

A Combined Remes-Differential Correction Algorithm for Rational Approximation

By Edwin H. Kaufman, Jr., David J. Leeming* and G. D. Taylor**

Abstract. In this paper a hybrid Remes-differential correction algorithm for computing best uniform rational approximants on a compact subset of the real line is developed. This algorithm differs from the classical multiple exchange Remes algorithm in two crucial aspects. First of all, the solving of a nonlinear system to find a best approximation on a given reference set in each iteration of the Remes algorithm is replaced with the differential correction algorithm to compute the desired best approximation on the reference set. Secondly, the exchange procedure itself has been modified to eliminate the possibility of cycling that can occur in the usual exchange procedure. This second modification is necessary to guarantee the convergence of this algorithm on a finite set without the usual normal and sufficiently dense assumptions that exist in other studies.

1. Introduction. This paper is divided into two parts. In the first part we consider X a compact subset of the real line with $\text{card}(X) \geq n + 2$. Let $C(X)$ denote the class of all continuous real valued functions defined on X , normed with the uniform norm, i.e., $\|f\| = \max\{|f(x)|: x \in X\}$. Let n be a positive integer and set

$$R_n^0(X) = \{r = 1/p: p \in \Pi_n, p(x) > 0 \text{ for all } x \in X\},$$

where Π_n denotes the set of all algebraic polynomials of degree $\leq n$. Note that $R_n^0(X)$ consists of only the positive elements of the set usually denoted by $R_n^0(X)$. In this setting we will give an algorithm for computing the best approximation for positive $f \in C(X)$ from $R_n^0(X)$. We believe this algorithm is the correct analog, for this setting, of the standard multiple exchange Remes algorithm for polynomials. We observe here that if $Y \subset X$, Y is compact and $\text{card}(Y) \geq n + 2$, then existence of a best approximant to positive $f \in C(Y)$ from $R_n^0(Y)$ is guaranteed by [6].

This algorithm contains some unique features including the incorporation of the differential correction algorithm [1], [4] to obtain a best approximation at each stage. This ensures that the denominator of the best approximation, p_k , on the k th reference set, X_k , will be positive on X_k . If, however, $p_k(x) \leq 0$ for some $x \in X \sim X_k$, we indicate two exchange procedures for selecting the next reference set. Note that in most studies this possibility is ignored by assuming (1) f is normal on some interval $[a, b]$ containing X ; (2) X is sufficiently dense in $[a, b]$; and, (3) X_k is sufficiently close to an alternating set of the best approximation to f on X . We shall also show that using our exchange procedure, there exists a $k_0 > 0$ such that for $k \geq k_0$, p_k

Received November 22, 1976.

AMS (MOS) subject classifications (1970). Primary 41A20, 41A50, 65D15.

*Research supported in part by National Research Council of Canada Grant A8061.

**Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant AFOSR-76-2878.

Copyright © 1978, American Mathematical Society

must be positive on X . From this point on, our exchange procedure will coincide with the standard multiple exchange procedure; and we can, therefore, guarantee convergence without the above assumptions. Our procedure could also be used to overcome the difficulty which Dunham [3] has pointed out in Williams' paper on interpolating rationals [8].

It should be further emphasized that a modified exchange procedure is actually necessary to guarantee the convergence of this algorithm without the assumptions (1)–(3) of above. Indeed, if one attempts to use the standard exchange procedure without regard to the possibility that $p_k \leq 0$ on $X \sim X_k$ may occur (and hoping that $p_k(x) = 0$ for $x \in X \sim X_k$ does not occur to give a divide fault), the usual proof that the error of approximation on the successive reference sets is strictly increasing is false. In fact, examples exist for which the error does not increase strictly and for which the algorithm actually cycles (i.e., $p_k = p_{k+2} = p_{k+4} = \dots$; $p_{k+1} = p_{k+3} = p_{k+5} = \dots$; $X_k = X_{k+2} = X_{k+4} = \dots$; $X_{k+1} = X_{k+3} = X_{k+5} = \dots$; starting at some k). Using either of the exchange procedures that we give, we are able to prove that the error of approximation on successive reference sets is strictly increasing.

