0$ for every plus extremal $t \in C(r)$ with $W[t, e(t)] = E$. Because $M(e_1) < E$ we have $W[t, e_1(t)] < W[t, e(t)] = E$. Using (1b), we get $e_1(t) < e(t)$, i.e., $r_1(t) < r(t)$. Since $q_1(t) > 0$, this implies $v(t) < 0$.

(c) $v(t) \geqslant 0$ for every minus extremal $t \in C(r)$ with $W[t, e(t)] > -E$. The proof is similar to that of (a).

(d) $v(t) > 0$ for every minus extremal $t \in C(r)$ with $W[t, e(t)] = -E$. This is proved using the same method as in (b).

(e) $v(t) = 0$ for every zero extremal $t \in C(r)$. Suppose $v(t) > 0$. Using the same method as in (a), and definition (2a), we reach a contradiction to (3.2). The possibility $v(t) < 0$ is handled in the same way as (c), yielding again a contradiction to (3.2).

In the case where $W(x, y) = y$, the last theorem reduces to a known result ([7, p. 128]) because, then, $|W[t, e(t)]| = E$ for every $t \in C(r)$. A theorem giving necessary conditions for best approximation is only possible under some further restrictions. We prove such a theorem using some of the ideas suggested by Moursund in [8, pp. 448–449]. First we prove a lemma which will be needed in the proof of Theorem 5.

LEMMA 4. *Suppose* $\epsilon > 0$, $\delta > 0$, $t \in X$, $V = U'(t, \delta/2)$, $u \in C(X)$ *and* $M(u) = E < \infty$.

(3a) *If* $u(x) - 2\epsilon \geqslant 0$ *for all* $x \in U(t, \delta)$, *then* $\sup_V W[x, u(x) - \epsilon] < E$.

(3b) *If* $u(x) + 2\epsilon \leqslant 0$ *for all* $x \in U(t, \delta)$, *then* $\inf_V W[x, u(x) + \epsilon] > -E$.

(3c) *If* $|W[x, u(x) \pm 2\epsilon]| \leqslant E$ *and* $u(x) - \epsilon \leqslant h(x) \leqslant u(x) + \epsilon$ *for all* $x \in U(t, \delta)$, *then* $\sup_V |W[x, h(x)]| < E$.

Proof.

(a) Suppose $u(x) - 2\epsilon \geqslant 0$ for all $x \in U(t, \delta)$. Let h^- be defined as in Lemma 3. Then

$$\begin{aligned} &0 < h^-(x) \leqslant u(x) \qquad \text{for} \quad x \in U(t, \delta), \\ &u(x) = h^-(x) \qquad \text{for} \quad x \in X \setminus U(t, \delta). \end{aligned}$$

Consequently, using property (1b), $M(h^-) \leqslant E$. Applying property (1c), with g replaced by h^-, and Y by V, we get, using (1a),

$$\sup_V W[x, u(x) - \epsilon] < \max\{\sup_V W[x, u(x)], \sup_V [x, u(x) - 2\epsilon]\}.$$

Because $0 \leqslant u(x) - 2\epsilon \leqslant u(x)$ for all $x \in V$, we get

$$\sup_V W[x, u(x) - \epsilon] < \sup_V W[x, u(x)] \leqslant E.$$

(b) Suppose $u(x) + 2\epsilon \leqslant 0$ for all $x \in U(t, \delta)$. Defining h^+ as in lemma 3, we obtain $M(h^+) \leqslant E$, and as in (a),

$$\sup_V | W[x, u(x) + \epsilon]| < \sup_V | W[x, u(x)]| \leqslant E.$$

Because $u(x) < u(x) + \epsilon < 0$ for all $x \in V$, we get, using property (1a),

$$\inf_V W[x, u(x) + \epsilon] > -E.$$

(c) Part (3c) follows immediately from Lemma 3 and property (1b) of $W(x, y)$.

