JOURNAL OF APPROXIMATION THEORY 32, 9–26 (1981)

Uniform Approximation by Rational Functions Having Restricted Denominators

E. H. KAUFMAN, JR.

Department of Mathematics,
Central Michigan University, Mount Pleasant, Michigan 48859

AND

G. D. TAYLOR*

Department of Mathematics,
Colorado State University, Fort Collins, Colorado 80523

Communicated by Richard S. Varga

Received June 25, 1979

This paper considers approximation of continuous functions on a compact metric space by generalized rational functions for which the denominators have bounded coefficients and are bounded below by a fixed positive function. This lower bound alleviates numerical difficulties, and in some applications (e.g., digital filter design) has a useful physical interpretation. A "zero in the convex hull" characterization of best approximations is developed and used to prove uniqueness and de la Vallée Poussin results. Examples are given to illustrate this theory and its differences with the standard theory, where the denominators are merely required to be positive. A modified differential correction algorithm is presented and is proved to always converge at least linearly, and often quadratically.

1. INTRODUCTION

In this paper we consider approximation of continuous functions by generalized rational functions whose denominators are required to be bounded away from zero. This is in contrast to the standard theory, where the denominators are only required to be positive. There are at least four reasons for having this stronger requirement:

(1) Best approximations always exist.

* Supported in part by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant F-49620-79-C-0124, and by the National Science Foundation under Grant MCS-78-05847.

0021-9045/81/050009-18$02.00/0

(2) Iterative procedures for computing rational approximations may experience numerical difficulties if the denominators of the rational functions become very small.

(3) Even if a good rational approximation is found, it may not be very useful if its denominator is too small at some point.

(4) There may be some physical advantage in being able to control the denominator. For example, McCallig [6] used a version of the algorithm described in this paper to compute approximations to the desired magnitude-squared response of a digital filter; control of the denominator of the rational function amounts to control of the feedback gain of the resulting filter, which allows one to progress smoothly from "fully recursive" filter designs to nonrecursive designs, and to reduce sensitivity problems and hardware requirements.

Formally, the situation we are considering is as follows. X is a compact metric space, m and n are fixed positive integers, $\mathscr{P}$ and $\mathscr{Q}$ are subspaces of $C[X]$ with bases $\{\theta_1,..., \theta_m\}$ and $\{\psi_1,..., \psi_n\}$, respectively, and L is a strictly positive continuous function on X (which is often a constant in practice). Our family of approximating functions is then defined to be

$$\mathscr{R}_L = \{P/Q: P = p_1\theta_1 + \cdots + p_m\theta_m \in \mathscr{P}, Q = q_1\psi_1 + \cdots + q_n\psi_n \in \mathscr{Q},$$
$$Q \geqslant L \text{ on } X, |q_j| \leqslant 1 \text{ for } j = 1,..., n\}.$$

We note in passing that without the restrictions $|q_j| \leqslant 1$, the restriction $Q \geqslant L$ would be no stronger than the more usual restriction $Q > 0$,, since it could always be satisfied by multiplying P and Q by a sufficiently large positive constant.

In order that $\mathscr{R}_L$ be nonempty, $\mathscr{Q}$ must contain at least one strictly positive function, so without loss of generality we will assume $\psi_1 > 0$ on X. To insure $\mathscr{R}_L \neq \varnothing$, and for other reasons which will be clearer later, we will also assume $\max_{x\in X} L(x) < \min_{x\in X} \psi_1(x)$; this requirement is no restriction in practice, since it can always be obtained by multiplying ψ_1 by a suitable positive constant.

Given $f \in C[X]$, a best approximation to f is defined to be a function $R^* \in \mathscr{R}_L$ such that $\|f - R^*\| \leqslant \|f - R\|$ for all $R \in \mathscr{R}_L$, where for $g \in C[X]$, $\|g\| = \max_{x\in X} |g(x)|$. The following theorem can be proved by standard techniques.

THEOREM 1. *If $f \in C[X]$, then there exists a best approximation to f from $\mathscr{R}_L$.*

In the remaining sections we will consider characterization of best approximations, uniqueness, de La Vallée Poussin results, and computation of

best approximations by a modified differential correction algorithm. The theory differs in some respects from the restricted range situation, where the entire rational function rather than just the denominator is restricted.

Suppose $f \in C[X]$ and $R^* \in \mathscr{R}_L$, where $R^* = P^*/Q^* = (p_1^*\theta_1 + \cdots + p_m^*\theta_m)/(q_1^*\psi_1 + \cdots + q_n^*\psi_n)$. The following notation will be useful.

$$\sigma(x) = \operatorname{sgn}(f(x) - R^*(x)) \qquad \forall x \in X;$$

$$\mathscr{P} + R^*\mathscr{Q} = \{P + R^*Q : P \in \mathscr{P}, Q \in \mathscr{Q}\};$$

$$X_0 = \{x \in X : |f(x) - R^*(x)| = \|f - R^*\|\};$$

$$Y_0 = \{y \in X : Q^*(y) = L(y)\};$$

$$I_0 = \{j \in \{1,\ldots, n\} : |q_j^*| = 1\};$$

$$S = \{\sigma(x)\,\hat{\mathbf{x}} : x \in X_0\} \cup \{\boldsymbol{\psi}(y) : y \in Y_0\} \cup \{q_j^*\mathbf{e}_{m+j} : j \in I_0\},$$

where

$$\hat{\mathbf{x}} = (\theta_1(x),\ldots, \theta_m(x), R^*(x)\,\psi_1(x),\ldots, R^*(x)\,\psi_n(x))^T,$$

$$\boldsymbol{\psi}(y) = (0,\ldots, 0, -\psi_1(y),\ldots, -\psi_n(y))^T$$

and

$$\mathbf{e}_k = (\delta_{1k},\ldots, \delta_{m+n,k})^T,$$

$$\delta_{ij} = \text{Kronecker delta};$$

$$\mathscr{K}(S) = \text{the convex hull of } S = \left\{ \sum_{i=1}^{k} \lambda_i s_i : \right.$$

$$\left. k \text{ is a positive integer, } s_i \in S\ \forall i, \lambda_i \geqslant 0\ \forall i, \sum_{i=1}^{k} \lambda_i = 1 \right\};$$

$$\operatorname{int}\mathscr{K}(S) = \text{the interior of } \mathscr{K}(S).$$

2. Characterization

We first prove a Kolmogorov-type characterization theorem.