The second part of the paper is devoted to the description of the Remes-Difcor algorithm (the name of our algorithm) for obtaining the best approximation to $f \in C(X)$, X a finite subset of the real line of at least $n + m + 2$ points, by elements of $R_n^m(X)$, $m \geq 0$, $n \geq 0$, where

$$R_n^m(X) = \{r = p/q: p \in \Pi_m, q \in \Pi_n, q(x) \geq \epsilon \text{ for all } x \in X\},$$

and the ϵ is some small positive number. A proof of the convergence of this algorithm is given, along with a flow chart. Finally, a brief discussion of some numerical results will be given. A complete discussion of the numerical results and comparison with both the Remes algorithm and the differential correction algorithm is planned for in a separate paper.

2. Approximating with $R_n^0(X)$. Let $f \in C(X) \sim R_n^0(X)$, with $f > 0$ on X . We first consider the case where X is a finite subset of the real line, with $\text{card}(X) \geq n + 2$. For each k , $k = 1, 2, \dots$, $X_k \subset X$ shall denote a reference set of $n + 2$ or $n + 3$ points and $r_k = 1/p_k \in R_n^0(X_k)$ will denote the best approximation to f on X_k from $R_n^0(X_k)$. This best approximation, r_k , is obtained by using the differential correction algorithm applied to the point set X_k . There are three advantages to finding r_k via the differential correction algorithm rather than via solving a nonlinear system of equations: a solution is guaranteed, we are assured that $p_k > 0$ on X_k , and no extra complications will arise if X_k has $n + 3$ points. After computing r_k , if X_k has $n + 3$ points we delete one point of X_k to get a new set Y_k of $n + 2$ points, taking care that $f - r_k$ alternates on Y_k . If X_k consists of $n + 2$ points, then we set $Y_k = X_k$. Set $e_k = \max\{|f(x) - r_k(x)|: x \in Y_k\}$, $Z_k = \{x \in X: p_k(x) > 0\}$ and consider the following two exchange procedures for constructing the next reference set X_{k+1} :

Exchange I: (The positive exchange). If r_k is not the best approximation to f on Z_k from $R_n^0(Z_k)$, X_{k+1} is constructed from Y_k by doing an ordinary Remes multiple exchange on the points of Z_k . If r_k is the best approximation to f on Z_k from

$R_n^0(Z_k)$, then the algorithm terminates if $Z_k = X$. If $Z_k \neq X$, then $y \in X$ satisfying $p_k(y) = \min\{p_k(x) : x \in X\}$ is found and X_{k+1} is defined to be $X_{k+1} = \{y\} \cup \{Y_k\}$. Note that in this case we have that $p_k(y) \leq 0$ and X_{k+1} consists of $n + 3$ distinct points of X .

Exchange II: (The negative exchange). In this exchange procedure, the algorithm first does a standard Remes multiple exchange on the point set Z_k with respect to $f - r_k$ and Y_k getting $W_k \subset Z_k$, where W_k consists of $n + 2$ points on which $f - r_k$ alternates in sign, $|f(w) - r_k(w)| \geq e_k$ for all $w \in W_k$ and $\max\{|f(w) - r_k(w)| : w \in W_k\} = \max\{|f(x) - r_k(x)| : x \in Z_k\}$. If $W_k = Y_k$ and $Z_k = X$, then the algorithm terminates as r_k is the desired best approximation to f on X . If this does not happen, then X_{k+1} is defined to be $W_k \cup \{y\}$ if $Z_k \neq X$ where y satisfies $p_k(y) = \min\{p_k(x) : x \in X\} \leq 0$ and W_k if $Z_k = X$.

Note that this exchange procedure differs from the first one in that whenever $Z_k \neq X$ an additional point where p_k takes on its minimum is added to the reference set. In the first exchange procedure this additional point is added only when r_k is the best approximation to f on Z_k from $R_n^0(Z_k)$. Also, note that whenever $Z_k = X$ both of these procedures coincide with the standard Remes multiple exchange procedure. For both of these exchange procedures the following theorem holds. (The set $X_1 \subset X$ is chosen so that it has $n + 2$ points and $e_1 > 0$.)

THEOREM 1. *If X is finite and the algorithm described above using either of the two exchange procedures is applied, then $\{e_k\}$ is strictly increasing. Furthermore, the algorithm eventually terminates at a best approximation to f on X from $R_n^0(X)$.*