THEOREM 5. *If $r = p/q$ is a best approximation to f, with $M(r - f) = E > 0$, and if for each $t \in C(r)$ with $r(t) = f(t)$, there exist constants $\delta_1 > 0$, $N > 0$, $s > 0$ (which may depend on t) such that $x \in U(t, \delta_1)$, $x \neq t$ imply*

$$| W(x, y)| \leqslant N \cdot | y |^s, \tag{3.4}$$

then there exists no $v \in P + rQ$ satisfying:

$$\begin{array}{ll} v(t) < 0 & \text{for every plus extremal } t \in C(r) \text{ with } e(t) > 0; \\ v(t) \leqslant 0 & \text{for every plus extremal } t \in C(r) \text{ with } e(t) = 0; \\ v(t) > 0 & \text{for every minus extremal } t \in C(r) \text{ with } e(t) < 0; \\ v(t) \geqslant 0 & \text{for every minus extremal } t \in C(r) \text{ with } e(t) = 0; \\ v(t) = 0 & \text{for every zero extremal } t \in C(r). \end{array} \tag{3.5}$$

Proof. Suppose there exists an element $v = p_0 + rq_0 \in P + rQ$ satisfying (3.5). As in ordinary uniform approximation ([2, p. 159]) we seek a constant $k > 0$ such that $M(r_k - f) < E$, with $r_k = (p + kp_0)/(q - kq_0)$. To this end, we show that for each $t \in X$ there exist constants $\delta_t > 0$, $m_t > 0$ and a function $E_1(t, k)$ such that if $x \in U(t, \delta_t)$ and $k \in (0, m_t]$ then

$$|W[x, r_k - f(x)]| \leqslant E_1(t, k) < E.$$

Observe that $r_k = r + k \cdot v/(q - kq_0)$. Set

$$T = \max_{x \in X} |v(x)|; \qquad n_0 = \max_{x \in X} |q_0(x)|; \qquad m_0 = \min_{x \in X} q(x) > 0.$$

(a) Let t be a plus extremal of $C(r)$ with $e(t) = r(t) - f(t) \neq 0$. By Theorem 1, $e(t) = A > 0$; therefore, by (3.5), $v(t) = B < 0$. By continuity, there exist $\delta_1 > 0$, $\delta_2 > 0$ such that

$$\begin{aligned} x \in U(t, \delta_1) &\quad \text{implies} \quad A/2 < r(x) - f(x), \\ x \in U(t, \delta_2) &\quad \text{implies} \quad v(x) < B/2. \end{aligned}$$

Let $\delta = \min(\delta_1, \delta_2)$, $V = U'(t, \delta/2)$,

$$n_1 = \max\{n_0, 1\}, \quad m_1 = \frac{m_0}{2n_1}; \qquad \text{and} \qquad T_1 = m_0 - m_1 \cdot n_0 .$$

Then $T_1 \geqslant m_0/2 > 0$. If we set

$$m_t = \min\left\{\frac{A \cdot T_1}{2 \cdot T}, m_1\right\}, \qquad \text{then} \qquad m_t > 0, \qquad \text{and,}$$

for $k \in (0, m_t]$ and $x \in V$, we have

$$r_k(x) - f(x) > \frac{A}{2} - k \cdot \frac{T}{T_1} \geqslant 0. \tag{3.6}$$

Set

$$T_2 = [\sup_{x \in U(t,\delta)} q(x)] + m_1 \cdot n_0 \qquad \text{and} \qquad \epsilon = \frac{k \cdot |B|}{2 \cdot T_2}.$$

For $k \in (0, m_t]$ and $x \in U(t, \delta)$, we have

$$r_k(x) - f(x) \leqslant r(x) - f(x) - \epsilon, \tag{3.7}$$

$$r(x) - f(x) - 2\epsilon > 0. \tag{3.8}$$

Using (3a) of Lemma 4, with $u = e$, (3.8) implies

$$\sup_V W[x, r(x) - f(x) - \epsilon] < E. \tag{3.9}$$

From (3.6), (3.7), and property (1b) it follows that if $x \in V$ and $k \in (0, m_t]$,

$$0 < W[x, r_k(x) - f(x)] \leqslant \sup_V W[x, e(x) - \epsilon]. \tag{3.10}$$

Set $E_1(t, k) = \sup_V W[x, r_k(x) - f(x)]$ and $\delta_t = \delta/2$. Then (3.9) and (3.10) imply

$$|W[x, r_k(x) - f(x)]| \leqslant E_1(t, k) < E \quad \text{for } x \in U(t, \delta_t) \quad \text{and} \quad k \in (0, m_t]. \tag{3.11}$$