Theorem 2. *Suppose $f \in C[X] - \mathscr{R}_L$. Then $R^* = P^*/Q^* \in \mathscr{R}_L$ is a best approximation to f iff there is no $\bar{P} = \bar{p}_1\theta_1 + \cdots + \bar{p}_m\theta_m \in \mathscr{P}$, $\bar{Q} = \bar{q}_1\psi_1 + \cdots + \bar{q}_n\psi_n \in \mathscr{Q}$ satisfying*

(i) $\operatorname{sgn}(\bar{P} + R^*\bar{Q})(x) = \operatorname{sgn}(f(x) - R^*(x))$, $\forall x \in X_0$;

(ii) $\bar{Q}(y) < 0$, $\forall y \in Y_0$;

(iii) $\bar{q}_j > 0$ *if* $q_j^* = 1$;

(iv) $\bar{q}_j < 0$ *if* $q_j^* = -1$.

Proof. ($\Leftarrow$) Suppose R^* is not a best approximation. Then there is a better approximation $R = P/Q \in \mathscr{R}_L$. By our assumptions in the previous section, we have $(\max_{x\in X} L(x))/(\min_{x\in X} \psi_1(x)) < 1$; letting a be any number satisfying $(\max_{x\in X} L(x))/(\min_{x\in X} \psi_1(x)) < a < 1$, we define $\tilde{Q} = \tilde{q}_1\psi_1 + \cdots + \tilde{q}_n\psi_n \in \mathscr{Q}$ by $\tilde{Q} = Q^* - a\psi_1$. Then $\tilde{Q}$ satisfies $\tilde{Q}(y) < 0 \quad \forall y \in Y_0$, $\tilde{q}_j > 0$ if $q_j^* = 1$, and $\tilde{q}_j < 0$ if $q_j^* = -1$. Now define $\bar{P} \equiv P - P^*$, $\bar{Q} \equiv Q^* - Q + \eta\tilde{Q}$, where η is a positive number. For $x \in X_0$, we have

$$\begin{aligned}
\operatorname{sgn}(\bar{P} + R^*\bar{Q})(x) &= \operatorname{sgn}(P - P^* + R^*(Q^* - Q + \eta\tilde{Q}))(x) \\
&= \operatorname{sgn}(P - R^*Q + \eta R^*\tilde{Q})(x) \\
&= \operatorname{sgn}\left\{Q(x)\left[\left(R(x) - R^*(x) + \eta\frac{R^*(x)\,\tilde{Q}(x)}{Q(x)}\right)\right]\right\} \\
&= \operatorname{sgn}\left[f(x) - R^*(x) - (f(x) - R(x)) + \eta\frac{R^*(x)\,\tilde{Q}(x)}{Q(x)}\right].
\end{aligned}$$

Now choosing η so small that $|\eta(R^*(x)\,\tilde{Q}(x)/Q(x))| < \|f - R^*\| - \|f - R\|$ $\forall x \in X_0$, we have

$$\operatorname{sgn}(\bar{P} + R^*\bar{Q})(x) = \operatorname{sgn}(f(x) - R^*(x)) \qquad \forall x \in X_0,$$

so (i) holds. For $y \in Y_0$, we have

$$\bar{Q}(y) = Q^*(y) - Q(y) + \eta\tilde{Q}(y) \leqslant L(y) - L(y) + \eta\tilde{Q}(y) < 0,$$

so (ii) holds. If $q_j^* = 1$, we have

$$\bar{q}_j = q_j^* - q_j + \eta\tilde{q}_j \geqslant 1 - 1 + \eta\tilde{q}_j > 0,$$

so (iii) holds. If $q_j^* = -1$, we have

$$\bar{q}_j = q_j^* - q_j + \eta\tilde{q}_j \leqslant -1 - (-1) + \eta\tilde{q}_j < 0,$$

so (iv) holds.

($\Rightarrow$) Suppose there exists $\bar{P} \in \mathscr{P}$, $\bar{Q} \in \mathscr{Q}$ satisfying (i)–(iv) above. Let λ be a small positive number. Then

$$\frac{P^* + \lambda\bar{P}}{Q^* - \lambda\bar{Q}} - \frac{P^*}{Q^*} = \lambda\frac{\bar{P}Q^* + P^*\bar{Q}}{(Q^* - \lambda\bar{Q})\,Q^*} = \lambda\frac{\bar{P} + R^*\bar{Q}}{Q^* - \lambda\bar{Q}}.$$

Thus

$$f-\frac{P^*+\lambda\bar{P}}{Q^*-\lambda\bar{Q}}=f-\frac{P^*}{Q^*}-\lambda\frac{\bar{P}+R^*\bar{Q}}{Q^*-\lambda\bar{Q}},$$

and using the arguments of [2, pp. 159, 160] it can be shown that

$$\left\|f-\frac{P^*+\lambda\bar{P}}{Q^*-\lambda\bar{Q}}\right\|<\left\|f-\frac{P^*}{Q^*}\right\|$$

for all λ sufficiently small. It remains only to show that

$$\frac{P^*+\lambda\bar{P}}{Q^*-\lambda\bar{Q}}\in\mathscr{R}_L$$

for λ sufficiently small.

For all $y\in Y_0$, we have

$$Q^*(y)-\lambda\bar{Q}(y)>Q^*(y)=L(y).$$

Since Y_0 is compact, there exists $\zeta>0$ such that $\bar{Q}(y)\leqslant-\zeta\ \forall y\in Y_0$. Let $Y_1=\{x\in X:\bar{Q}(x)<-\zeta/2\}$, $Y_2=X-Y_1$. Then Y_2 is compact, with $Y_2\cap Y_0=\varnothing$. Let $\mu=\min\{Q^*(x)-L(x):x\in Y_2\}>0$. Choose λ to satisfy $0<\lambda<\mu/\max(1,\|\bar{Q}\|)$. Then if $x\in Y_1$, we have $Q^*(x)-\lambda\bar{Q}(x)>Q^*(x)\geqslant L(x)$; if $x\in Y_2$, we have $Q^*(x)-\lambda\bar{Q}(x)\geqslant L(x)+\mu-\lambda\bar{Q}(x)\geqslant L(x)$, so $Q^*(x)-\lambda\bar{Q}(x)\geqslant L(x)\ \forall x\in X$. Finally, if $q_j^*=1$ we have

$$q_j^*-\lambda\bar{q}_j=1-\lambda\bar{q}_j<1,$$

and if $q_j^*=-1$ we have

$$q_j^*-\lambda\bar{q}_j=-1-\lambda\bar{q}_j>-1,$$

so choosing λ sufficiently small will insure that $|q_j^*-\lambda\bar{q}_j|\leqslant 1$, $j=1,\ldots,n$. Thus for λ sufficiently small we have

$$\frac{P^*+\lambda\bar{Q}}{Q^*-\lambda\bar{Q}}\in\mathscr{R}_L\qquad\text{and}\qquad\left\|f-\frac{P^*+\lambda\bar{Q}}{Q^*-\lambda\bar{Q}}\right\|<\left\|f-\frac{P^*}{Q^*}\right\|,$$

so P^*/Q^* is not a best approximation. Q.E.D.