Proof. Assume the algorithm does not terminate at stage k . To show that $e_k < e_{k+1}$ one must consider two cases. The first is when X_{k+1} is constructed only from points of Z_k . In this case p_k and p_{k+1} are both positive on X_{k+1} and a standard de la Vallée-Poussin type of argument (zero counting) shows that $e_k < e_{k+1}$ since $p_k \neq p_{k+1}$. In the case that $X_{k+1} = W_k \cup \{y\}$ where $W_k = Y_k$ or W_k is the result of a standard Remes multiple exchange in the points Z_k with respect to $f - r_k$ and Y_k , and $y \in X$ satisfies $p_k(y) = \min\{p_k(x) : x \in X\} \leq 0$ we first note that both p_k and p_{k+1} are positive on W_k . Also, $|f - r_{k+1}| \leq e_{k+1}$ on W_k and $f - r_k$ alternates in sign on W_k with $|f - r_k| \geq e_k$ on W_k . Thus, by zero counting we must once again have that $\max\{|f(z) - r_{k+1}(z)| : z \in W_k\} > \min\{|f(z) - r_k(z)| : z \in W_k\}$ since $p_k \neq p_{k+1}$ implying that $e_k < e_{k+1}$. (For a more careful treatment of the de la Vallée-Poussin type of argument see the proof of Lemma 2 later in the paper.) The rest of the theorem now follows since X is finite, and no reference set can occur more than once.

Although in actual computation one only encounters finite sets, it is of interest to consider the behavior of this algorithm if X is only required to be compact. In the remainder of this section we shall only consider Exchange I (the positive exchange). It can be shown that similar results are true for Exchange II. We first note that in this case the set $Z_k = \{x \in X : p_k(x) > 0\}$ may fail to be compact. If this happens, then it may not be possible to carry out the Remes multiple exchange on Z_k with respect to $f - r_k$ and Y_k . Thus, the algorithm must be modified by choosing some $\epsilon > 0$ and

setting $Z_k = \{x \in X: p_k(x) \geq \epsilon\}$. The elements of the set $G_k = \{x \in X: p_k(x) < \epsilon\}$ will be called g -poles (generalized poles) of r_k . The number ϵ should be chosen so that p_k has no g -poles on X_k . Since

$$\|1/p_k\|_{X_k} \leq \|f - r_k\|_{X_k} + \|f\|_{X_k} \leq 2\|f\|_{X_k} \leq 2\|f\|,$$

it suffices to choose any ϵ with $0 < \epsilon \leq 1/(2\|f\|)$. For such a choice of ϵ , the algorithm is defined as above with either of the two exchanges. We now prove that this modified algorithm converges globally and at least linearly.

THEOREM 2. *For X a compact subset of $[a, b]$, and $0 < \epsilon \leq 1/(2\|f\|)$, and $f \in C(X) \sim R_n^0(X)$, the rational functions r_k generated by the modified algorithm described above have no g -poles on X for $k \geq k_0$ (say) and converge uniformly to the best approximation r^* to f on X according to an inequality of the form $\|r_k - r^*\|_X \leq A\theta^k$, $0 < \theta < 1$, for $k \geq k_0$.*

Proof. Since the conclusion follows trivially if the algorithm terminates, we assume that this is not the case. The method of proof is to show that $\{e_k\}_{k=1}^\infty$ is increasing and to actually estimate this rate of increase. To prove that $e_k < e_{k+1}$ holds for all k , one simply uses the arguments of Theorem 1. Also, note that $\{e_k\}_{k=1}^\infty$ is bounded (otherwise $r \equiv 1$ would be a better approximation than r_k on X_k for some k). Hence, there exists e^* such that $e_k \uparrow e^*$. The remainder of this proof is broken into seven lemmas; the first of these, which proves that the points in Y_k cannot cluster is proved by arguments similar to Wendroff [7, p. 65]. Complete proofs of all the lemmas are available in [5].

LEMMA 1. *There exists $\delta > 0$ such that for every k , if $Y_k = \{x_0^k, \dots, x_{n+1}^k\}$, then $x_i^k \leq x_{i+1}^k - \delta$ for $i = 0, 1, \dots, n$.*

LEMMA 2. *Let \bar{X} be a compact set of real numbers containing at least $m + n + 2$ points, and let $f \in C(\bar{X})$. Suppose $r^* = p^*/q^* \in R_n^m(\bar{X})$ has defect $d = \min(m - \partial p^*, n - \partial q^*)$ and let $N = m + n + 2 - d$. Suppose that $f - r^*$ alternates in sign on $\{x_i\}_{i=1}^N \subset \bar{X}$ where $x_1 < x_2 < \dots < x_N$, and that $f(x_i) - r^*(x_i) \neq 0$, for $i = 1, \dots, N$. Then if $r = p/q \in R_n^m(\bar{X})$, $r \neq r^*$ on \bar{X} we have*

$$\max_{1 \leq i \leq N} |f(x_i) - r(x_i)| > \min_{1 \leq i \leq N} |f(x_i) - r^*(x_i)|.$$

