

A Note on the Semi-Infinite Programming Approach to Complex Approximation

By Roy L. Streit and Albert H. Nuttall

Abstract. Several observations are made about a recently proposed semi-infinite programming (SIP) method for computation of linear Chebyshev approximations to complex-valued functions. A particular discretization of the SIP problem is shown to be equivalent to replacing the usual absolute value of a complex number with related estimates, resulting in a class of quasi-norms on the complex number field \mathbf{C} , and consequently a class of quasi-norms on the space $C(Q)$ consisting of all continuous functions defined on $Q \subset \mathbf{C}$, Q compact. These quasi-norms on $C(Q)$ are estimates of the L_∞ norm on $C(Q)$ and are useful because the best approximation problem in each quasi-norm can be solved by solving (i) an ordinary linear program if Q is finite or (ii) a simplified SIP if Q is not finite.

Glashoff and Roleff [1] solve a semi-infinite program (SIP) which is shown to be equivalent to the linear approximation problem for functions in $C(Q)$, where $C(Q)$ is the space of complex-valued continuous functions on a compact (and not necessarily finite) subset Q of the complex plane \mathbf{C} and is equipped with the uniform (L_∞) norm

$$(1) \quad \|f\|_\infty = \max_{z \in Q} |f(z)|.$$

Their method is a two-step procedure: the first step applies the usual simplex method of linear programming to solve a discrete approximation of the SIP; the second step uses the end result of the first step as the initial starting point in a Newton-Raphson iteration to solve a certain system of nonlinear algebraic equations whose solution (if feasible) is a solution to the linear approximation problem (Problem 1 below). The purpose of this note is to make some observations about the linear program of their discrete first step, which closely connects its solution with the solution of the approximation problem. A knowledge of the SIP definition and solution method is not needed to understand the results presented here. The interested reader is referred to [1], [2], [3], and to their bibliographies. We point out that Theorems 1 and 2 were first proved in [4], where a method identical to the first step of Glashoff-Roleff's procedure for finite Q was discovered independently of knowledge of [1] and of semi-infinite programming. Readers interested in practical examples and an engineering application of linear complex approximation are also referred to [4].

Received February 24, 1982.

1980 *Mathematics Subject Classification*. Primary 65D15, 65E05, 65K05; Secondary 30C30, 41A50, 30A82.

©1983 American Mathematical Society
0025-5718/82/0000-1068/\$02.50

Let $h_1(z), \dots, h_n(z)$ and $f(z)$ be given functions in $C(Q)$. For any set of complex parameters $a = \{a_1, \dots, a_n\}$, define

$$(2) \quad L(a; z) = \sum_{k=1}^n a_k h_k(z).$$

Problem 1. Compute a set of complex parameters $a^* = \{a_1^*, \dots, a_n^*\}$ such that, for all parameter sets a ,

$$(3) \quad \|f - L(a^*; z)\|_\infty \leq \|f - L(a; z)\|_\infty.$$

We set

$$(4) \quad E_n(f) = \|f - L(a^*; z)\|_\infty.$$

Let $p \geq 2$ be a positive integer. Define the angles

$$\theta_j = \pi(j-1)/p; \quad j = 1, 2, \dots, 2p,$$

and let $S_p = \{\theta_j\}$. Define, for any complex number z ,

$$(5) \quad |z|_p = \max_{1 \leq j \leq 2p} \{\operatorname{Re}(z) \cos \theta_j + \operatorname{Im}(z) \sin \theta_j\}.$$

It may be readily verified that:

- (i) $|z|_p \geq 0$ and $|z|_p = 0$ if and only if $z = 0$.
- (ii) $|z + w|_p \leq |z|_p + |w|_p$ for all complex z and w .
- (iii) Given α complex, $|\alpha z|_p = |\alpha| \cdot |z|_p$ for all z if and only if $\arg \alpha \in S_p$.
- (iv) For α and α_n complex, $|-z|_p = |z|_p$, $\lim_{\alpha_n \rightarrow 0} |\alpha_n z|_p = 0$, and $\lim_{|z|_p \rightarrow 0} |\alpha z|_p = 0$.

