

Linear Chebyshev Approximation of Complex-Valued Functions

By I. Barrodale, L. M. Delves and J. C. Mason

Abstract. This paper is concerned with Chebyshev approximation by linear functions to complex-valued data. The problem is nonlinear, and we present a convergent algorithm for its solution. We also pose a related linear problem which is simple to solve, and which produces approximations which are near-best in the Chebyshev sense within a factor of $\sqrt{2}$. Some numerical examples are provided.

1. Introduction. The discrete best linear Chebyshev approximation problem for complex-valued functions can be stated as follows: Let $Z = \{z_1, z_2, \dots, z_m\}$ be a given discrete subset of the complex plane,* $f(z)$ and $\phi_1(z), \phi_2(z), \dots, \phi_n(z)$ be given complex-valued functions defined on Z , and for any set of n ($\leq m$) complex parameters $A = \{a_1, a_2, \dots, a_n\}$ let $L(A, z) = \sum_{j=1}^n a_j \phi_j(z)$. Then the problem is to determine a best set of parameters, A_∞ say, satisfying

$$(1) \quad \|r(A_\infty, z)\|_\infty \leq \|r(A, z)\|_\infty \quad \text{for all } A,$$

where $r(A, z) = f(z) - L(A, z)$, and, for any complex-valued function $g(z)$ defined on Z , the Chebyshev norm $\|g\|_\infty$ is given by

$$(2) \quad \|g\|_\infty = \max_{z \in Z} |g(z)|.$$

This problem has been considered by many authors (for example, see [2], [7], [8], and [11]): a best approximation $L(A_\infty, z)$ always exists, but it may not be unique. Most recently Ellacott and Williams [3] have applied Lawson's algorithm [6] to the problem; and although its ultimate convergence rate is very slow, they state that they are unaware of an algorithm which is faster. In this paper we present a globally convergent algorithm which (a) is easy to implement, and (b) often exhibits a rapid rate of convergence. Moreover, we direct attention to a simpler related problem which can be solved far more efficiently using existing numerical software, and which provides an approximation that is near-best in norm within a factor of $\sqrt{2}$.

We note that problem (1) can be restated as the mathematical programming problem

$$(3) \quad \min_{A, \tilde{w}} \{ \tilde{w} \mid \|r(A, z)\|_\infty \leq \tilde{w} \}.$$

Received August 29, 1977.

AMS (MOS) subject classifications (1970). Primary 65D15, 65E05, 65K05.

*More generally, Z is a discrete subset of any domain on which complex-valued functions can be defined.

Denoting the real and imaginary parts of $r(A, z_t)$ by $R(A, z_t)$ and $I(A, z_t)$, respectively, and squaring the norm in (3), it follows that an A_∞ can be determined by solving the (nonlinear) convex programming problem ([2], [8], [11])

$$(4) \quad \min_{A, w} \{w \mid [R(A, z_t)]^2 + [I(A, z_t)]^2 \leq w, \text{ for } t = 1, 2, \dots, m\}.$$

Letting \tilde{w}_∞ and w_∞ denote the optimal values of \tilde{w} and w in (3) and (4), respectively, it is clear that $\tilde{w}_\infty = \sqrt{w_\infty}$.

We describe in Section 3 an algorithm with guaranteed convergence which solves problem (4) iteratively via an obvious linearization. However, it is apparent that the nonlinearity in (4) is due to the choice of norm (2). In the next section we propose the use of a different norm which (a) leads to a single linear programming problem in place of (4), and (b) yields a near-solution to (4). In addition, this alternative norm is perhaps more natural than (2), since in complex approximation the Chebyshev norm is a combination of the real l_∞ and l_2 norms.

2. A Simpler Related Problem. Let v be any vector with complex components $v_t = x_t + iy_t$, for $t = 1, 2, \dots, m$; then $\|v\|_\infty = \max_t |v_t| = \max_t \sqrt{x_t^2 + y_t^2}$. If we now define a norm $\|\cdot\|_*$ as

$$(5) \quad \|v\|_* = \max_t \{\max(|x_t|, |y_t|)\},$$

then it is clear that

$$(6) \quad \|v\|_* \leq \|v\|_\infty \leq \sqrt{2} \|v\|_*.$$

Furthermore, if A_* is a best (in the sense of (5)) set of parameters satisfying

$$(7) \quad \|r(A_*, z)\|_* \leq \|r(A, z)\|_* \quad \text{for all } A,$$

we have the following result.