Proof. Suppose

$$\max_{1 \leq i \leq N} |f(x_i) - r(x_i)| \leq \min_{1 \leq i \leq N} |f(x_i) - r^*(x_i)|.$$

Let $\Delta(x) \equiv r(x) - r^*(x) = (f(x) - r^*(x)) - (f(x) - r(x))$, for all $x \in \bar{X}$. Assume (without loss of generality) that $f(x_1) - r^*(x_1) > 0$; then we have $(-1)^i \Delta(x_i) \leq 0$, $i = 1, \dots, N$. Now for all $x \in \bar{X}$,

$$\Delta(x) = \frac{p(x)}{q(x)} - \frac{p^*(x)}{q^*(x)} = \frac{p(x)q^*(x) - p^*(x)q(x)}{q(x)q^*(x)} \equiv \frac{S(x)}{q(x)q^*(x)}$$

so that $(-1)^i S(x_i) \leq 0$, $i = 1, \dots, N$. But $\partial S \leq m + n - d = N - 2$ so $S \equiv 0$.

Therefore, $r \equiv r^*$ on \bar{X} ; and this contradiction proves the lemma. \square

The next lemma is proved by showing that in the contrary case p_k has too many relative extrema for some k .

LEMMA 3. *There exists a constant c such that for every k , if $p_k(x) = p_0^k + p_1^k x + \dots + p_n^k x^n$, then $|p_i^k| \leq c$ for $i = 0, 1, \dots, n$.*

COROLLARY. *There exists a constant $c^* > 0$ such that $|p_k(x)| \leq c^*$ for $k = 1, 2, \dots$ and all $x \in X$.*

Before proceeding to Lemma 4, we introduce some new notation and make a few remarks. We shall call the exchange from Y_k to X_{k+1} an augmented exchange if $X_{k+1} = Y_k \cup \{y\}$ (recall that $Y_k \subset X_k$ is a set of $n + 2$ points on which $f - r_k$ alternates with error e_k). Also, in this case the point $y \in X$ satisfies $p_k(y) = \min\{p_k(x) : x \in X\} < \epsilon$. Writing $X_{k+1} = \{y_0^k, \dots, y_{n+2}^k\}$, we have that X_{k+1} contains exactly one g -pole of r_k . Call this point y_σ^k . As stated earlier, we let r_{k+1} denote the best approximation to f from $R_n^0(X_{k+1})$ on X_{k+1} (found via the differential correction algorithm) and we define Y_{k+1} to be a subset of X_{k+1} on which $f - r_{k+1}$ alternates in sign with modulus e_{k+1} . Note that $y_\sigma^k \in Y_{k+1}$. This follows from Lemma 2. For $e_k \leq \lambda \leq e_{k+1}$ construct $\bar{r}_\lambda = 1/\bar{p}_\lambda$ by requiring that

$$f(\tilde{y}_i^k) - \bar{r}_\lambda(\tilde{y}_i^k) = \eta_i \lambda, \quad i = 0, 1, \dots, n,$$

where $\tilde{X}_{k+1} = Y_{k+1} \sim \{y_\sigma^k\} = \{\tilde{y}_0^k, \dots, \tilde{y}_n^k\}$ and $\eta_i = \text{sgn}[f(\tilde{y}_i^k) - r_{k+1}(\tilde{y}_i^k)]$. Observe that for $\eta_i = -1$, $f(\tilde{y}_i^k) - \eta_i \lambda = f(\tilde{y}_i^k) + \lambda > 0$ and for $\eta_i = +1$,

$$f(\tilde{y}_i^k) - \eta_i \lambda = f(\tilde{y}_i^k) - \lambda \geq f(\tilde{y}_i^k) - e_{k+1} = r_{k+1}(\tilde{y}_i^k) \geq \frac{1}{c^*} > 0.$$

Thus, \bar{p}_λ is well defined by these equations and for all $x \in X$, $\bar{p}_\lambda(x)$ is a continuous function of λ . Finally, let $\Delta = \inf\{\|f - r\| : r \in R_n^0(X)\}$. Note that $\Delta > 0$ since $f \in C(X) \sim R_n^0(X)$. Then,

LEMMA 4. *If at the k th exchange, an augmented exchange occurs and*

$$\text{sgn}[f(\tilde{y}_i^k) - r_k(\tilde{y}_i^k)] = \text{sgn}[f(\hat{y}_i^k) - r_{k+1}(\hat{y}_i^k)], \quad i = 0, 1, \dots, n,$$

then $e_{k+1} - e_k \geq \Omega(\|f\| - \Delta)$, where Ω is a constant independent of k .