Thus $|z|_p$ is not a norm on \mathbf{C} because (iii) is not sufficiently strong; however, it is a quasi-norm because of (i), (ii), and (iv). See [5, pp. 30–32]. From the well-known identity

$$(6) \quad |z| = \max_{0 \leq \theta < 2\pi} \{\operatorname{Re}(z) \cos \theta + \operatorname{Im}(z) \sin \theta\},$$

it follows that $|z|_p \leq |z|$. In addition it can be shown that

$$(7) \quad |z|_p \leq |z| \leq |z|_p \sec\left(\frac{\pi}{2p}\right), \quad p \geq 2,$$

for all complex z . To see (7), it is helpful to visualize the set of all z in \mathbf{C} such that $|z|_p = 1$ as an equilateral polygon of $2p$ sides whose inscribed circle is the unit circle $|z| = 1$.

It is easy to verify that

$$(8) \quad \|f\|_p \equiv \max_{z \in Q} |f(z)|_p$$

is a quasi-norm on $C(Q)$ for each integer $p \geq 2$. Further, from (7),

$$(9) \quad \|f\|_p \leq \|f\|_\infty \leq \|f\|_p \sec\left(\frac{\pi}{2p}\right).$$

We now define a new (partially) discretized version of Problem 1.

Problem 2. Fix $p \geq 2$. Compute a set of complex parameters $a^{**} = \{a_1^{**}, \dots, a_n^{**}\}$ such that, for all parameter sets a ,

$$(10) \quad \|f - L(a^{**}; z)\|_p \leq \|f - L(a; z)\|_p.$$

We set

$$(11) \quad E_{np}(f) = \|f - L(a^{**}; z)\|_p.$$

THEOREM 1. $E_{np}(f) \leq E_n(f) \leq E_{np}(f) \sec\left(\frac{\pi}{2p}\right)$.

Proof. We have

$$\begin{aligned} E_{np}(f) &= \|f - L(a^*; z)\|_p \leq \|f - L(a^{**}; z)\|_p \leq \|f - L(a^*; z)\|_\infty = E_n(f) \\ &\leq \|f - L(a^{**}; z)\|_\infty \leq \|f - L(a^{**}; z)\|_p \sec\left(\frac{\pi}{2p}\right) = E_{np}(f) \sec\left(\frac{\pi}{2p}\right). \end{aligned}$$

THEOREM 2. $E_n(f) \leq \|f - L(a^{**}; z)\|_\infty \leq E_n(f) \sec\left(\frac{\pi}{2p}\right)$.

Proof.

$$\begin{aligned} E_n(f) &= \|f - L(a^*; z)\|_\infty \leq \|f - L(a^{**}; z)\|_\infty \leq \|f - L(a^{**}; z)\|_p \sec\left(\frac{\pi}{2p}\right) \\ &\leq \|f - L(a^*; z)\|_p \sec\left(\frac{\pi}{2p}\right) \leq E_n(f) \sec\left(\frac{\pi}{2p}\right). \end{aligned}$$

COROLLARY 1. $E_{np}(f) \leq \|f - L(a^{**}; z)\|_\infty \leq E_{np}(f) \sec\left(\frac{\pi}{2p}\right)$.

COROLLARY 2. For each $p \geq 2$, $E_n(f) = 0$ if and only if $E_{np}(f) = 0$.

COROLLARY 3. If $E_n(f) \neq 0$,

$$(12) \quad 0 \leq \frac{\|f - L(a^{**}; z)\|_\infty - E_n(f)}{E_n(f)} \leq \frac{\pi^2}{8p^2} + O\left(\frac{1}{p^4}\right), \quad p \rightarrow \infty,$$

and the upper bound is independent of the compact set Q , n , f , and the functions h_1, \dots, h_n .

Proof. The indicated ratio is bounded above by the constant $-1 + \sec\left(\frac{\pi}{2p}\right)$.

It is not necessary that the domain of approximation Q be a subset of the complex plane \mathbf{C} . All that is required is that f and h_1, \dots, h_n be defined on a common domain Q and that a solution to Problem 1 exists.