We can now prove a "zero in the convex hull" characterization of best approximations.

THEOREM 3. *Suppose $f\in C[X]-\mathscr{R}_L$. Then $R^*=P^*/Q^*\in\mathscr{R}_L$ is a best approximation to f iff the origin of $(m+n)$-space lies in the convex hull of S (where S is the set defined in the Introduction).*

Proof. By the theorem on linear inequalities [2, p. 19], $\mathbf{0} \notin \mathscr{H}(S)$ iff the system of inequalities

$$\langle z, s \rangle > 0, \qquad s \in S,$$

is consistent (here $\langle\ ,\ \rangle$ denotes inner product). But this is true iff $\exists$ a vector $z = [z_1, \ldots, z_{m+n}]^T$ satisfying

$$\begin{aligned} \sigma(x)[z_1\theta_1(x) + \cdots + z_m\theta_m(x) + R^*(x)(z_{m+1}\psi_1(x) + \cdots + z_{m+n}\psi_n(x))] \\ > 0 \qquad \forall x \in X_0; \\ z_{m+1}\psi_1(y) + \cdots + z_{m+n}\psi_n(\,) < 0 \qquad \forall y \in Y_0; \\ z_{m+j} > 0 \qquad \text{if} \quad q_j^* = 1; \end{aligned}$$

and

$$z_{m+j} < 0 \qquad \text{if} \quad q_j^* = -1.$$

Letting $\bar{P} = z_1\theta_1 + \cdots + z_m\theta_m$ and $\bar{Q} = z_{m+1}\psi_1 + \cdots + z_{m+n}\psi_n$, we see by Theorem 2 that this is true iff R^* is not a best approximation. Thus $\mathbf{0} \in \mathscr{H}(S)$ iff R^* is a best approximation. Q.E.D.

We illustrate the application of this theorem with the following example.

EXAMPLE 1. Let $X = \{0, 1\}$, $f(x) = x$, $\mathscr{P} = \Pi_0 =$ the set of all polynomials of degree $\leqslant 0$, $\mathscr{Q} = \Pi_1$, $L(x) \equiv 0.1$, $R^*(x) = (1/11)/(1 - 0.9x)$. We have $X_0 = \{0, 1\}$, $Y_0 = \{1\}$, $I_0 = \{1\}$, $\sigma(0) = -1$, $\sigma(1) = 1$. Thus

$$\sigma(x)\begin{bmatrix} \theta_1(x) \\ (R^*\psi_1)(x) \\ (R^*\psi_2)(x) \end{bmatrix} = \sigma(x)\begin{bmatrix} 1 \\ \frac{1}{11}\Big/(1 - 0.9x) \\ \frac{x}{11}\Big/(1 - 0.9x) \end{bmatrix},$$

so

$$S = \left\{ \begin{bmatrix} -1 \\ -\frac{1}{11} \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ \frac{10}{11} \\ \frac{10}{11} \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

Solving the linear system

$$\begin{aligned} -\lambda_1 + \quad \lambda_2 \qquad\qquad &= 0, \\ -\frac{1}{11}\lambda_1 + \frac{10}{11}\lambda_2 - \lambda_3 + \lambda_4 &= 0, \\ \frac{10}{11}\lambda_2 - \lambda_3 \qquad &= 0, \\ \lambda_1 + \quad \lambda_2 + \lambda_3 + \lambda_4 &= 1 \end{aligned}$$

yields the solution $\lambda_1 = 1/3$, $\lambda_2 = 1/3$, $\lambda_3 = 10/33$, $\lambda_4 = 1/33$. Since we have $\lambda_i \geqslant 0\ \forall i$, we have $\mathbf{0} \in \mathscr{H}(S)$, so R^* is a best approximation.

We may observe that in this example $\mathbf{0}$ is a positive convex combination of exactly $m + n + 1 = 4$ vectors in S, and the coefficient matrix used is nonsingular. Thus by Cramer's rule $(\delta_1, \delta_2, \delta_3)^T \in \mathscr{H}(S)$ if $|\delta_1|, |\delta_2|, |\delta_3|$ are sufficiently small, so $\mathbf{0}$ is actually in the *interior* of $\mathscr{H}(S)$; this distinction will be important in the next two sections.

The next example illustrates what can happen if our assumption that $\max_{x \in X} L(x) < \min_{x \in X} \psi_1(x)$ is violated.

EXAMPLE 2. Let $X = [0, 1]$, $\mathscr{P} = \Pi_0$, $\mathscr{Q} = \Pi_1$, $L(x) \equiv 1$, $R^*(x) = 1/(1 + 0.5x)$. We have

$$\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \in S \qquad \text{and} \qquad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \in S,$$

so $\mathbf{0} \in \mathscr{H}(S)$ regardless of whether R^* is a best approximation to f or not. Intuitively, the trouble is that we are imposing a double restraint on the denominator at $x = 0$ which ties it down completely there.

The next example shows that the standard alternation characterization of best approximations does not hold in our setting.

EXAMPLE 3. Let $X = [0, 3]$,

$$\begin{aligned} f(x) &= 3.5x, && 0 \leqslant x \leqslant 1, \\ &= 2 + 1.5x, && 1 \leqslant x \leqslant 2, \\ &= 8 - 1.5x, && 2 \leqslant x \leqslant 3, \end{aligned}$$

$\mathscr{P} = \Pi_0$, $Q = \Pi_2$, $L(x) \equiv 0.2$, $R^*(x) = 1/(1 - 0.8x + 0.2x^2)$. We have $X_0 = \{0, 1, 3\}$, $Y_0 = \{2\}$, $I_0 = \{1\}$, $\sigma(0) = -1$, $\sigma(1) = 1$, $\sigma(3) = 1$. The fact that R^* is a best approximation to f is shown by the equality

$$\frac{4}{17}\begin{bmatrix}-1\\-1\\0\\0\end{bmatrix}+\frac{3}{17}\begin{bmatrix}1\\2.5\\2.5\\2.5\end{bmatrix}+\frac{1}{17}\begin{bmatrix}1\\2.5\\7.5\\22.5\end{bmatrix}+\frac{15}{34}\begin{bmatrix}0\\-1\\-2\\-4\end{bmatrix}$$

$$+\frac{3}{34}\begin{bmatrix}0\\1\\0\\0\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}.$$

Without denominator restrictions we would expect $(m-1)+(n-1)+2-\min(m-1-\text{degree of } P^*,\ n-1-\text{degree of } Q^*)=$ four alternating points for $f-R^*$; since there is only one denominator restriction we might have hoped for three alternating extreme points, but there are only two. It is tempting to conjecture that denominator constraints act as extreme points with "sign of error" opposite that of the previous extreme point, but there are examples which show that there may still be fewer than expected alternating extreme points. Thus there does not appear to be a simple alternation theorem in this setting. Some partial results under suitable Haar assumptions are possible; however, they do not seem to add much insight.