Now let us assume there exists a subsequence of positive integers $\{k_m\}_{m=1}^\infty$ satisfying the following:

1. An augmented exchange occurs between Y_{k_m} and X_{k_m+1} .
2. $e_{k_m+1} - e_{k_m} < \Omega(\|f\| - \Delta)$ (since $e_{k_m} \uparrow e^* \leq \Delta$, where Δ is the error of best approximation to f from $R_n^0(X)$).

By our assumption 2, we see that the sign condition of Lemma 4 cannot hold, hence we have for each k_m , the additional condition:

3. $\text{sgn}[f(\tilde{y}_i^{k_m}) - r_{k_m+1}(\tilde{y}_i^{k_m})] \neq \text{sgn}[f(\hat{y}_i^{k_m}) - r_{k_m}(\hat{y}_i^{k_m})]$ for some i , $0 \leq i \leq n$.

Recall that for an augmented exchange between Y_{k_m} and X_{k_m+1} , that Y_{k_m+1} denotes a subset of X_{k_m+1} consisting of $n + 2$ points on which $f - r_{k_m+1}$ alternates with error e_{k_m+1} . Define $y_\beta^{k_m}$ by $\{y_\beta^{k_m}\} = Y_{k_m} \sim Y_{k_m+1}$. That is, $y_\beta^{k_m}$ is the point of X_{k_m+1} which is deleted in forming Y_{k_m+1} . Since we are considering Exchange Procedure I, we have that W_{k_m} may be taken to be Y_{k_m} whenever an augmented exchange is performed. Under these assumptions we prove the following two lemmas.

LEMMA 5. *If $\bar{p}_{e_k}(y_\beta^k) < \epsilon/(1 + \epsilon)$, then $e_{k+1} - e_k \geq \Omega''$ where Ω'' is a constant independent of k .*

LEMMA 6. *An augmented exchange can occur only a finite number of times.*

We now turn our attention to the case that the exchange from Y_k to X_{k+1} is not an augmented exchange. In this case, r_k is not the best approximation to f on Z_k from $R_n^0(Z_k)$ and $X_{k+1} = Y_{k+1} = \{x_0^{k+1}, \dots, x_{n+1}^{k+1}\}$ with no g -pole of r_k in X_{k+1} . Setting

$$\gamma_{k+1} = \min_{0 \leq i \leq n+1} |f(x_i^{k+1}) - r_k(x_i^{k+1})|$$

and

$$\beta_{k+1} = \max_{0 \leq i \leq n+1} |f(x_i^{k+1}) - r_k(x_i^{k+1})|,$$

we observe that $\gamma_{k+1} \geq e_k$ and $\beta_{k+1} > e_{k+1}$.

LEMMA 7. *There exists a constant $\Omega > 0$ (independent of k) such that if X_{k+1} is not obtained by an augmented exchange, then $e_{k+1} - \gamma_{k+1} \geq \Omega(\beta_{k+1} - e_{k+1})$.*

3. **Approximation from $R_n^m(X)$.** We now turn to the second objective of this paper. Here our approximating family is taken to be

$$R_n^m(X) = \{r = p/q: p \in \Pi_m, q \in \Pi_n, q > 0 \text{ on } X\},$$

and we require $\text{card}(X) \geq m + n + 2$ ($m \geq 0, n \geq 0$). g -poles are defined as before, i.e., $x \in X$ is said to be a g -pole of $r = p/q$ if $q(x) < \epsilon$ where $\epsilon > 0$. This concept is useful even when X is finite, since it enables us to avoid division by very small positive numbers. We have used $\epsilon = 10^{-16}$ on a UNIVAC 1106, which has roughly 18-digit accuracy in double precision. Unfortunately, we can no longer be sure that r_k will be g -pole free on its reference set, although this condition can be enforced by inserting additional constraints into the linear programming part of the differential correction algorithm (we will return to this point later). The algorithm we used (with $\text{card}(X) = \text{NUMGR} < \infty$) is described by the Flowcharts 1 and 2.

4. **Convergence of the Remes-Difcor Algorithm.** In this section we prove that if the 20-step stopping criterion is deleted from the Remes-Difcor flow chart, then under certain existence assumptions the algorithm will terminate at a best approximation to f from $R_n^m(X)$.