THEOREM 1. $\|r(A_\infty, z)\|_\infty \leq \|r(A_*, z)\|_\infty \leq \sqrt{2} \|r(A_\infty, z)\|_\infty$.

Proof.

$$\begin{aligned} \|r(A_\infty, z)\|_\infty &\leq \|r(A_*, z)\|_\infty && \text{from (1)} \\ &\leq \sqrt{2} \|r(A_*, z)\|_* && \text{from (6)} \\ &\leq \sqrt{2} \|r(A_\infty, z)\|_* && \text{from (7)} \\ &\leq \sqrt{2} \|r(A_\infty, z)\|_\infty && \text{from (6)}. \end{aligned}$$

An A_* satisfying (7) can be determined by solving the linear programming problem

$$(8) \quad \min_{A, \hat{w}} \{\hat{w} \mid |R(A, z_t)| \leq \hat{w}, |I(A, z_t)| \leq \hat{w}, \text{ for } t = 1, 2, \dots, m\}.$$

(The existence theorem of linear programming guarantees that an A_* exists, but it may not be unique.) Also, the optimal value of \hat{w} obtained by solving (8) is equal to $\|r(A_*, z)\|_*$, and this quantity could be used to bound $\|r(A_\infty, z)\|_\infty$ as follows.

THEOREM 2. $\|r(A_*, z)\|_* \leq \|r(A_\infty, z)\|_\infty \leq \sqrt{2} \|r(A_*, z)\|_*$.

Proof.

$$\begin{aligned} \|r(A_*, z)\|_* &\leq \|r(A_\infty, z)\|_* && \text{from (7)} \\ &\leq \|r(A_\infty, z)\|_\infty && \text{from (6)} \\ &\leq \|r(A_*, z)\|_\infty && \text{from (1)} \\ &\leq \sqrt{2} \|r(A_*, z)\|_* && \text{from (6)}. \end{aligned}$$

Problem (8) can be solved very efficiently using any algorithm which determines the (real) I_∞ solution to an overdetermined system of M linear equations in N unknowns: in this case $M = 2m$ and $N = 2n$ (unless the parameters a_j are required to be real, in which case $N = n$). In Section 5 we explain in more detail how to solve (8) in this manner, by means of a numerical example.

In view of Theorem 1, we feel that in most practical applications it is sufficient to solve problem (8) rather than the nonlinear problem (4). Indeed, it seems likely that in many applications the approximation problem will be as naturally, or more naturally, posed in the norm (5) as in the norm (2). However, for users who insist on solving problem (4) we present a suitable algorithm in the next section: it requires an initial estimate $A^{(0)}$ of A_∞ , and we normally set $A^{(0)}$ equal to the set A_* obtained by solving problem (8).

3. The Chebyshev Approximation Problem. One advantage of restating problem (1) as problem (4) is that there are several algorithms described in the literature for solving convex programming problems ([5], [9], [10], [11]). However, we describe in this section an alternative method which takes advantage of the special structure of problem (4). It is a simple algorithm with guaranteed convergence, and in practice it often converges rapidly.

Briefly, the method consists of replacing each quadratic constraint by its linear Taylor approximation, solving the resulting linear programming problem, and repeating this procedure until convergence occurs. The details are as follows.

Putting

$$(9) \quad a_j = b_j + ic_j,$$

$$(10) \quad f(z_t) = d(z_t) + ie(z_t) = d_t + ie_t,$$

and

$$(11) \quad \phi_j(z_t) = h_j(z_t) + ik_j(z_t) = h_{j,t} + ik_{j,t},$$

it follows that

$$(12) \quad R(A, z_t) = d_t - \sum_{j=1}^n (b_j h_{j,t} - c_j k_{j,t}),$$

and

$$(13) \quad I(A, z_t) = e_t - \sum_{j=1}^n (b_j k_{j,t} + c_j h_{j,t}).$$

Now suppose we have a current estimate $A^{(s)} = \{a_1^{(s)}, a_2^{(s)}, \dots, a_n^{(s)}\}$ of an A_∞ and we wish to change this to $A' = \{a_1^{(s)} + \alpha_1, a_2^{(s)} + \alpha_2, \dots, a_n^{(s)} + \alpha_n\}$, where