3. Uniqueness and de la Vallée Poussin Results

Best approximations from $\mathscr{R}_L$ need not be unique, as shown by the following example.

Example 4. Let $X=\{0,1\}$, $f(x)=x$, $\mathscr{P}=\Pi_0$, $\mathscr{Q}=\Pi_2$, $L(x)\equiv 0.1$, $R^*(x)=(1/11)/(1-0.9x)$. We have $X_0=\{0,1\}$, $Y_0=\{1\}$, $I_0=\{1\}$, $\sigma(0)=-1$, $\sigma(1)=1$. Thus

$$\sigma(x)\begin{bmatrix}\theta_1(x)\\(R^*\psi_1)(x)\\(R^*\psi_2)(x)\\(R^*\psi_3)(x)\end{bmatrix}=\sigma(x)\begin{bmatrix}1\\\frac{1}{11}\Big/(1-0.9x)\\\frac{x}{11}\Big/(1-0.9x)\\\frac{x^2}{11}\Big/(1-0.9x)\end{bmatrix}$$

so

$$S = \left\{ \begin{bmatrix} -1 \\ -\frac{1}{11} \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ \frac{10}{11} \\ \frac{10}{11} \\ \frac{10}{11} \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

and R^* is shown to be a best approximation by the equality

$$\frac{1}{3}\begin{bmatrix} -1 \\ \frac{1}{11} \\ 0 \\ 0 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 1 \\ \frac{10}{11} \\ \frac{10}{11} \\ \frac{10}{11} \end{bmatrix} + \frac{10}{33}\begin{bmatrix} 0 \\ -1 \\ -1 \\ -1 \end{bmatrix} + \frac{1}{33}\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

But $R(x) = (1/11)/(1 + q_2 x + q_3 x^2)$ is also a best approximation for any q_2, q_3 satisfying $|q_2| \leqslant 1$, $|q_3| \leqslant 1$, $q_2 + q_3 = -0.9$ since $\|f - R\| = 1/11 = \|f - R^*\|$.

It turns out uniqueness is assured if, unlike the situation in this example, $\mathbf{0}$ is in the interior of $\mathscr{H}(S)$. To prove this, we need the following lemma.

LEMMA 1. *Suppose $f \in C[X]$ and $R^* \in R_L$. Suppose $\tilde{X}_0$ is an arbitrary compact subset of X, and $\tilde{S}$ is the S of earlier theorems with X_0 replaced by $\tilde{X}_0$. If $\mathbf{0} \in \operatorname{int} \mathscr{H}(\tilde{S})$, then $\bar{P} \equiv 0$, $\bar{Q} \equiv 0$ is the only solution in $\mathscr{P}$, $\mathscr{Q}$ to the inequalities*

(a) $\sigma(x)(\bar{P} + R^*\bar{Q})(x) \geqslant 0$, $\forall x \in \tilde{X}_0$;

(b) $\bar{Q}(y) \leqslant 0$, $\forall y \in Y_0$;

(c) $\bar{q}_j \geqslant 0$ *if* $q_j^* = 1$;

(d) $\bar{q}_j \leqslant 0$ *if* $q_j^* = -1$.

Proof. Suppose $\bar{P} \in \mathscr{P}$, $\bar{Q} \in \mathscr{Q}$ satisfy (a)–(d). Letting $\mathbf{z} = [\bar{p}_1, \ldots, \bar{p}_m, \bar{q}_1, \ldots, \bar{q}_n]^T$, the system (a)–(d) may be rewritten as $\langle \mathbf{z}, \mathbf{s} \rangle \geqslant 0$ $\forall \mathbf{s} \in \tilde{S}$. Suppose that there is a $\mathbf{z} \neq \mathbf{0}$ satisfying these inequalities. Since $\mathbf{0} \in \operatorname{int} \mathscr{H}(\tilde{S})$, for

$\delta > 0$ sufficiently small we have $-\delta \mathbf{z} \in \mathscr{K}(\bar{S})$. By Caratheodory's theorem [2, p. 17] for some integer $k \leqslant m+n+1$, $\exists \mathbf{s}_1, \ldots, \mathbf{s}_k \in \bar{S}$, $\lambda_1, \ldots, \lambda_k \geqslant 0$ with $\sum_{i=1}^k \lambda_i = 1$ such that $-\delta \mathbf{z} = \sum_{i=1}^k \lambda_i \mathbf{s}_i$. Thus $-\delta\langle \mathbf{z}, \mathbf{z}\rangle = \langle \mathbf{z}, -\delta \mathbf{z}\rangle = \langle \mathbf{z}, \sum_{i=1}^k \lambda_i \mathbf{s}_i\rangle = \sum_{i=1}^k \lambda_i \langle \mathbf{z}, \mathbf{s}_i\rangle \geqslant 0$, which is contradiction. Thus $\mathbf{z} = \mathbf{0}$, so $\bar{P} \equiv 0$, $\bar{Q} \equiv 0$. Q.E.D.

Although we will not use it in this paper, the converse of this lemma can also be shown to be true.

We can now prove uniqueness of $\mathbf{0} \in \operatorname{int} \mathscr{K}(S)$.

THEOREM 4. *Suppose* $f \in C[X] - \mathscr{R}_L$ *and* $R^* = P^*/Q^* \in \mathscr{R}_L$. *If* $\mathbf{0} \in \operatorname{int} \mathscr{K}(S)$, *then* R^* *is the unique best approximation to* f *from* $\mathscr{R}_L$.

Proof. By Theorem 3, R^* is a best approximation. Suppose $R = P/Q \in \mathscr{R}_L$ were another best approximation. Let $\bar{P} = P - P^*$, $\bar{Q} = Q^* - Q$. For $x \in X_0$, we have

$$\begin{aligned}\sigma(x)(\bar{P} + R^*\bar{Q})(x) &= \sigma(x)(P - R^*Q)(x) \\ &= \sigma(x)\, Q(x)[f(x) - R^*(x) - (f(x) - R(x))] \geqslant 0.\end{aligned}$$

For $y \in Y_0$, we have

$$\bar{Q}(y) = Q^*(y) - Q(y) = L(y) - Q(y) \leqslant 0.$$

Finally, if $q_j^* = 1$, we have

$$\bar{q}_j = q_j^* - q_j = 1 - q_j \geqslant 0$$

and if $q_j^* = -1$, we have

$$\bar{q}_j = q_j^* - q_j = -1 - q_j \leqslant 0.$$

Thus by Lemma 1, $\bar{P} \equiv 0$ and $\bar{Q} \equiv 0$. Thus $P \equiv P^*$ and $Q \equiv Q^*$, so R^* is unique. Q.E.D.