THEOREM 3. *Let X be a finite set of real numbers containing at least $m + n + 2$ points, and let $f \in C(X)$. Suppose that for each subset $Y \subset X$ containing exactly $m + n + 2$ or $m + n + 3$ points, a best approximation $r = p/q \in R_n^m(Y)$ exists for f , for this approximation we have $q \geq \epsilon$ on Y and that the differential correction algorithm will produce r when applied to Y . Then the Remes-Difcor algorithm will terminate at a best approximation r^* to f on X .*

Proof. Let X_0 be the initial reference set and let X_k be the reference set at the k th stage. Let r_k be the best approximation to f on X_k with

$$e_k = \max\{|f(x) - r_k(x)|: x \in X_k\}.$$

If the algorithm terminates at stage k , then there are no g -poles and the maximum error occurs in X_k ; thus, $e_k = \|f - r_k\|$ and r_k is the best approximation to f on X from $R_n^m(X)$.

Now suppose the algorithm does not terminate at the k th stage ($k \geq 1$). If r_{k-1} has g -poles in X , then at least one of these is included in X_k by construction, so that $r_k \neq r_{k-1}$. Also, if r_{k-1} has no g -poles in X , then $r_k \neq r_{k-1}$; since otherwise the maximum error for r_k in X would occur at the same point as the maximum error for r_{k-1} , and thus would be included in X_k . This would contradict the fact that the algorithm does not terminate at the k th stage.

Now $f - r_{k-1}$ must alternate on some set $\{x_1, x_2, \dots, x_{m+n+2-d_{k-1}}\} \subset X_{k-1}$, where d_{k-1} is the defect of r_{k-1} , and so by construction $f - r_{k-1}$ must alternate in sign on some set $\{x'_1, x'_2, \dots, x'_{m+n+2-d_{k-1}}\} \subset X_k$ with $|f(x'_i) - r_{k-1}(x'_i)| \geq e_{k-1}, i = 1, 2, \dots, m + n + 2 - d_{k-1}$. So by Lemma 2 we have

$$e_k = \max_{x \in X_k} |f(x) - r_k(x)| \geq \max_i |f(x'_i) - r_k(x'_i)| > \min_i |f(x'_i) - r_{k-1}(x'_i)| \geq e_{k-1}.$$

Therefore, $\{e_k\}$ is strictly increasing, so since there are only a finite number of possible reference sets contained in X the algorithm must terminate. \square

5. Examples and Conclusions. In order to get a time comparison of the Remes-Difcor algorithm with the ordinary differential correction algorithm alone, we ran the following digital filter design problem:

Let

$$X = [0, 0.2\pi] \cup [0.4\pi, \pi],$$

$$f(x) = \begin{cases} 1, & 0 \leq x \leq 0.2\pi, \\ 0.0123, & 0.4\pi \leq x \leq \pi. \end{cases}$$

We approximate from

$$\tilde{R}_2^q(X) = \left\{ \frac{p(x)}{q(x)} = (a_0 + a_1 \cos x + \dots + a_9 \cos 9x) / (b_0 + b_1 \cos x + b_2 \cos 2x) : \right.$$