$$(14) \quad \alpha_j = \beta_j + i\gamma_j.$$

Then expanding $[R(A', z_t)]^2$ and $[I(A', z_t)]^2$ and ignoring second order terms in α_j , (of course, there are no higher powers of α_j), problem (4) yields the subproblem (15):

$$(15) \quad \left\{ \begin{array}{l} \text{minimize } w, \quad \text{where for each } t = 1, 2, \dots, m, \\ [R(A^{(s)}, z_t)]^2 - 2R(A^{(s)}, z_t) \sum_{j=1}^n (\beta_j h_{j,t} - \gamma_j k_{j,t}) \\ + [I(A^{(s)}, z_t)]^2 - 2I(A^{(s)}, z_t) \sum_{j=1}^n (\beta_j k_{j,t} + \gamma_j h_{j,t}) \leq w. \end{array} \right.$$

If $A^{(s)}$ is not a good estimate of an A_∞ , the constraints in (15) may be satisfied by some extremely large values of the variables β_j and γ_j . Hence, it is advisable to insert a lower bound of zero on the left side of each constraint, since this is obviously valid in problem (4), and to add the bounding constraints (for some large constant K) $|\beta_j| \leq K$ and $|\gamma_j| \leq K$, for $j = 1, 2, \dots, n$. With these insertions, and adopting the more concise notation

$$(16) \quad u_t^{(s)} = R(A^{(s)}, z_t),$$

$$(17) \quad v_t^{(s)} = I(A^{(s)}, z_t),$$

and

$$(18) \quad y_t^{(s)} = [u_t^{(s)}]^2 + [v_t^{(s)}]^2,$$

we arrive at the following LP problem. Choose w (nonnegative) and β_j, γ_j (unrestricted in sign, but bounded) to

$$(19) \quad \left\{ \begin{array}{l} \text{minimize } w \\ \text{subject to, for } t = 1, 2, \dots, m, \\ w + 2 \sum_{j=1}^n \beta_j (u_t^{(s)} h_{j,t} + v_t^{(s)} k_{j,t}) + 2 \sum_{j=1}^n \gamma_j (v_t^{(s)} h_{j,t} - u_t^{(s)} k_{j,t}) \geq y_t^{(s)}, \text{ and} \\ - 2 \sum_{j=1}^n \beta_j (u_t^{(s)} h_{j,t} + v_t^{(s)} k_{j,t}) - 2 \sum_{j=1}^n \gamma_j (v_t^{(s)} h_{j,t} - u_t^{(s)} k_{j,t}) \geq -y_t^{(s)}. \end{array} \right.$$

Letting $\beta_j^{(s)}$ and $\gamma_j^{(s)}$ denote optimal values of the variables in problem (19), $\alpha_j^{(s)}$ is formed from (14) and we set

$$(20) \quad A^{(s+1)} = \{a_1^{(s)} + \lambda^{(s)} \alpha_1^{(s)}, a_2^{(s)} + \lambda^{(s)} \alpha_2^{(s)}, \dots, a_n^{(s)} + \lambda^{(s)} \alpha_n^{(s)}\},$$

where $\lambda^{(s)}$ is a step length determined as follows. Setting

$$\delta^{(s)}(\lambda) = \max_{1 \leq t \leq m} \left| f(z_t) - \sum_{j=1}^n (a_j^{(s)} + \lambda \alpha_j^{(s)}) \phi_j(z_t) \right|^2$$

and denoting $\delta^{(s)}(0)$ by $\delta^{(s)}$, we distinguish two strategies:

(21) Choose $\lambda^{(s)}$ to minimize $\delta^{(s)}(\lambda)$, $0 \leq \lambda \leq \Lambda$, for some large constant Λ .

(22) Choose $\lambda^{(s)}$ to be the largest number λ_l of the form $\lambda_l = 1/2^{l-1}$ for $l = 1, 2, 3, \dots$, such that $\delta^{(s)}(\lambda_l) < \delta^{(s)}$ and $\delta^{(s)}(\lambda_l) < \delta^{(s)}(\lambda_{l+1})$.