(b) Let t be a minus extremal of $C(r)$ with $e(t) \neq 0$. From Theorem 1 and (3.5) follows, $e(t) = A < 0$ and $v(t) = B > 0$. In the same way as in (a), using (3b) of Lemma 4, we get, there exist $\delta_t > 0$, $m_t > 0$ and $E_1(t, k)$ such that

$$|W[x, r_k(x) - f(x)]| \leqslant E_1(t, k) < E \quad \text{for } x \in U(t, \delta_t) \quad \text{and} \quad k \in (0, m_t]. \tag{3.12}$$

(c) Let t be a point of $C(r)$ with $e(t) = 0$ and $v(t) = 0$. In this case, (3.4) is satisfied. With $S = (E/2N)^{1/s}$,

$$|y| \leqslant S \quad \text{implies} \quad |W(x, y)| \leqslant E/2, \tag{3.13}$$

for $x \in U(t, \delta_1)$, $x \neq t$. Because $e(t) = 0$, there exists $\delta_2 > 0$ such that $x \in U(t, \delta_2)$ implies $|e(x)| < S/2$. Let δ, T_1, V and m_1 be defined as in (a). Set

$$m_t = \min \left\{\frac{S \cdot T_1}{2 \cdot T}, m_1\right\};$$

then $x \in V$ and $k \in (0, m_t]$ imply

$$|r_k(x) - f(x)| \leqslant |e(x)| + \frac{k\,|v(x)|}{q(x) - k \cdot q_0(x)},$$

$$|r_k(x) - f(x)| < S/2 + k \cdot \frac{T}{T_1} \leqslant S.$$

Set $E_1(t, k) = E/2$, $\delta_t = \delta/2$, and note that $r_k(t) - f(t) = 0$. Using (3.13) and property (1a), we have

$$|W[x, r_k(x) - f(x)]| \leqslant E_1(t, k) < E \quad \text{for } x \in U(t, \delta_t) \quad \text{and} \quad k \in (0, m_t]. \tag{3.14}$$

(d) Let t be a point of $C(r)$ with $e(t) = 0$ and $v(t) \neq 0$. Suppose that $v(t) > 0$ (the case $v(t) < 0$ can be handled in the same way). Then because of (3.5), the point t is a minus extremal. Definition (2c) implies that there exists $\epsilon > 0$ such that

$$W[t, \epsilon] = W[t, e(t) + \epsilon] < E. \tag{3.15}$$

We proceed now in the same way as in (c), but with

$$m_t = \min\left\{\frac{S \cdot T_1}{2 \cdot T}, m_1, \frac{\epsilon \cdot T_1}{T}\right\}.$$

Then, for $k \in (0, m_t]$ and $x \in U(t, \delta)$, $x \neq t$, we have

$$|W[x, r_k(x) - f(x)]| \leqslant E/2. \tag{3.16}$$

Since, for $k \in (0, m_t]$,

$$0 < r_k(t) - f(t) \leqslant e(t) + k \cdot \frac{T}{T_1} < \epsilon,$$

we have, using (3.15) and property (1b),

$$W[t, r_k(t) - f(t)] < E. \tag{3.17}$$

Setting $E_1(t, k) = \max\{E/2, W[t, r_k(t) - f(t)]\}$ and $\delta_t = \delta/2$, (3.16) and (3.17) imply

$$|W[x, r_k(x) - f(x)]| \leqslant E_1(t, k) < E \quad \text{for } x \in U(t, \delta_t) \quad \text{and} \quad k \in (0, m_t]. \tag{3.18}$$

(e) Let $t \in X \setminus C(r)$. Because $t \notin C(r)$, there exist $\delta > 0$ and $\epsilon > 0$ such that

$$x \in U(t, \delta) \qquad \text{implies} \quad |W[x, e(x) \pm \epsilon]| \leqslant E. \tag{3.19}$$