$$\left. q > 0 \text{ on } X \right\}.$$

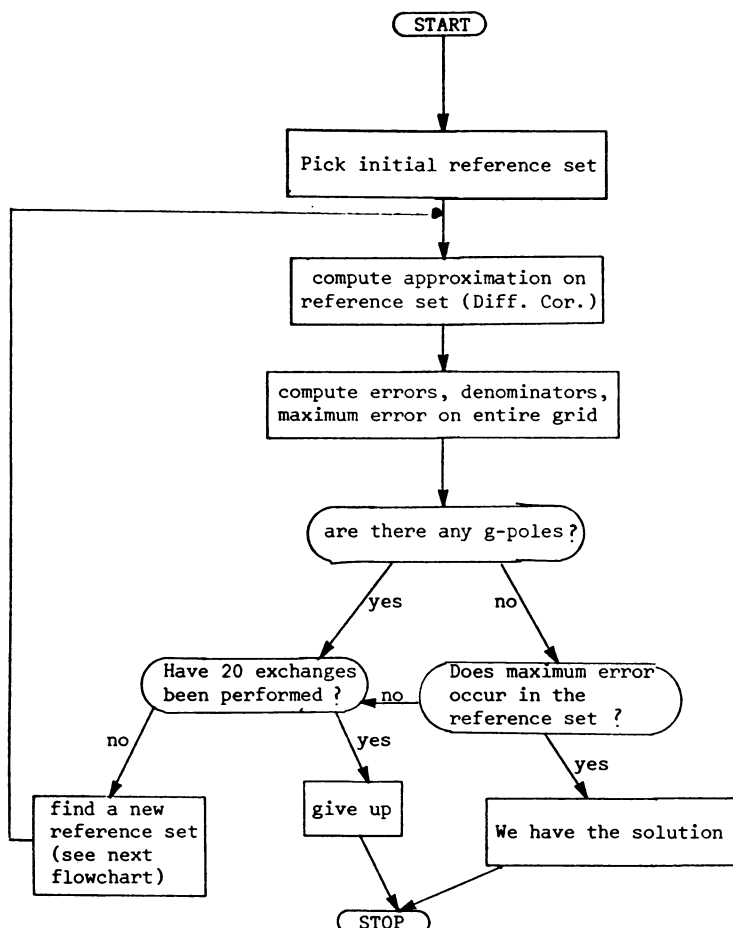
We also want $q > 0$ on $[0, \pi]$ and $p/q \geq 0$ on $[0, \pi]$, but in this example it is not necessary to do anything extra to force this. Although we are not using ordinary algebraic rational functions, we do have the alternating theory in this situation, and that is all that is required. To run this example we replaced X with an equally-spaced mesh (spacing $\pi/256$) containing 206 points. Using as our initial reference set five (roughly) equally spaced points in $[0, 0.2\pi]$ and eight (roughly) equally spaced points in $[0.4\pi, \pi]$, we obtained convergence after four exchanges and 60.0 seconds; $\|f - r^*\|$ was 1.83914×10^{-4} (where r^* is best). (Note. The final alternating set does have five points in $[0, 0.2\pi]$ and eight points in $[0.4\pi, \pi]$, but they are not equally spaced.) Starting with eight equally spaced points in $[0, 0.2\pi]$ and five in $[0.4\pi, \pi]$, eight exchanges and 1 minute 11.7 seconds were required; starting with all reference points pushed to the extreme right of $[0.4\pi, \pi]$ (which is one of the worst possible starting reference sets), fifteen exchanges and 2 minutes 0.5 seconds were required. On the other hand, running this problem with differential correction alone required 5 minutes

45.5 seconds. One would expect the time difference to increase if a finer mesh were used.

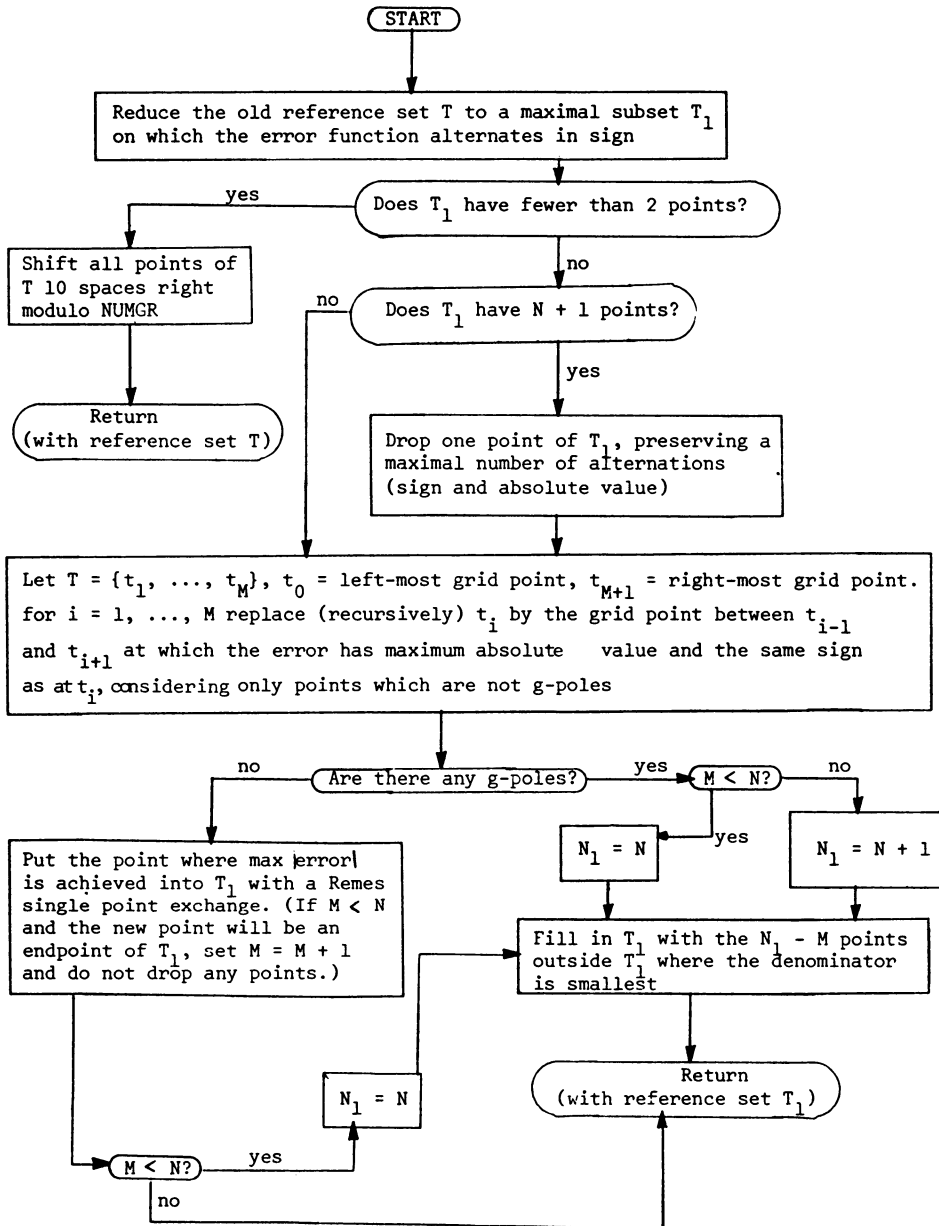
We also ran the Remes-Difcor program on an example for which best approximations did not exist on some reference sets, although a best approximation did exist on X . Here convergence depended on the choice of initial reference set, although we were able to obtain convergence even with a bad initial reference set if we "helped the program over the bad spots" by forcing $q \geq \epsilon$ on the reference set; this (as opposed to forcing $q \geq \epsilon$ on all of X) did not require much additional work.