If the point set Q is not finite, then both Problems 1 and 2 can be readily transformed into linear SIP's with linear objective functions and infinitely many linear constraints and then can be solved in the manner of Glashoff and Roleff [1]. The difference is that, for Problem 1, there is one constraint for each element of the Cartesian product $S \times Q$, where $S \equiv \{\eta \text{ in } \mathbf{C}: |\eta| = 1\}$; whereas for Problem 2, there is only one constraint for each element of $S_p \times Q$, where $S_p \equiv \{\eta \text{ in } \mathbf{C}: \eta^{2p} = 1\}$. It can happen in certain applications that the bounds proved above show that Problem 2 is adequate for some fixed $p \geq 2$. The numerical solution procedures of Glashoff and Roleff may then be appropriately, and potentially significantly, simplified.

On the other hand, if Q is finite, Problem 2 becomes an ordinary linear program, although Problem 1 remains an SIP. The finite Q case is precisely the first step of the Glashoff and Roleff method for solving Problem 1. It is not hard to see that, for $Q = \{z, \dots, z_m\} \subset \mathbf{C}$, Problem 2 may be reformulated as solving an overdetermined system of mp real linear algebraic equations in $2n$ real unknowns in the usual Chebyshev (l_∞) norm. Full details for setting up the linear equations can be found in [4]. (This formulation works for any choice of $T \equiv \{\theta_k\}$ provided only that $\theta_k \in T$ if

and only if $\theta_k + \pi \in T$.) This real system may be written in the following block-partitioned form:

$$(13) \quad \begin{bmatrix} R \cos \theta_1 + S \sin \theta_1 & R \sin \theta_1 - S \cos \theta_1 \\ R \cos \theta_2 + S \sin \theta_2 & R \sin \theta_2 - S \cos \theta_2 \\ \vdots & \vdots \\ R \cos \theta_p + S \sin \theta_p & R \sin \theta_p - S \cos \theta_p \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} u \cos \theta_1 + v \sin \theta_1 \\ u \cos \theta_2 + v \sin \theta_2 \\ \vdots \\ u \cos \theta_p + v \sin \theta_p \end{bmatrix},$$

where we define

$$\begin{aligned} x &= [\operatorname{Re}(a_k)] \in R^n, & y &= [\operatorname{Im}(a_k)] \in R^n, \\ u &= [\operatorname{Re}(f(z_k))] \in R^m, & v &= [\operatorname{Im}(f(z_k))] \in R^m. \end{aligned}$$

and the two $m \times n$ matrices

$$R = [r_{jk}] = [\operatorname{Re}(h_k(z_j))], \quad S = [s_{jk}] = [\operatorname{Im}(h_k(z_j))].$$

Computer CPU time and storage requirements may present severe practical limitations on the numerical solution of (13) in certain problems of genuine interest. See Streit and Nuttall [4] for an antenna array example with $n = 44$, $p = 8$, and $m = 501$ which required 1262 simplex iterations and 179 minutes on the DEC VAX 11/780 to solve (13) using the general purpose algorithm [6]. If, however, the special structure of (13) is exploited, very significant reductions in both time and storage requirements are possible; see [7].

At least two situations might arise where the effective use of the structure of (13) in its solution would be important. First, the Glashoff-Roleff method for any given Q requires the solution (by Newton-Raphson or any other workable iterative method) of a nonlinear system of algebraic equations. If the initial point is not sufficiently good, then this procedure either does not converge or it converges to a nonfeasible (hence, incorrect) point. Since initial points are constructed by solving (13), it is conceivable that very large systems may have to be solved (even for small n) to get a sufficiently good initial point. The other reason for studying the special structure of (13) is simply that n may be very large to begin with. In the kind of applications mentioned in [4], it would not be at all unreasonable to find $n \geq 100$. Even for small p , the system (13) is then very large. Either case presents an interesting problem with a large 100% dense linear program having special structure, instead of the more typical situation of a large sparse linear program having relatively little special structure other than sparsity.