The next example shows that the converse of this theorem if false.

EXAMPLE 5. Let $X = [0, 1]$, $f(x) = 2 - 2x$, $\mathscr{P} = \mathscr{Q} = \Pi_0$, $L(x) \equiv 0.1$, $R^*(x) \equiv 1/1$. We have $X_0 = \{0, 1\}$, $Y_0 = \varnothing$, $I_0 = (1)$, $S = \{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\}$. Here R^* is the unique best approximation, but $\mathbf{0} \notin \operatorname{int} \mathscr{K}(S)$.

In general, we always have $\mathbf{0} \notin \operatorname{int} \mathscr{K}(S)$ if R^* is a best approximation with $Y_0 = \varnothing$; the reason is that $\mathbf{0} \in \operatorname{int} \mathscr{K}(S)$ implies that the *coefficients* of the best approximation are unique, but $Y_0 = \varnothing$ implies that for some α with $0 < \alpha < 1$, $(\alpha P^*)/(\alpha Q^*)$ is another best approximation in $\mathscr{R}_L$ with different coefficients.

The hypotheses of Theorem 4 are actually sufficient to prove strong uniqueness.

THEOREM 5. *Suppose* R^* *is a best approximation to* $f \in C[X]$ *with* $0 \in \operatorname{int} \mathscr{H}(S)$. *Then there is a constant* $\gamma > 0$ *such that for any* $R \in \mathscr{R}_L$, $\|f - R\| \geqslant \|f - R^*\| + \gamma \|R - R^*\|$.

Proof (sketch). The proof follows the same general lines as the proof of strong uniqueness in the setting where the denominator is merely required to be positive [2, p. 165], with $P - P^*$ playing the role of P in the standard proof and $Q^* - Q$ playing the role of Q, and Lemma 1 of this paper used in place of Lemmas 1 and 2 in [2]. The other major change is that the definition of c on p. 166 of [2] is replaced by

$$c = \inf\{\max[-\sigma(x)(\bar{P} + R^*\bar{Q})(x)]: \bar{p} \in \mathscr{P}, \bar{Q} \in \mathscr{Q}, \bar{Q}(y) \leqslant 0 \ \forall y \in Y_0$$
$$\bar{q}_j \geqslant 0 \text{ if } q_j^* = 1, \bar{q}_j \leqslant 0 \text{ if } q_j^* = -1,$$
$$\max\{\|\bar{P} + R^*\bar{Q}\|, \|\bar{P}\|, \|\bar{Q}\|\} = 1\};$$

the extra complication in the last equality of this definition is needed since otherwise we could have $\|\bar{P} + R^*\bar{Q}\| = 1$ with $\|\bar{Q}\|$ and $\|\bar{P}\|$ arbitrarily large. Q.E.D.

We finish this section with two de la Vallée Poussin estimates, which give lower bounds on error norms.

THEOREM 6. *Suppose* $f \in C[X] - \mathscr{R}_L$, $R^* \in \mathscr{R}_L$ *(not necessarily a best approximation), and* $\tilde{X}_0 \neq \emptyset$ *is some compact subset of* X. *Suppose* $\mathbf{0} \in \mathscr{H}(\tilde{S})$, *where* $\tilde{S}$ *is the* S *of earlier theorems with* X_0 *replaced by* $\tilde{X}_0$. *Then* $\inf\{\|f - R\|: R \in \mathscr{R}_L\} \geqslant \min\{|f(x) - R^*(x)|: x \in \tilde{X}_0\}$.

Proof. Without loss of generality, we may assume $f(x) \neq R^*(x)$ $\forall x \in \tilde{X}_0$. Suppose the conclusion of the theorem is false. Then $\exists R = P/Q \in \mathscr{R}_L$ with $\|f - R\| < \min\{|f(x) - R^*(x)|: x \in \tilde{X}_0\}$. Let $\bar{P} \equiv P - P^*$, $\bar{Q} \equiv Q^* - Q$. Proceeding as in the proof of Theorem 4, we get

(a′) $\sigma(x)(\bar{P} + R^*\bar{Q})(x) > 0 \ \forall x \in \tilde{X}_0$;

(b) $\bar{Q}(y) \leqslant 0 \ \forall y \in Y_0$;

(c) $\bar{q}_j \geqslant 0$ if $q_j^* = 1$;

(d) $\bar{q}_j \leqslant 0$ if $q_j^* = -1$.

Since $\mathbf{0} \in \mathscr{H}(\tilde{S})$, Caratheodory's theorem [2, p. 17] implies that for some integer $k \leqslant m + n + 1$, $\exists \mathbf{s}_1, \ldots, \mathbf{s}_k \in \tilde{S}$, $\lambda_1, \ldots, \lambda_k \geqslant 0$ with $\sum_{l=1}^k \lambda_l = 1$ such that $\sum_{l=1}^k \lambda_l \mathbf{s}_l = \mathbf{0}$. Thus there are nonnegative induces u, v, w with $u + v + w = k$ such that $\sum_{l=1}^u \lambda_l \sigma(x_l)\, \hat{\mathbf{x}}_l + \sum_{l=1}^v \lambda_{u+l} \boldsymbol{\psi}(y_l) + \sum_{l=1}^w \lambda_{u+v+l} q_{i_l}^* \mathbf{e}_{m+i_l} = \mathbf{0}$,

where $x_1,\ldots, x_u \in \bar{X}_0$; $y_1,\ldots, y_v \in Y_0$; $i_1,\ldots, i_w \in I_0$ and $\hat{\mathbf{x}}_l = (\theta_1(x_l),\ldots, \theta_m(x_l), R^*(x_l)\,\psi_1(x_l),\ldots, R^*(x_l)\,\psi_n(x_l))^T$. We will next show that

$$\sum_{l=1}^{u} \lambda_l \sigma(x_l)(\bar{P} + R^*\bar{Q})(x_l) \leqslant 0.$$

We have

$$\sum_{l=1}^{u} \lambda_l \sigma(x_l)\,\theta_i(x_l) = 0 \qquad \text{for } i = 1,\ldots, m,$$

$$\sum_{l=1}^{u} \lambda_l \sigma(x_l)\,R^*(x_l)\,\psi_j(x_l) = \sum_{l=1}^{v} \lambda_{u+l}\psi_j(y_l) - \hat{\lambda}_j \qquad \text{for } j = 1,\ldots, n,$$

where

$$\hat{\lambda}_j = \lambda_{u+v+l} q_j^* \qquad \text{if } \; j = i_l \in I_0$$
$$= 0 \qquad \text{otherwise.}$$