In general, the relative merits of Remes, Remes-Difcor, and Difcor for finite X can be summarized as follows. When Remes works, so does Remes-Difcor, and with comparable speed. Remes-Difcor will usually still work when Remes fails due to problems in finding a new approximation on a reference set, and is much faster than Difcor if $\text{card}(X)$ is large. Difcor is theoretically more robust than Remes-Difcor since it does not require an alternating theory, and $\|f - r_k\|$ will converge to $\inf_r \|f - r\|$ even if there is no best approximation, but round-off and storage problems may be prohibitive if $\text{card}(X)$ is too large.

Remes-Difcor Flowchart #1 (excluding input-output), for a fixed number (NUMGR) of grid points



Flowchart #2. Finding a new reference set
 ($N = n + m + 2$, NUMGR = number of grid points)



Department of Mathematics
 Central Michigan University
 Mount Pleasant, Michigan 48859

Department of Mathematics
 University of Victoria
 Victoria, B. C., Canada V8W 2Y2

Department of Mathematics
Colorado State University
Fort Collins, Colorado 80523

1. I. BARRODALE, M. J. D. POWELL & F. D. K. ROBERTS, "The differential correction algorithm for rational l_∞ approximation," *SIAM J. Numer. Anal.*, v. 9, 1972, pp. 493–504.
2. E. W. CHENEY, *Introduction to Approximation Theory*, McGraw-Hill, New York, 1966.
3. C. B. DUNHAM, "A difficulty in Williams' algorithm for interpolating rationals," *Math. Comp.*, v. 29, 1975, pp. 552–553.
4. E. H. KAUFMAN, JR. & G. D. TAYLOR, "Uniform rational approximation by functions of several variables," *Internat. J. Numer. Meth. Engrg.*, v. 9, 1975, pp. 297–323.
5. E. H. KAUFMAN, JR., D. J. LEEMING & G. D. TAYLOR, *A Combined Remes-Differential Correction Algorithm for Rational Approximation*, Univ. of Victoria, Dept. of Math. Report, IM-106-IR.
6. D. J. LEEMING & G. D. TAYLOR, "Approximation with reciprocals of polynomials on compact sets," *J. Approximation Theory*, v. 21, 1977.
7. B. WENDROFF, *Theoretical Numerical Analysis*, Academic Press, New York, 1966.
8. J. WILLIAMS, "Numerical Chebyshev approximation by interpolating rationals," *Math. Comp.*, v. 26, 1972, pp. 199–206.