Solving the overdetermined system (13), while requiring nonnegative residuals, can have interesting geometrical interpretations. For example, take $p = 2$ so that $\theta_1 = 0$ and $\theta_2 = \pi/2$. Thus, the $2m$ components of the residual vector of (13) are precisely the real and imaginary parts of the complex error $e(z) = f(z) - L(a; z)$ evaluated at all m data points. Requiring nonnegative residuals means that we have forced the error curve $e(z)$ to lie entirely in the first quadrant of the complex plane. Furthermore, it is easy to see that we may force $e(z)$ to lie in any convex wedge-shaped sector \mathcal{W} of the complex plane by making appropriate alternative choices of the two

angles θ_1 and θ_2 . Further exploration of this idea shows that upper and lower bounds for the error $W_n(f)$ defined by

$$W_n(f) \equiv \min_{a \in \mathbb{C}^n} \max_{z \in Q} |f(z) - L(a; z)|$$

subject to: $f(z) - L(a; z) \in \mathcal{U}, z \in Q,$

can be obtained in terms of $W_{np}(f)$ defined by

$$W_{np}(f) \equiv \min_{a \in \mathbb{C}^n} \max_{z \in Q} |f(z) - L(a; z)|_p$$

subject to: $f(z) - L(a; z) \in \mathcal{U}, z \in Q.$

This technique requires an appropriately modified set of angles $\theta_1, \dots, \theta_{2p}$. A solution of $W_{np}(f)$ can then be found numerically by computing the l_∞ solution of an overdetermined system of the form (13) with the additional requirement of nonnegative residuals.

LEMMA. *Let Q be finite. The $2n$ columns of the coefficient matrix in (13) are linearly dependent (over the real number field) if and only if the n functions $\{h_1, \dots, h_n\}$ are linearly dependent on Q (over the complex number field).*

Proof. There exist complex numbers $a_k = x_k + iy_k$, $1 \leq k \leq n$, not all zero, satisfying $\|\sum_k a_k h_k\|_\infty = 0$ if and only if $\|\sum_k a_k h_k\|_p = 0$. This latter equation is true if and only if

$$\max_{z \in Q} \left| \operatorname{Re} \left(\sum_{k=1}^n a_k h_k(z) \right) + i \operatorname{Im} \left(\sum_{k=1}^n a_k h_k(z) \right) \right|_p = 0,$$

which holds if and only if, for each $z_t \in Q$ and $\theta_j \in S_p$ ($1 \leq j \leq p$),

$$\left| \left(\sum_{k=1}^n x_k \operatorname{Re} h_k(z_t) - y_k \operatorname{Im} h_k(z_t) \right) \cos \theta_j \right. \\ \left. + \left(\sum_{k=1}^n x_k \operatorname{Im} h_k(z_t) + y_k \operatorname{Re} h_k(z_t) \right) \sin \theta_j \right| = 0.$$

Rearranging and using the notation of (13) gives

$$(R \cos \theta_j + S \sin \theta_j)x + (R \sin \theta_j - S \cos \theta_j)y = 0, \quad j = 1, \dots, p,$$

which means that the columns of the coefficient matrix in (13) are linearly dependent. This completes the proof.

THEOREM 3. *Let Q contain $m \leq 2n$ distinct points, let the functions $\{h_1, \dots, h_n\}$ be linearly independent on Q , and let a^{**} satisfy (10) where $p \geq 2$. Then*

$$(14) \quad \|f - L(a^{**}, z)\|_\infty = E_{np}(f) \sec\left(\frac{\pi}{2p}\right).$$

Proof. If f is linearly dependent on h_1, \dots, h_n , then $E_n(f) = 0$ and, from Corollary 2, $E_{np}(f) = 0$ and (14) is trivially true. Suppose then that f is linearly independent of h_1, \dots, h_n . Let a^{**} satisfy (10). Then a^{**} is a Chebyshev (l_∞) solution of the system (13), and the maximum residual has magnitude $E_{np}(f) > 0$. From the

preceding Lemma, the rank of the coefficient matrix in (13) is $2n$. Hence there exists [8, p. 29] another solution \hat{a} of (13) such that

$$\|f - L(\hat{a}; z)\|_p = E_{np}(f),$$

and a subset of at least $2n + 1$ of the mp equations (13) has residuals equal in magnitude to $E_{np}(f)$. (We cannot take $\hat{a} = a^{**}$ in general, because we have not assumed that the coefficient matrix in (13) satisfies the Haar condition for matrices.) Now these $2n + 1$ extremal equations must be distributed among the $m \leq 2n$ points of Q . Therefore, at least one point z in Q is assigned at least two equations.