By solving (19) for $s = 0, 1, 2, \dots$, either $\lambda^{(s)} = 0$ because $A^{(s)} = A_\infty$ or we obtain from (20) a sequence $\delta^{(0)}, \delta^{(1)}, \delta^{(2)}, \dots$. We show in the next section that with strategy (21) this sequence converges monotonically to the optimal value w_∞ of problem (4). In practice, we have instead used strategy (22) without (so far) encountering a single example of nonconvergence.

Finally, we note that problem (19) can be solved more efficiently in its dual linear programming form, which is exhibited in the Appendix.

4. Convergence Theory. It is convenient in this section to adopt vector notation. Let us therefore define

$$\mathbf{A} = [b_1, b_2, \dots, b_n, c_1, c_2, \dots, c_n]^T,$$

and

$$\boldsymbol{\alpha} = [\beta_1, \beta_2, \dots, \beta_n, \gamma_1, \gamma_2, \dots, \gamma_n]^T,$$

where the elements of these vectors are the real quantities defined in (9) and (14). In addition, for each $t = 1, 2, \dots, m$, we define

$$\begin{aligned} F_t(\mathbf{A}) &= \left| f(z_t) - \sum_{j=1}^n a_j \phi_j(z_t) \right|^2 \\ &= \left[d_t - \sum_{j=1}^n (b_j h_{j,t} - c_j k_{j,t}) \right]^2 + \left[e_t - \sum_{j=1}^n (b_j k_{j,t} + c_j h_{j,t}) \right]^2, \end{aligned}$$

in view of (10)–(13). Note that each $F_t(\mathbf{A})$ is a convex function bounded below by zero, and that all the partial derivatives $\partial F_t(\mathbf{A})/\partial b_j$ and $\partial F_t(\mathbf{A})/\partial c_j$ exist. Thus, putting

$$\nabla F_t(\mathbf{A}) = [\partial F_t(\mathbf{A})/\partial b_1, \dots, \partial F_t(\mathbf{A})/\partial b_n, \partial F_t(\mathbf{A})/\partial c_1, \dots, \partial F_t(\mathbf{A})/\partial c_n],$$

our algorithm can be stated in vector notation as follows.

Algorithm. Given $\mathbf{A}^{(s)}$ and $\delta^{(s)} = \max_{1 \leq t \leq m} F_t(\mathbf{A}^{(s)})$, solve the LP problem

$$(23) \min_{\boldsymbol{\alpha}, w} \{w \mid 0 \leq F_t(\mathbf{A}^{(s)}) + \nabla F_t(\mathbf{A}^{(s)}) \cdot \boldsymbol{\alpha} \leq w, 1 \leq t \leq m, \text{ and } \|\boldsymbol{\alpha}\|_\infty \leq K\}.$$

Then, letting $\boldsymbol{\alpha}^{(s)}$ and $w^{(s)}$ denote optimal values of the variables in (23), set

$$(24) \quad \mathbf{A}^{(s+1)} = \mathbf{A}^{(s)} + \lambda^{(s)} \boldsymbol{\alpha}^{(s)},$$

where $\lambda^{(s)}$ is chosen to

$$(25) \quad \text{minimize } \max_{1 \leq t \leq m} F_t(\mathbf{A}^{(s)} + \lambda \boldsymbol{\alpha}^{(s)}).$$

THEOREM 3. *Solving (23)–(25) for $s = 0, 1, 2, \dots$ produces a sequence $\delta^{(0)}, \delta^{(1)}, \delta^{(2)}, \dots$ which converges monotonically to the optimal value w_∞ of problem (4).*

Proof. Let $J = \{t \mid F_t(\mathbf{A}^{(s)}) = \delta^{(s)}\}$. Firstly, we note that for any given value of w , the constraints in (23) and (25) impose bounds on $\|\boldsymbol{\alpha}\|_\infty$ and on $|\lambda|$. Then, following Zangwill [10, p. 105], since the maps

$$D : \mathbf{A}^{(s)} \rightarrow (\mathbf{A}^{(s)}, \boldsymbol{\alpha}^{(s)}) \quad \text{and} \quad M : (\mathbf{A}^{(s)}, \boldsymbol{\alpha}^{(s)}) \rightarrow \lambda^{(s)}$$

involve minimizing a continuous function over a compact set, this implies that both M and D are closed. Also, since $\boldsymbol{\alpha} = \mathbf{0}$ and $w = \delta^{(s)}$ satisfy the constraints in (23), it follows that $w^{(s)} \leq \delta^{(s)}$.