Let T_1, V and m_1 be defined as in (a). We take now

$$m_t = \min\left\{\frac{\epsilon \cdot T_1}{2 \cdot T}, m_1\right\}.$$

If $x \in U(t, \delta)$ and $k \in (0, m_t]$ then

$$r(x) - f(x) - \frac{\epsilon}{2} \leqslant r_k(x) - f(x) \leqslant r(x) - f(x) + \frac{\epsilon}{2}.$$

Set $E_1(t, k) = \sup_V |W[x, r_k(x) - f(x)]|$ and $\delta_t = \delta/2$; then, using (3.19) and (3c) of Lemma 4, we get

$$|W[x, r_k(x) - f(x)]| \leqslant E_1(t, k) < E \quad \text{for } x \in U(t, \delta_t) \quad \text{and} \quad k \in (0, m_t], \tag{3.20}$$

(f) Conclusion. From (3.11), (3.12), (3.14), (3.18), and (3.20) follows: for each $t \in X$ there exist constants $\delta_t > 0$, $m_t > 0$, and a function $E_1(t, k)$

such that $x \in U(t, \delta_t)$ and $k \in (0, m_t]$ imply $|W[x, r_k(x) - f(x)]| \leqslant E_1(t, k) < E$. The set of neighborhoods $U(t, \delta_t)$ form an open cover of X. Since X is a compact metric space, there exists a finite subcover, say,

$$\{U(t_i, \delta_{t_i}) : i = 1, 2, ..., p\}.$$

If $k_1 = \min_{1 \leqslant i \leqslant p} m_{t_i}$ and $E_1 = \max_{1 \leqslant i \leqslant p} E_1(t_i, k_1)$, then $k_1 > 0$ and $E_1 < E$. For $k \in (0, k_1]$ and all $x \in X$, we have then

$$|W[x, r_k(x) - f(x)]| \leqslant E_1 < E.$$

This means that such r_k are better approximations to f than r, contradicting the hypothesis that r is a best approximation to f. This concludes the proof.

One can combine Theorems 4 and 5, to obtain the following:

COROLLARY 1. *If r is an approximation to f, with $M(r - f) = E > 0$, and if*

(i) *$W(x, y)$ satisfies a condition of the form* (3.4) *at each $t \in C(r)$ with $r(t) = f(t)$;*

(ii) *$|W[t, r(t) - f(t)]| = E$ for each $t \in C(r)$ with $r(t) \neq f(t)$;*

then r is a best approximation to f if and only if there exists no $v \in P + rQ$ such that (3.5) *holds.*

We remark that Theorems 4 and 5 are different from the corresponding result in [8, Theorem 10, p. 447]; see also [17]. As an illustration of our results we give two examples.

EXAMPLE 1. Suppose $f(x) = x$ and $X = [1/2, 2]$. Let P be the linear subspace spanned by $p(x) = 1$, and Q the linear subspace spanned by $q(x) = x$. Then r must be of the form α/x, where α is a constant. We want to approximate f so that $r(2) - f(2) \geqslant -1$. The corresponding generalized weight function $W(x, y) = y$ if $x \neq 2$, $W(2, y) = y$ if $y \geqslant -1$, and $W(x, y) = -\infty$ if $y < -1$. Consider $r(x) = 2/x$. The point $x_1 = 1/2$ is a plus extremal and $x_2 = 2$ is a minus extremal because $W[x_1, e(x_1)] = M(r - f) = 7/2$ and $W[x_2, e(x_2) - \epsilon] = -\infty$, for each $\epsilon > 0$. It is clear that the approximation $r(x) = 2/x$ is a best approximation. This is in agreement with Theorems 4 and 5 because there exists no $v \in P + rQ$ (a constant function) such that $v(x_1) < 0$ and $v(x_2) \geqslant 0$.