Claim. No point in Q can be assigned more than two extremal equations. Note first that the residuals of the p equations in (13) corresponding to a given point z in Q are precisely

$$r_j \equiv A \cos \theta_j + B \sin \theta_j, \quad j = 1, \dots, p,$$

where A and B are the real and imaginary parts of $f(z) - L(\hat{a}; z)$, respectively. Let $K(z)$ denote the set of indices j of the extremal equations assigned to the point z . If $K(z)$ is not empty, then the equations

$$(15) \quad |A \cos \theta_j + B \sin \theta_j| = E_{np}(f), \quad j \in K(z),$$

must hold simultaneously. Since $E_{np}(f) > 0$, it is clear that, if $K(z)$ contains more than two indices, the system (15) is inconsistent. This proves our claim.

Thus, let z be a point in Q which is assigned two extremal equations. Let $K(z) = \{j, k\}$ with $j \neq k$. Then the equations (15) imply

$$|A + iB| = E_{np}(f) \sec(\phi/2),$$

where ϕ is the smallest angle measured between the four angles $\{\theta_j, \theta_k, \theta_j + \pi, \theta_k + \pi\}$. Since Theorem 1 cannot be violated, we must have $\phi = \pi/p$. This concludes the proof.

If the coefficient matrix in (13) satisfies the Haar condition, then the norm (14) is attained for at least $t = \min\{2n + 1 - m, m\}$ distinct points z in Q . In this case, $\hat{a} = a^{**}$, so every point having two of the $2n + 1$ extremal equations has the residual (14). There must be at least t such points, considering the claim established in the proof of Theorem 3.

The first author is employed by the Naval Underwater Systems Center, New London Laboratory, New London, CT 06320. This paper was written during his stay as a Visiting Scholar in the Department of Operations Research, Stanford University, Stanford, CA 94305. His work was supported jointly by the Office of Naval Research Project RR014-07-01 and by The Independent Research Program of the Naval Underwater Systems Center.

Naval Underwater Systems Center
New London, Connecticut 06320

1. K. GLASHOFF & K. ROLEFF, "A new method for Chebyshev approximation of complex-valued functions," *Math. Comp.*, v. 36, 1981, pp. 233-239.
2. A. CHARNES, W. W. COOPER & K. O. KORTANEK, "Duality, Haar programs and finite sequence spaces," *Proc. Nat. Acad. Sci. U.S.A.*, v. 48, 1962, pp. 783-786.
3. S.-A. GUSTAFSON, "On semi-infinite programming in numerical analysis," in *Semi-Infinite Programming* (R. Hettich, Ed.), Lecture Notes in Control and Information Sciences, Vol. 15, Springer-Verlag, Berlin and New York, 1979, pp. 137-153.

4. R. L. STREIT & A. H. NUTTALL, "A general Chebyshev complex function approximation procedure and an application to beamforming," *J. Acoust. Soc. Amer.*, v. 72, 1982, pp. 181–190; Also NUSC Technical Report 6403, 26 February 1981. (Naval Underwater Systems Center, New London, Connecticut, U.S.A.)
5. K. YOSIDA, *Functional Analysis*, 2nd ed., Springer-Verlag, Berlin and New York, 1968.
6. I. BARRODALE & C. PHILLIPS, "Solution of an overdetermined system of linear equations in the Chebyshev norm," Algorithm 495, *ACM Trans. Math. Software*, v. 1, 1975, pp. 264–270.
7. R. L. STREIT, *Numerical Solutions of Systems of Complex Linear Equations with Constraints on the Unknowns*, Stanford University Department of Operations Research SOL Report. (To appear.)
8. G. A. WATSON, *Approximation Theory and Numerical Methods*, Wiley, New York, 1980.