Case (i). Assume $w^{(s)} = \delta^{(s)}$. Then (23) implies that no $\boldsymbol{\alpha}$ exists for which

$$\nabla F_j(\mathbf{A}^{(s)}) \cdot \boldsymbol{\alpha} < 0 \quad \text{for every } j \in J.$$

Hence, for any $\boldsymbol{\alpha}$ there exists $p \in J$ such that the directional derivative

$$(26) \quad \nabla F_p(\mathbf{A}^{(s)}) \cdot \boldsymbol{\alpha} \geq 0,$$

and from (26), and the definition of a directional derivative, it follows that

$$(27) \quad F_p(\mathbf{A}^{(s)} + \lambda \boldsymbol{\alpha}) \geq F_p(\mathbf{A}^{(s)}) = \delta^{(s)} \quad \text{for all } \lambda, 0 < \lambda \leq \lambda_p \text{ (say)}.$$

Finally, (27) implies that $\mathbf{A}^{(s)}$ is a local minimum of the convex function

$$(28) \quad F(\mathbf{A}) = \max_{1 \leq t \leq m} F_t(\mathbf{A}),$$

and so $\mathbf{A}^{(s)}$ is a global minimum of $F(\mathbf{A})$ and $\delta^{(s)} = w_\infty$.

Case (ii). Assume $w^{(s)} < \delta^{(s)}$. Then (23) implies that for every $j \in J$

$$(29) \quad \nabla F_j(\mathbf{A}^{(s)}) \cdot \boldsymbol{\alpha}^{(s)} < 0;$$

and so, for every $j \in J$

$$(30) \quad F_j(\mathbf{A}^{(s)} + \lambda \boldsymbol{\alpha}^{(s)}) < F_j(\mathbf{A}^{(s)}) = \delta^{(s)} \quad \text{for all } \lambda, 0 < \lambda \leq \lambda_j \text{ (say)}.$$

Now, if we define a positive number ϵ as

$$\epsilon = \min_{t \notin J} (\delta^{(s)} - F_t(\mathbf{A}^{(s)})),$$

since each $F_t(\mathbf{A})$ is continuous, it follows that for every $t \notin J$

$$(31) \quad F_t(\mathbf{A}^{(s)} + \lambda \boldsymbol{\alpha}^{(s)}) < F_t(\mathbf{A}^{(s)}) + \epsilon \leq \delta^{(s)} \quad \text{for all } \lambda, 0 < \lambda \leq \lambda_t \text{ (say)}.$$

From (30) and (31) we see that

$$(32) \quad \max_{1 \leq t \leq m} F_t(A^{(s)} + \lambda \alpha^{(s)}) < \delta^{(s)},$$

for all λ satisfying $0 < \lambda \leq \min(\min_{j \in J} \lambda_j, \min_{t \notin J} \lambda_t)$. In particular, we can select a $\lambda^{(s)}$ as prescribed in (25), and from (32) we see that $\delta^{(s+1)} < \delta^{(s)}$. Finally, since the convex function $F(A)$ defined in (28) is nonnegative and $\delta^{(s)} = F(A^{(s)})$, it follows that the sequence $\delta^{(0)}, \delta^{(1)}, \delta^{(2)}, \dots$ converges monotonically to δ (say). But since the map MD is closed, Zangwill's Theorem A [10, p. 91] tells us that $\delta = w_\infty = \min F(A)$.