So we get

$$\begin{aligned}
&\sum_{l=1}^{u} \lambda_l \sigma(x_l)(\bar{P} + R^*\bar{Q})(x_l) \\
&\quad = \sum_{l=1}^{u} \lambda_l \sigma(x_l) \left[\sum_{i=1}^{m} \bar{p}_i \theta_i(x_l) + R^*(x_l) \sum_{j=1}^{n} \bar{q}_j \psi_j(x_l) \right] \\
&\quad = \sum_{i=1}^{m} \bar{p}_i \left[\sum_{l=1}^{u} \lambda_l \sigma(x_l)\,\theta_i(x_l) \right] + \sum_{j=1}^{n} \bar{q}_j \left[\sum_{l=1}^{u} \lambda_l \sigma(x_l)\,R^*(x_l)\,\psi_j(x_l) \right] \\
&\quad = 0 + \sum_{j=1}^{n} \bar{q}_j \left[\sum_{l=1}^{v} \lambda_{u+l}\psi_j(y_l) - \hat{\lambda}_j \right] \\
&\quad = \sum_{l=1}^{v} \lambda_{u+l} \left[\sum_{j=1}^{n} \bar{q}_j \psi_j(y_l) \right] - \sum_{j=1}^{n} \hat{\lambda}_j \bar{q}_j \\
&\quad = \sum_{l=1}^{v} \lambda_{u+l} \bar{Q}(y_l) - \sum_{l=1}^{w} \lambda_{u+v+l} q_{i_l}^* \bar{q}_{i_l} \leqslant 0
\end{aligned}$$

by properties (b)–(d), as claimed. But by properties (a′), this implies $\lambda_1 = \cdots = \lambda_u = 0$. But this in turn implies that $\mathbf{0} \in \mathscr{H}(S')$, where $S' = \{\boldsymbol{\psi}(y): y \in Y_0\} \cup \{q_j^* \mathbf{e}_{m+j}: j \in I_0\}$. Thus, if $X_0 = \{x \in X: |f(x) - R^*(x)| = \|f - R^*\|\}$ as before, we have $\mathbf{0} \in \mathscr{H}(\{\sigma(x)\,\hat{\mathbf{x}}: x \in X_0\} \cup S')$. Thus, by Theorem 3, R^* is a best approximation to f. Thus $\inf\{\|f - R\|: R \in \mathscr{R}_L\} = \|f - R^*\| = \max_{x \in X} |f(x) - R^*(x)| \geqslant \min\{|f(x) - R^*(x)|: x \in \bar{X}_0\}$, contradicting the assumption that the conclusion of the theorem is false. Q.E.D.

If we assume that $\mathbf{0} \in \operatorname{int} \mathscr{H}(\tilde{S})$, we can prove the following stronger result.

THEOREM 7. *Suppose* $f \in C[X]$, $R^* = P^*/Q^* \in \mathscr{R}_L$ *(not necessarily a best approximation)*, and $\tilde{X}_0 \neq \varnothing$ *is some compact subset of* X. *Suppose* $\mathbf{0} \in \operatorname{int} \mathscr{H}(\tilde{S})$, *where* $\tilde{S}$ *is the* S *of earlier theorems with* X_0 *replaced by* $\tilde{X}_0$. *Then for every* $R = P/Q \in \mathscr{R}_L$ *with* $R \not\equiv R^*$ *we have* $\max\{|f(x) - R(x)|: x \in \tilde{X}_0\} > \min\{|f(x) - R^*(x)|: x \in \tilde{X}_0\}$.

Proof. Suppose the conclusion is false. Let $\bar{P} \equiv P - P^*$, $\bar{Q} \equiv Q^* - Q$. Then $\exists R \in \mathscr{R}_L$ with $R \not\equiv R^*$ and $\max\{|f(x) - R(x)|: x \in \tilde{X}_0\} \leqslant \min\{|f(x) - R^*(x)|: x \in \tilde{X}_0\}$. Proceeding as in the proof of Theorem 4, we get

(a″) $\sigma(x)(\bar{P} + R^*\bar{Q})(x) \geqslant 0$, $\forall x \in \tilde{X}_0$;

(b) $\bar{Q}(y) \leqslant 0$, $\forall y \in Y_0$;

(c) $\bar{q}_j \geqslant 0$ if $q_j^* = 1$;

(d) $\bar{q}_j \leqslant 0$ if $q_j^* = -1$.

Thus Lemma 1 implies $\bar{P} \equiv 0$, $\bar{Q} \equiv 0$. Thus $R \equiv R^*$, contrary to assumption. Q.E.D.

We observe that Example 4 with X replaced by $\{0, 0.1, 1\}$ and $X_0 = \{0, 1\}$ shows that the conclusion of the theorem may fail if $\mathbf{0} \notin \operatorname{int} \mathscr{H}(\tilde{S})$.

4. Computation of Best Approximations

The differential correction algorithm introduced by Cheney and Loeb [3] and discussed further by Barrodale *et al.* [1] can be modified to compute approximations from $\mathscr{R}_L$ by inserting extra constraints to force $Q(x) \geqslant L(x)$. We have

ALGORITHM (Restricted-denominator differential correction–RDDC).

(i) Choose $P_0/Q_0 \in \mathscr{R}_L$;

(ii) Having found $P_k/Q_k \in \mathscr{R}_L$ with $\|f - R_k\| = \Delta_k$, choose P_{k+1}, Q_{k+1} as a solution to the problem

$$\text{minimize:} \quad \max_{x \in X} \frac{|f(x)\,Q(x) - P(x)| - \Delta_k Q(x)}{Q_k(x)}$$

$$\text{subjec to:} \quad |q_j| \leqslant 1, \quad j = 1,\ldots, n, \quad \text{and} \quad Q(x) \geqslant L(x), \quad \forall x \in X;$$

(iii) continue until some stopping criterion is met.

One common stopping criterion is to stop when $(\Delta_k - \Delta_{k+1})/\Delta_k < \varepsilon$ for some prescribed $\varepsilon > 0$, selecting R_{k+1} as the approximation returned by the algorithm if $\Delta_{k+1} < \Delta_k$, and selecting R_k otherwise. A convenient way of choosing P_0/Q_0, which often is considerably more efficient than making some arbitrary choice such as $P_0/Q_0 \equiv 1/1$ (see Lee and Roberts [5] for numerical evidence in the unrestricted-denominator case), is to minimize $\max_{x\in X} |f(x)\,Q(x) - P(x)|$ subject to $Q(x) \geqslant L(x)$, $\forall x \in X$ and $|q_j| \leqslant 1$ for $j = 1,\ldots, n$.