EXAMPLE 2. *Let* $f(x) = (1 + x)/2$ and $X = [-1, 1]$. Let us approximate f by a rational function of the form $r(x) = ax/(b + cx)$, constrained by $r(0) - f(0) \geqslant -1/2$. The corresponding generalized weight function

$W(x, y) = y$ if $x \neq 0$, $W(0, y) = y$ if $y \geqslant -1/2$, and $W(0, y) = -\infty$ if $y < -1/2$. Consider $r(x) = 4x/(3 - x)$, for which $M(r - f) = 1$. The points $x_1 = -1$ and $x_2 = 0$ are minus extremals; the point $x_3 = 1$ is a plus extremal. There exists an element $v \in P + rQ$, namely, $v(x) = -x - 8x/(3 - x)$ such that $v(1) < 0$, $v(0) \geqslant 0$, $v(-1) > 0$. According to Theorem 3, r is not a best approximation; indeed, a better one is

$$r_1(x) = \frac{x}{2}, \qquad \text{with} \qquad M(r_1 - f) = 1/2.$$

4. Equivalent Statements of the Characterization Theorems

It is possible to give equivalent statements for Theorems 4 and 5, under some restrictions on $P + rQ$ and $C(r)$, using methods of the theory of ordinary uniform approximation. First we prove a lemma, to be used in the next theorem. Set $C^0 = \{t : t \in C(r) \text{ and } r(t) = f(t)\}$ and

$$C' = \{t : t \in C(r) \text{ and } r(t) \neq f(t)\}.$$

Lemma 5. *If r is an approximation to f with $M(r - f) = E > 0$, and if $W(x, y)$ satisfies a condition of the form* (3.4) *for every $t \in C^0$, then C' is closed in X and not empty.*

Proof. First we prove that C' is not empty. If C' were empty then we would have for every $t \in C(r)$, $W[t, r(t) - f(t)] = 0$, because of property (1a) of $W(x, y)$. For such t, there exist $\delta_{1t} > 0$, $N > 0$, and $s > 0$ such that if $|y| \leqslant S$ (where $S = (2N/E)^{-1/s}$, $x \in U(t, \delta_{1t})$ and $x \neq t$, we have $|W(x, y)| \leqslant E/2$ [because of condition (3.4)]. For every $t \in C(r)$ there exists $\delta_{2t} > 0$ such that $|r(x) - f(x)| < S$ for all $x \in U(t, \delta_{2t})$. Set $\delta_t = \min\{\delta_{1t}, \delta_{2t}\}$; then $|W[x, r(x) - f(x)]| \leqslant E/2 = d_0(t)$ whenever $x \in U(t, \delta_t)$ and $x \neq t$. If $t \notin C(r)$ then according to Theorem 2 there exist $\delta_t > 0$, $\epsilon_t > 0$ and $d_0(t) > 0$ such that $|W[x, r(x) - f(x) \pm \epsilon_t]| \leqslant d_0(t) < E$ for every $x \in U(t, \delta_t)$. Consider now all the neighborhoods $U(t, \delta_t)$, $t \in X$. They form an open cover of X. Since X is a compact metric space, there exists a finite subcover, say, $\{U(t_i, \delta_{t_i}) : i = 1, 2, \ldots, p\}$. Put $d_1 = \max_{1 \leqslant i \leqslant p} d_0(t_i)$; then $d_1 < E$. It follows that $|W[x, r(x) - f(x)]| \leqslant d_1$ for every $x \in X$, i.e., $M(r - f) \leqslant d_1 < E$. This contradicts $M(r - f) = E$; thus, C' can not be empty.

We prove now that C' is closed. To see this, note that every $t \in C^0$ is an isolated point of $C(r)$. Suppose $t_0 \in C^0$ is not isolated in $C(r)$. From the continuity of r and f in X follows the existence of $\delta_3 > 0$ such that $x \in U(t_0, \delta_3)$ implies $|r(x) - f(x) \pm S/2| < S$. Set

$$\delta' = \min\{\delta_{1t_0}, \delta_3\} \qquad \text{and} \qquad \epsilon = \frac{S}{2}.$$

Then

$$|W[x, r(x) - f(x) \pm \epsilon]| \leqslant \frac{E}{2} \qquad \text{for every} \quad x \in U(t_0, \delta'), \quad x \neq t_0 .$$

There exists a point $t_1 \in C(r)$ such that $t_1 \in U(t_0, \delta'/2)$ and $t_1 \neq t_0$. Let $\delta_4 = d(t_0, t_1)$, the distance between t_0 and t_1 in X; then for all $x \in U(t_1, \delta_4)$ we have $|W[x, r(x) - f(x) \pm \epsilon]| \leqslant E/2$. This contradicts the fact that $t_1 \in C(r)$. Thus, every point of C^0 is an isolated point of $C(r)$. From this and Theorem 3 it clearly follows that C' is closed.