5. Numerical Examples. Although (8) is a linear programming problem, a solution A_* can also be obtained from any algorithm which solves an overdetermined system of M (real) linear equations in N (real) unknowns, say

$$(33) \quad \mathbf{Bx} = \mathbf{g},$$

in the l_∞ sense. Thus, defining two $m \times n$ matrices

$$(34) \quad \mathbf{H} = [h_{j,t}]^T \quad \text{and} \quad \mathbf{K} = [k_{j,t}]^T,$$

where the elements in (34) are given by (11), and putting

$$(35) \quad \mathbf{d} = [d_1, d_2, \dots, d_m]^T \quad \text{and} \quad \mathbf{e} = [e_1, e_2, \dots, e_m]^T,$$

where the elements in (35) are given by (10), problem (8) can be solved in the form (33) by setting

$$(36) \quad \mathbf{B} = \left[\begin{array}{c|c} \mathbf{H} & -\mathbf{K} \\ \hline \mathbf{K} & \mathbf{H} \end{array} \right],$$

and

$$(37) \quad \mathbf{g} = \begin{bmatrix} \mathbf{d} \\ \mathbf{e} \end{bmatrix}.$$

Here, $M = 2m$ and $N = 2n$, and an l_∞ solution \mathbf{x}^* to (33) has the form

$$(38) \quad \mathbf{x}^* = [b_1^*, b_2^*, \dots, b_n^*, c_1^*, c_2^*, \dots, c_n^*]^T.$$

A solution A_* to (8) is formed from the elements in (38) by using (9). Finally, we note that if the parameters a_j are required to be real, then $N = n$ and \mathbf{B} is set equal to the left half of the partitioned matrix in (36).

In the numerical examples which follow we exhibit various polynomial approximations to

$$(39) \quad f(z) = 1/(z - \xi),$$

where ξ is a complex constant with $|\xi| > 1$. For this function it is known [7, p. 33] that on the disc $|z| \leq 1$, the polynomial of degree $n - 1$ of best approximation in the Chebyshev sense satisfies

$$(40) \quad \|r(A_\infty, z)\|_\infty = |\xi^{1-n} / (|\xi|^2 - 1)|.$$

The approximations below were all computed on 100 points evenly distributed around the boundary of the disc $|z| \leq 1$, and the approximating functions are polynomials of degree $n - 1$; thus

$$(41) \quad Z = \{z_t \mid z_t = e^{i\pi(t-1)/50}, \text{ for } t = 1, 2, \dots, 100\},$$

and

$$(42) \quad L(A, z) = \sum_{j=1}^n a_j z^{j-1}.$$

Problem (33) was solved using the Fortran subroutine in Barrodale and Phillips [1]; this algorithm does not require \mathbf{B} to satisfy any special hypotheses, such as the Haar condition. All the calculations were performed in double precision (16 decimal digits) on an IBM 370 computer.

Example 1. Solve problem (8) for $\xi = 2 + i$, $n = 3$, and complex-valued a_j 's. We obtained the following results:

$$a_1^* = (-.4000623603(+0), .1999973128(+0)),$$

$$a_2^* = (-.1200095730(+0), .1600037836(+0)),$$

$$a_3^* = (-.2001397696(-1), .1099618568(+0)),$$

$$\|r(A_*, z)\|_* = .4995538598(-1), \text{ and}$$

$$\|r(A_*, z)\|_\infty = .5009811947(-1).$$

In Examples 2 and 3 we illustrate the algorithm (23)–(25) for solving problem (4). In each case the initial estimate $A^{(0)}$ of A_∞ is in fact the set A_* which results from solving the corresponding problem (8), and then each iteration produces a new estimate $A^{(s)}$ via the dual of problem (23). See the Appendix for further details. In practice we have used strategy (22) to determine $\lambda^{(s)}$; and we have omitted the constraint $\|\alpha\|_\infty \leq K$ in (23), since for large K its purpose is subsumed by the zero lower bound on the constraints.

Example 2. Solve problem (4) for $\xi = 2 + i$, $n = 3, 5$, and 7 , and complex-valued a_j 's and terminate the iterations when

$$(43) \quad \max_j (|\beta_j|, |\gamma_j|) \leq 10^{-10}.$$

We obtained the following values for $\|r(A^{(s)}, z)\|_\infty$:

iteration # (=s)	$n = 3$	$n = 5$	$n = 7$
0	.5009811947(-1)	.1007252663(-1)	.2021313394(-2)
1	.5000006319(-1)	.1000040228(-1)	.2000092876(-2)
2	.5000000000(-1)	.1000000001(-1)	.2000000001(-2)
3	.5000000000(-1)	.1000000000(-1)	.2000000000(-2)

For each iteration $\lambda^{(s)} = 1$, and criterion (43) is satisfied when $s = 3$.