Using the techniques of Barrodale *et al.* [1], we prove

THEOREM 8. *The RDDC algorithm converges monotonically and at least linearly.*

Proof. Let $M = \max_{x\in X} \sum_{j=1}^{n} |\psi_j(x)|$ (thus $\|Q\| \leqslant M$ for all $Q \in \mathscr{Q}$ with $|q_j| \leqslant 1$). Suppose R_k is not a best approximation. Let R^* be a best approximation with $\Delta^* = \|f - R^*\|$. Let $\delta = \min_{x\in X} L(x) > 0$. We have

$$\max_{x\in X} \frac{|f(x)\,Q_{k+1}(x) - P_{k+1}(x)| - \Delta_k Q_{k+1}(x)}{Q_k(x)}$$

$$\leqslant \max_{x\in X} \frac{|f(x)\,Q^*(x) - P^*(x)| - \Delta_k Q^*(x)}{Q_k(x)}$$

$$= \max_{x\in X} \left\{ [|f(x) - R^*(x)| - \Delta_k] \cdot \frac{Q^*(x)}{Q_k(x)} \right\} \leqslant [\Delta^* - \Delta_k] \cdot \frac{\delta}{M} < 0,$$

$\therefore \forall x \in X,$

$$\left[\left| f(x) - \frac{P_{k+1}(x)}{Q_{k+1}(x)} \right| - \Delta_k \right] \cdot \frac{Q_{k+1}(x)}{Q_k(x)} \leqslant [\Delta^* - \Delta_k] \cdot \frac{\delta}{M} < 0,$$

$\therefore \forall x \in X,$ $$\left| f(x) - \frac{P_{k+1}(x)}{Q_{k+1}(x)} \right| - \Delta_k < 0,$$

$\therefore \Delta_{k+1} < \Delta_k,$ so the convergence is monotonic.

$\forall x \in X$, we have

$$\left| f(x) - \frac{P_{k+1}(x)}{Q_{k+1}(x)} \right| - \Delta_k \leqslant [\Delta^* - \Delta_k] \cdot \frac{\delta}{M} \cdot \frac{Q_k(x)}{Q_{k+1}(x)}$$

$$\leqslant [\Delta^* - \Delta_k] \cdot \left(\frac{\delta}{M} \cdot \frac{\delta}{M} \right),$$

$$\therefore \Delta_{k+1} - \Delta_k \leqslant [\Delta^* - \Delta_k] \cdot \frac{\delta^2}{M^2},$$

$$\therefore \Delta_{k+1} - \Delta^* - (\Delta_k - \Delta^*) \leqslant -\frac{\delta^2}{M^2}(\Delta_k - \Delta^*),$$

$$\therefore \Delta_{k+1} - \Delta^* \leqslant \left(1 - \frac{\delta^2}{M^2}\right)(\Delta_k - \Delta^*),$$

$\therefore \Delta_k$ converges at least linearly to Δ^*. Q.E.D.

We observe that this theorem is stronger than the corresponding theorem in the unrestricted-denominator case [1, Theorems 1, 2] in that it gives information on the rate of convergence, and finiteness of X is not required to prove convergence. Finiteness of X is required, however, in order to run the algorithm in the usual way.

The following lemma was proved and used by Barrodale *et al.* [1] for the case $\mathscr{P} = \Pi_{m-1}$ and $\mathscr{Q} = \Pi_{n-1}$. The Haar subspace assumption of our lemma is equivalent to their assumption that $\min(m - 1 - \text{degree of } P^*, n - 1 - \text{degree of } Q^*) = 0$, with P^* and Q^* having no common nonconstant factors in their setting.

LEMMA 2. *Suppose X contains at least $m + n + 1$ distinct points and $R^* = P^*/Q^* \in \mathscr{R}_0 \equiv \{P/Q: P = p_1\theta_1 + \cdots + p_m\theta_m \in \mathscr{P}, \quad Q = q_1\psi_1 + \cdots + q_n\psi_n \in \mathscr{Q}, Q > 0$ on $X, \max_{1 \leqslant j \leqslant n} |q_j| = 1\}$. Suppose that the space spanned by $\{\theta_1, \ldots, \theta_m, R^*\psi_1, \ldots, R^*\psi_n\}$ is a Haar subspace of dimension $m + n - 1$; that is, the space has dimension $m + n - 1$, and no nontrivial element of it has more than $m + n - 2$ distinct zeros in X. Then $\exists \theta > 0$ such that for all $R = P/Q \in \mathscr{R}_0$ we have $\|Q - Q^*\| \leqslant \theta \|R - R^*\|$.*

Proof. Dua and Loeb [4] prove this lemma in the case where $X = [0, 1]$, $\mathscr{P} = \Pi_{m-1}$, and $\mathscr{Q} = \Pi_{n-1}$, but their proof requires these extra conditions only in proving that if $Q \geqslant 0$ on X and $P \equiv R^*Q$ on X, then $P \equiv P^*$ on X and $Q \equiv Q^*$ on X. This fact, however, follows from an argument of the type given by Cheney [2, p. 165]. Q.E.D.

We can now prove quadratic convergence of the RDDC in some circumstances.

THEOREM 9. *Suppose X contains at least $m + n + 1$ distinct points. $R^* = P^*/Q^* \in \mathscr{R}_L$ is a best approximation to $f \in C[X] - \mathscr{R}_L$, and the space spanned by $\{\theta_1, \ldots, \theta_m, R^*\psi_1, \ldots, R^*\psi_n\}$ is a Haar subspace of dimension $m + n - 1$. Suppose that either of the following two conditions holds:*

(A) $Y_0 = \varnothing$ *or*

(B) $\mathbf{0} \in \operatorname{int} \mathscr{H}(S)$.

Then the rate of convergence of the restricted-denominator differential correction algorithm is at least quadratic.

Proof. If some R_k is a best approximation the conclusion of the theorem is true, so we assume this is not the case.