COROLLARY 2. *Set*

$$C^+ = \{t : t \in C' \text{ and } t \text{ is a plus extremal}\}$$

and

$$C^- = \{t : t \in C' \text{ and } t \text{ is a minus extremal}\}.$$

Then under the conditions of Lemma 5, the sets C^+ and C^- are closed.

Proof. Suppose $t = \lim_{i\to\infty} t_i$ with $t_i \in C^+$. Using Lemma 5, we have $t \in C'$, and, because of Theorem 1, $r(t_i) - f(t_i) > 0$. By continuity,

$$r(t) - f(t) \geqslant 0.$$

Since $t \in C'$, we must have $r(t) - f(t) > 0$ and, therefore, t is a plus extremal. Consequently, C^+ is closed. In the same way we can prove that C^- is closed.

In the rest of this paper we assume:

(i) $P + rQ$ is a Haar subspace of dimension k; this means that 0 is the only function in $P + rQ$ which has k or more zeros in X;

(ii) The number m of points in C^0 is at most $k - 1$;

(iii) $W(x, y)$ satisfies a condition of the form (3.4) for every $t \in C^0$;

(iv) C^0 contains only zero extremals and $|W[t, r(t) - f(t)]| = E$ for every $t \in C'$.

THEOREM 6. *If r is a best approximation to f with $M(r - f) = E > 0$ and if the assumptions* (i)–(iii) *are satisfied then there exist $n \leqslant k - m + 1$ (distinct) points $t_j \in C'$, and positive numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that*

$$\sum_{j=1}^{n} \lambda_j \cdot [r(t_j) - f(t_j)] \cdot v(t_j) = 0 \qquad \text{for every } v \in K, \tag{4.1}$$

where $K = \{v : v \in P + rQ, \text{ and } v(t) = 0 \text{ if } t \in C^0\}$.

Proof. From Theorem 5 follows: there exists no $v \in K$ with

$$v(t) \cdot [f(t) - r(t)] > 0 \qquad \text{for every} \quad t \in C'.$$

Using a Theorem on linear inequalities ([2, p. 19]) and a Theorem of Carathéodory ([2, p. 17]), the result follows. We note, however, that in order to be able to apply the Theorem on linear inequalities, the set C' must be compact. The compactness of C' follows from Lemma 5 and the compactness of X. Note also that the set K is not empty because of the conditions (i) and (ii).

THEOREM 7. *If r is an approximation to f with $M(r - f) = E > 0$, and if the assumptions* (i), (ii), (iv) *are satisfied, then r is a best approximation to f if there exist $n \leqslant k - m + 1$ (distinct) points $t_1, t_2, \dots, t_n$ in C', and positive numbers $\lambda_1, \lambda_2, \dots, \lambda_n$ such that* (4.1) *is satisfied.*

Proof. We may assume that $\sum_{j=1}^{n} \lambda_j = 1$. Let $a(x)$ be the vector with components $a_i(x) = [f(x) - r(x)] \cdot v_i(x)$, $i = 1, 2 \cdots p$, where $\{v_1, v_2 \cdots v_p\}$ is a basis for K. Note that $p = k - m$. Let $T = \{t_j : j = 1, 2 \cdots n\}$ and $Z = \{a(t) : t \in T\}$. From (4.1) follows: $0 = \sum_{j=1}^{n} \lambda_j \cdot a(t_j)$; therefore 0 [the origin of (real) Euclidean p-space] belongs to the convex hull of Z. Using the above Theorem on linear inequalities and the compactness of Z, we get that there exists no $v \in K$ with $[f(t) - r(t)] \cdot v(t) > 0$ for every $t \in T$. Therefore, there exists no $v \in P + rQ$ which satisfies (3.1), because of the conditions (iv). According to Theorem 4, r is a best approximation to f.