Example 3. Solve problem (4) for $\xi = 2$, $n = 3, 5$, and 7 , and real-valued a_j 's, and terminate the iteration when

$$(44) \quad \max_j |\beta_j| \leq 10^{-10}.$$

We obtained the following values for $\|r(A^{(s)}, z)\|_\infty$:

iteration # (=s)	n = 3	n = 5	n = 7
0	.8360106268(-1)	.2112277615(-1)	.5234157415(-2)
1	.8333343077(-1)	.2083614349(-1)	.5208367354(-2)
2	.8333333333(-1)	.2083333344(-1)	.5208333333(-2)
3	.8333333333(-1)	.2083333333(-1)	.5208333333(-2)

For each iteration $\lambda^{(s)} = 1$, and criterion (44) is satisfied when $s = 3$.

In Examples 2 and 3 the values of $\|r(A^{(3)}, z)\|_\infty$, which of course relate to approximations of the form (42) computed on the discrete set (41), agree with expression (40) to within the ten digits shown.

6. Discussion. In the above examples the values of $\|r(A_*, z)\|_\infty / \|r(A_\infty, z)\|_\infty$ are well within the bound dictated by Theorem 1, although we have produced other numerical examples where this ratio is equal to $\sqrt{2}$. However, in spite of the efficacy and simplicity of the algorithm of this paper for solving the Chebyshev problem (4), we are convinced that for all practical purposes there is little to be lost (and much to be gained in efficiency) by solving instead the linear problem (8). Interestingly, the resulting approximation $L(A_*, z)$ is not only near-best in the (complex) Chebyshev sense (2), but also the following extremal property is satisfied: it is analogous to an extremal property which pertains to best approximation in the (real) l_∞ sense.

THEOREM 4. *Let the rank of the matrix B defined in (36) be ρ (and note that, normally, $\rho = 2n$ for complex-valued a_j 's, and $\rho = n$ for real-valued a_j 's). Then for some approximation $L(A_*, z)$ obtained by solving problem (8), the $2m$ inequalities*

$$|R(A_*, z_t)| \leq \hat{w} \quad \text{and} \quad |I(A_*, z_t)| \leq \hat{w}$$

become strict equalities at least $\rho + 1$ times when \hat{w} assumes its minimum value $\|r(A_, z)\|_*$.*

(The complementary slackness theorem of linear programming can be used to prove Theorem 4, but we omit the details here.)

Finally, we note that Geddes and Mason [4] show how near-best approximations in the Chebyshev sense to within a factor of $O(\log n)$ can be easily obtained via projection techniques, for simple regions and polynomial approximating functions of degree n .

Acknowledgements. We wish to acknowledge the programming assistance of Mr. K. B. Wilson, and the financial support provided by NRC Grant No. A5251. The work of the second-named author was carried out while he was Visiting Professor at

the University of Victoria, and the financial support provided during this period is also gratefully acknowledged.

Appendix. Problem (19), omitting the upper bound K on the variables β_j and γ_j , can be solved more efficiently as the following dual linear programming problem in nonnegative variables p_t and q_t :

$$(45) \quad \left\{ \begin{array}{l} \text{maximize } \sum_{i=1}^m (p_t - q_t)y_t^{(s)} \text{ subject to, for } j = 1, 2, \dots, n, \\ 2 \sum_{t=1}^m (p_t - q_t)(u_t^{(s)}h_{j,t} + v_t^{(s)}k_{j,t}) = 0, \\ 2 \sum_{t=1}^m (p_t - q_t)(v_t^{(s)}h_{j,t} - u_t^{(s)}k_{j,t}) = 0, \\ \text{and } \sum_{t=1}^m p_t \leq 1. \end{array} \right.$$

The variables β_j , γ_j , and w of problem (19) correspond to the artificial variables and slack variable which must be inserted in problem (45) prior to its solution by the simplex method. The known quantities in (45) are, of course, all real, and they are defined by Eqs. (10)–(13) and (16)–(18).