Then the approximations produced by the RDDC algorithm (except possibly P_0/Q_0) satisfy the normalization $\max_{1 \leqslant j \leqslant n} |q_j| = 1$, since otherwise the (negative) minimum computed in step (ii) of the algorithm could be decreased by renormalization. Since we may also assume R^* satisfies this normalization, the hypotheses of Lemma 2 hold for R^* and R_k, $k \geqslant 1$. We now claim that if condition (A) holds, then R^* is a best approximation to f from $\mathscr{R}_0$. To see this, we note that by Theorem 3 we have $\mathbf{0} \in \mathscr{H}(\{\sigma(x)\,\hat{\mathbf{x}}: x \in X_0\} \cup \{q_j^* \mathbf{e}_{m+j}: j \in I_0\})$ (see Introduction). If there were $R = P/Q \in \mathscr{R}_0$ satisfying $\|f - R\| < \|f - R^*\|$, letting $\tilde{L}(x) \equiv \frac{1}{2}\min(\min_{x \in X} L(x), \min_{x \in X} Q(x))$ we have $R^* \in \mathscr{R}_{\tilde{L}}$ and $R \in \mathscr{R}_{\tilde{L}}$, with $Q^* > \tilde{L}$ and $Q > \tilde{L}$. But the convex hull statement above and Theorem 3 now imply that R^* is a best approximation from $\mathscr{R}_{\tilde{L}}$, which is a contradiction. Thus, if condition (A) holds, the strong uniqueness of R^* holds by a theorem in Cheney [2, p. 165], while if condition (B) holds, strong uniqueness holds by Theorem 5. With the strong uniqueness of R^* and the conclusion of Lemma 2 available, the rest of the proof is as given by Barrodale *et al.* [1, Theorem 3]. Q.E.D.

As noted earlier, condition (A) and (B) of this theorem are mutually exclusive; Example 4 shows that they are not exhaustive. Under the hypotheses of Theorem 9 the absolute values of the differences of the coefficients of R_k and R^* can be shown to be bounded by sequences which converge quadratically to zero.

It is sometimes desirable to ignore the function f at some points of X, but still apply the denominator restrictions on all of X; the theory of this paper goes through unchanged if the subset of X on which f is to be approximated is compact. This situation is illustrated in the following example, where f is the desired magnitude squared response of a digital filter, and we do not wish to approximate f in the "transition band" (0.1, 0.11).

EXAMPLE 6. Let $X = \{0, 0.005, 0.01, 0.015, \ldots, 0.5\}$,

$$\begin{aligned} f(x) &= 1, && 1 \leqslant x \leqslant 0.1, \\ &= \text{undefined}, && 0.1 < x < 0.11, \\ &= 0.14, && 0.11 \leqslant x \leqslant 0.5, \end{aligned}$$

$$\mathscr{P} = \mathscr{Q} = \left\{ \sum_{i=1}^{6} a_i \cos(2\pi(i-1)\,x) \right\},$$

$L(x) \equiv 0.125$. Applying the RDDC algorithm on a CDC CYBER 172, which has roughly 14 digits of accuracy, we get (rounded to five places)

$$R^*(x) = (0.34408 + 0.35185 \cos 2\pi x + 0.08270 \cos 4\pi x + 0.15552 \cos 6\pi x + 0.31629 \cos 8\pi x + 0.23541 \cos 10\pi x)/(1.00000 + 0.27473 \cos 2\pi x - 0.56073 \cos 4\pi x - 0.04879 \cos 6\pi x + 0.68555 \cos 8\pi x + 0.37588 \cos 10\pi x),$$

with $\Delta^* = 0.13945$, $X_0 = \{0^+, 0.085^-, 0.1^+, 0.11^-, 0.125^+, 0.215^-, 0.34^+, 0.345^+, 0.405^-, 0.41^-, 0.5^+\}$ (where the sign indicates the sign of $f - R^*$), $Y_0 = \{0.105\}$, $I_0 = \{1\}$. Nine iterations were required, with the quantities $\Delta_k - \Delta^*$ (for $k = 0, 1, \ldots, 8$) being approximately 1×10^{-1}, 8×10^{-2}, 2×10^{-2}, 1×10^{-2}, 4×10^{4}, 1×10^{-6}, 8×10^{-12}, 2×10^{-14}, 8×10^{-15}; this sequence indicates the eventual quadratic nature of the convergence, up to machine accuracy.

We finally remark that the theory of this paper can also be extended to include restricted range conditions and a positive continuous multiplicative weight function w; that is, we may further require the functions R in $\mathscr{R}_L$ to satisfy $R \leqslant u$ on X_1 and $R \geqslant l$ on X_2, where u and l are given continuous functions defined on compact subsets X_1 and X_2 of X, respectively, and we wish to minimize $\|w \cdot (f - R)\|$ instead of $\|f - R\|$. The set S must be expanded to include the vectors of the form $\tilde{\sigma}(x)\,\hat{\mathbf{x}}$ $(x \in Z_0)$, where $Z_0 = \{x \in X: R^*(x) = u(x) \text{ or } R^*(x) = l(x)\}$ and

$$\begin{aligned} \tilde{\sigma}(x) &= -1 \qquad \text{if} \quad R^*(x) = u(x) \\ &= 1 \qquad \text{if} \quad R^*(x) = l(x). \end{aligned}$$

In this setting, the assumption $\max_{x \in X} L(x) < \min_{x \in X} \psi_1(x)$ must be replaced by the assumption that there exists a member of $\mathscr{R}_L$ satisfying all the restrictions (denominator and restricted range) strictly. The restrictions $P \leqslant Q \cdot u$ on X_1 and $P \geqslant Q \cdot l$ on X_1 must be added to the RDDC algorithm, and in the expression to be minimized in steps (i) and (ii) $P(x)$ must be replaced by $w(x)\,P(x)$. The results of this paper still hold essentially as stated (with a few minor changes in the proofs) except for Theorems 2 and 9 and Lemma 1; in Theorem 2 we must add the condition $\operatorname{sgn}(\bar{P} + R^*\bar{Q})(x) = \tilde{\sigma}(x)$ $\forall x \in Z_0$; in Lemma 1 we must add the condition $\tilde{\sigma}(x)(\bar{P} + R^*\bar{Q}) \geqslant 0$ $\forall x \in Z_0$; and in Theorem 9 we must add to condition (A) either $Z_0 = \varnothing$, or $f \leqslant u$ on X_1 and $f \geqslant l$ on X_2. For numerical examples in this extended setting, see McCallig [6].

REFERENCES

1. I. BARRODALE, M. J. D. POWELL, AND F. D. K. ROBERTS, The differential correction algorithm for rational l_∞ approximation, *SIAM J. Numer. Anal.* **9** (1972), 493–504.
2. E. W. CHENEY, "Introduction to Approximation Theory," McGraw–Hill, New York, 1966.

3. E. W. CHENEY AND H. L. LOEB, Two new algorithms for rational approximation, *Numer. Math.* **3** (1961), 72–75.
4. S. N. DUA AND H. L. LOEB, Further remarks on the differential correction algorithm, *SIAM J. Numer. Anal.* **10** (1973), 123–126.
5. C. M. LEE AND F. D. K. ROBERTS, A comparison of algorithms for rational l_∞ approximation, *Math. Comp.* **27** (1973), 111–121.
6. M. T. MCCALLIG, R. KURTH, AND B. STEEL, Recursive digital filters with low coefficient sensitivity, *in* "Proceedings, 1979 IEEE International Conference on Accoustics, Speech, and Signal Processing, Washington D.C., April 1979."