It is possible to combine Theorems 6 and 7, if we suppose that the conditions (i)–(iv) hold. We state the result as

COROLLARY 3. *Suppose r is an approximation to f with $M(r - f) = E > 0$ and suppose the assumptions* (i)–(iv) *are satisfied. Then r is a best approximation to f if and only if there exist $n \leqslant k - m + 1$ (distinct) points $t_j \in C'$, and positive numbers $\lambda_1, \lambda_2, \dots, \lambda_n$ such that* (4.1) *holds.*

This corollary can be formulated in several equivalent forms if $X = [a, b]$. For these, we refer the reader to [3, pp. 65–70]. The situation here and there is quite similar, although our theory is more general. We mention here just two of these equivalent statements.

COROLLARY 4. *Suppose that the conditions of Corollary* 3 *are satisfied, with $X = [a, b]$. Then r is a best approximation to f if and only if the origin of (real) Euclidean p-space is in the convex hull of the set $\{[r(t) - f(t)] \cdot \hat{t} : t \in C'\}$, where $\hat{t} = (v_1(t), \dots, v_p(t))$, and $\{v_1, \dots, v_p\}$ is any basis of K.*

COROLLARY 5. *Suppose that the conditions of Corollary* 3 *are satisfied, with* $X = [a, b]$. *Then* r *is a best approximation to* f *if and only if there exist in* C' *points* $t_1, t_2, ..., t_n$ $(n = k - m + 1)$ *satisfying*

$$\operatorname{sgn}([r(t_i) - f(t_i)] \cdot D_i) = (-1)^{i+1} \operatorname{sgn}([r(t_1) - f(t_1)] \cdot D_1)$$

for $i = 2, 3, ..., n$, *where*

$$D_i = \begin{vmatrix} v_1(t_1) \cdots v_1(t_{i-1}) & v_1(t_{i+1}) \cdots v_1(t_n) \\ v_2(t_1) \cdots v_2(t_{i-1}) & v_2(t_{i+1}) \cdots v_2(t_n) \\ \vdots \qquad \vdots & \vdots \qquad \vdots \\ v_p(t_1) \cdots v_p(t_{i-1}) & v_p(t_{i+1}) \cdots v_p(t_n) \end{vmatrix}$$

and $\{v_1, v_2 \cdots v_p\}$ *is any basis for* K $(p = k - m)$.

We remark that Corollary 5 includes the classical alternation theorem if $m = 0$, because then all D_i have the same sign. To illustrate the above results, we give two examples.

EXAMPLE 3. Suppose $X = [0, 4]$ and

$$f(x) = \begin{cases} -2 + 2x & \text{if } x \in [0, 2], \\ 6 - 2x & \text{if } x \in [2, 3], \\ -6 + 2x & \text{if } x \in [3, 4]. \end{cases}$$

Let us approximate f by a function of the form $r(x) = a + bx$ with $r(2) = f(2)$. The corresponding generalized weight function $W(x, y)$ for this problem satisfies

$$W(x, y) = y \quad \text{if } x \neq 2; \qquad W(2, y) = 0 \quad \text{if } y = 0;$$
$$W(2, y) = \infty \quad \text{if } y > 0; \qquad W(2, y) = -\infty \quad \text{if } y < 0.$$

A best approximation to f is $r(x) = (2 + 2x)/3$. We have $M(r - f) = 8/3$ and $C(r) = \{0, 2, 3\}$, $C' = \{0, 3\}$. Put $t_1 = 0$, $t_2 = 3$. Then if $\lambda_1 = 1$ and $\lambda_2 = 2$, Theorems 6 and 7 can be applied. Take $v_1(x) = -2 + x$; then $D_1 = 1$, $D_2 = -2$, and Corollary 5 is also applicable.