Now the Fortran subroutine [1] calculates an l_∞ solution to an $m \times N$ overdetermined system of linear equations

$$(46) \quad \mathbf{Bx} = \mathbf{g},$$

by applying a modification of the simplex method to the corresponding dual linear programming problem. The latter problem differs from the dual linear programming problem (45) only in that the constraint $\sum_{t=1}^m p_t \leq 1$ in (45) is replaced by the constraint $\sum_{t=1}^m (p_t + q_t) \leq 1$. Thus, by modifying the Fortran subroutine [1] so as to allow for this constraint replacement, we can solve via (45) the primal problem (19).

In turn, this requires that we express (19) in the form (46), which can be done formally as follows.

Defining two $m \times n$ matrices

$$(47) \quad \mathbf{B}_1 = [2u_t^{(s)}h_{j,t} + 2v_t^{(s)}k_{j,t}] \quad \text{and} \quad \mathbf{B}_2 = [2v_t^{(s)}h_{j,t} - 2u_t^{(s)}k_{j,t}]$$

where the elements in (47) are as in (19), we set

$$(48) \quad \mathbf{B} = [\mathbf{B}_1 \mid \mathbf{B}_2],$$

and

$$(49) \quad \mathbf{g} = [y_1^{(s)}, y_2^{(s)}, \dots, y_m^{(s)}]^T,$$

where the elements in (49) are given by (18). Here $N = 2n$, and for (48) and (49)

a solution $\mathbf{x}^{(s)}$ produced by the modified Fortran subroutine has the form

$$\mathbf{x}^{(s)} = [\beta_1^{(s)}, \beta_2^{(s)}, \dots, \beta_n^{(s)}, \gamma_1^{(s)}, \gamma_2^{(s)}, \dots, \gamma_n^{(s)}]^T.$$

Also, we note that if the parameters a_j are required to be real, then $N = n$ and \mathbf{B} is set equal to the left half of the partitioned matrix in (48).

Finally, and perhaps somewhat surprisingly, the modified Fortran subroutine for solving problem (19) via its dual form (45) is obtained merely by changing the constant 2. to 1. in lines CHE 1720, 1820, 2020, and 2230 in [1, p. 269]. More generally, this modified subroutine can be used to minimize the l_∞ norm of any (real) overdetermined system of linear equations subject to the constraint that the residual vector be nonnegative.

Department of Mathematics
University of Victoria
Victoria, B. C., Canada

Department of Computational and Statistical Science
University of Liverpool
Liverpool, England

Mathematics Division
The Royal Military College of Science
Shrivenham, Wiltshire, England

1. I. BARRODALE & C. PHILLIPS, *Algorithm 495—Solution of an Overdetermined System of Linear Equations in the Chebyshev Norm*, ACM Trans. Math. Software, v. 1, 1975, pp. 264–270.
2. L. COLLATZ & W. WETTERLING, *Optimization Problems*, Springer-Verlag, Berlin and New York, 1975, pp. 278–280.
3. S. ELLACOTT & JACK WILLIAMS, "Linear Chebyshev approximation in the complex plane using Lawson's algorithm," *Math. Comp.*, v. 30, 1976, pp. 35–44.
4. K. O. GEDDES & J. C. MASON, "Polynomial approximation by projections on the unit circle," *SIAM J. Numer. Anal.*, v. 12, 1975, pp. 111–120.
5. P. E. GILL & W. MURRAY, *Numerical Methods for Constrained Optimization*, Academic Press, New York, 1974.
6. C. L. LAWSON, *Contributions to the Theory of Linear Least Maximum Approximations*, Thesis, Univ. of California, Los Angeles, 1961.
7. G. G. LORENTZ, *Approximation of Functions*, Holt, Rinehart and Winston, New York, 1966.
8. P. RABINOWITZ, "Mathematical programming and approximation," in *Approximation Theory*, A. Talbot (Editor), Academic Press, New York, 1970, pp. 217–231.
9. D. M. SIMMONS, *Nonlinear Programming for Operations Research*, Prentice-Hall, Englewood Cliffs, N. J., 1975, (Chapter 9).
10. W. I. ZANGWILL, *Nonlinear Programming*, Prentice-Hall, Englewood Cliffs, N. J., 1969, (Chapter 14).
11. S. I. ZUKHOVITSKIY & L. I. AVDEYEVA, *Linear and Convex Programming*, Saunders, Philadelphia, 1966, (Chapter 6).