EXAMPLE 4. Here we illustrate the fact that, in Theorem 7, the condition that C^0 contains only zero-extremals is essential. Let $f(x) = x$ in $[0, 1]$, $f(x) = 2 - x$ in $[1, 2]$, and $f(x) = -2 + x$ in $[2, 3]$. Let us approximate f by a function of the form $r(x) = a + bx$ with $r(1) - f(1) \leqslant 0$. The corresponding generalized weight function $W(x, y)$ equals y if $x \neq 1$, $W(1, y) = y$ if $y \leqslant 0$, and $W(1, y) = \infty$ if $y > 0$. Consider $r_1(x) = 1$ in $X = [0, 3]$. Then $M(r_1 - f) = 1$, $C(r_1) = \{0, 1, 2\}$ and $C' = \{0, 2\}$. For $t_1 = 0$, $t_2 = 2$, $\lambda_1 = \lambda_2 = 1$, (4.1) is satisfied. One could try to apply Theorem 7 or

Corollary 5 (with $v_1(x) = 1 - x$, $D_1 = -1$, $D_2 = 1$). The conclusion that r_1 is a best approximation is, however, wrong because $r_2(x) = 1/2$ is a better approximation. The reason for the unavailability of Theorem 7 and Corollary 5 in this case is that C^0 contains a plus extremal ($t = 1$), and, thus, condition (iv) is not satisfied.

Acknowledgment

The author thanks Professor L. Buyst for his advice, criticism, and encouragement during work on this paper. He also acknowledges gratefully the Editor's helpful suggestions for improving the exposition of this paper.

References

1. P. Alexandroff and H. Hopf, "Topologie," Springer-Verlag, Berlin, 1935.
2. E. W. Cheney, "Introduction to Approximation Theory," McGraw-Hill, New York, 1966.
3. F. Deutsch, On Uniform Approximation with Interpolatory Constraints, *J. Math. Anal. Appl.* **24** (1968), 62–79.
4. J. Dieudonné, "Foundations of Modern Analysis," Academic Press, New York, 1960.
5. H. L. Loeb, D. G. Moursund, and G. D. Taylor, Uniform Rational Weighted Approximations Having Restricted Ranges, *J. Approximation Theory* **1** (1968), 401–411.
6. H. L. Loeb, D. G. Moursund, L. L. Schumaker, and G. D. Taylor, Uniform generalized weight function polynomial approximation with interpolation, *SIAM J. Numer. Anal.* **6** (1969), 284–293.
7. G. Meinardus, "Approximation von Funktionen und ihre numerische Behandlung," Springer-Verlag, Berlin, 1964.
8. D. G. Moursund, Chebyshev approximation using a generalized weight function, *SIAM J. Numer. Anal.* **3** (1966), 435–450.
9. D. G. Moursund, Optimal starting values for Newton–Raphson calculation of $\sqrt{x}$, *Comm. ACM.* **10** (1967), 430–432.
10. D. G. Moursund and G. D. Taylor, Optimal starting values for the Newton–Raphson calculation of inverses of certain functions, *SIAM J. Numer. Anal.* **5** (1968), 138–150.
11. D. G. Moursund and G. D. Taylor, Uniform rational approximation using a generalized weight function, *SIAM J. Numer. Anal.* **5** (1968), 882–889.
12. L. L. Schumaker and G. D. Taylor, On approximation by polynomials having restricted ranges. II, *SIAM J. Numer. Anal.* **6** (1969), 31–36.
13. G. D. Taylor, On approximation by polynomials having restricted ranges, *SIAM J. Numer. Anal.* **5** (1968), 258–268.
14. G. D. Taylor, Approximation by functions having restricted ranges. III, *J. Math. Anal. Appl.* **27** (1969), 241–248.
15. G. D. Taylor, Approximation by functions having restricted ranges: equality case, *Numer. Math.* **14** (1969), 71-78.
16. L. Wuytack, The existence of a solution in constrained rational approximation problems, *Simon Stevin* **43** (1969/70), 83–99.
17. L. Wuytack, Some remarks on a paper of D. G. Moursund, *SIAM J. Numer. Anal.* **7** (1970), 233-237.