REAL VS. COMPLEX RATIONAL CHEBYSHEV APPROXIMATION
ON AN INTERVAL

BY

LLOYD N. TREFETHEN1 AND MARTIN H. GUTKNECHT

Abstract. If \( f \in C[-1,1] \) is real-valued, let \( E'_r(f) \) and \( E'_c(f) \) be the errors in best approximation to \( f \) in the supremum norm by rational functions of type \((m, n)\) with real and complex coefficients, respectively. It has recently been observed that \( E'_c(f) < E'_r(f) \) can occur for any \( n \gg 1 \), but for no \( n \gg 1 \) is it known whether \( \gamma_{mn} = \inf_ f E'_c(f)/E'_r(f) \) is zero or strictly positive. Here we show that both are possible: \( \gamma_{01} > 0 \), but \( \gamma_{mn} = 0 \) for \( n \gg m + 3 \). Related results are obtained for approximation on regions in the plane.

1. Introduction. Let \( I \) be the unit interval \([-1,1] \), \( C' \) the set of continuous real functions on \( I \), and \( \| \cdot \| \) the supremum norm \( \| f \| = \sup_{x \in I} |f(x)| \). For nonnegative integers \( m \) and \( n \), let \( R_{mn} \) and \( R'_{mn} \subseteq R_{mn} \) be the spaces of rational functions of type \((m, n)\) with coefficients in \( \mathbb{C} \) and \( \mathbb{R} \), respectively. For \( f \in C' \), let \( E'_c(f) \) and \( E'_r(f) \) denote the infima

\[
E'_c(f) = \inf_{r \in R'_{mn}} \| f - r \|, \quad E'_r(f) = \inf_{r \in R_{mn}} \| f - r \|.
\]

It is known that both limits are attained, and a function that does so is called a best approximation (BA) to \( f \). In the real case the BA is unique [8], and in the complex case for \( n \gg 1 \) in general it is not [7, 10, 11, 14, 15].

Obviously \( E'_c \leq E'_r \) for any \( f \), but since \( f \) is real, it is not at first obvious whether a strict inequality can occur. However in 1971 Lungu [7], following a proposal of Gončar [16], published a class of examples showing that \( E'_c(f) < E'_r(f) \) is indeed possible if \( n \gg 1 \). Independently, Saff and Varga [10, 11] made the same discovery in 1977, and obtained more general sufficient conditions for \( E'_c(f) < E'_r(f) \) and also a sufficient condition for \( E'_c(f) = E'_r(f) \). The former was later sharpened by Ruttan [18] to the following statement: \( E'_c(f) < E'_r(f) \) must hold if the best real approximation to \( f \) attains its maximum error on no alternation set of length greater than \( m + n + 1 \) points. For a survey of such results, see [14].

But is \( E'_c \) ever much less than \( E'_r \)? If \( \gamma_{mn} \) denotes the infimum

\[
\gamma_{mn} = \inf_{f \in C' \setminus R'_{mn}} E'_c(f)/E'_r(f),
\]

then one would like to know whether \( \gamma_{mn} \) can be zero or is always positive, and if the latter, how small it is. In all of the examples devised to date, \( E'_c(f)/E'_r(f) \) has fallen

\[\text{License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use}\]
in the range \((\frac{1}{2}, 1]\), suggesting that \(\gamma_{mn} = \frac{1}{2}\) might be the minimum value. Saff and Varga posed in particular the question, is \(\gamma_{mn}\) positive or zero \([10, 11]\)? Ellacott has suggested that \(\gamma_{mn} = \frac{1}{2}\) may hold for \(m \geq n\) \([3]\). (For more on his argument see §2.) Some partial results for \((m, n) = (1, 1)\) have been obtained by Bennet, et al. \([1, 2]\) and by Ruttan \([9]\).

In this paper we resolve some of these questions, as follows. First, not only can \(\gamma_{mn} < \frac{1}{2}\) occur, but \(\gamma_{mn} = 0\) for all \(m \geq 0, \ n \geq m + 3\) (Theorem 1). Second, \(\gamma_{01} > 0\) (Theorem 2). We conjecture that \(\gamma_{mn} > 0\) holds whenever \(n < m + 3\). Finally, at least some of our arguments extend to approximation on complex regions, and we show: \(\gamma^\Delta_{0n} = 0\) for \(n \geq 4\) in approximation on the unit disk \(\Delta\) (Theorem 3). A similar result is obtained for approximation on a symmetric Jordan region.

2. \(\gamma_{mn} = 0\) for \(n \geq m + 3\).

**THEOREM 1.** \(\gamma_{mn} = 0\) for all \(m \geq 0, \ n \geq m + 3\).

**PROOF.** The idea of the construction is indicated in Figure 1, where crosses represent poles and circles represent zeros.

Given \(m \geq 0\), let \(\phi \in R_{m,n+3}\) be defined by

\[
\phi(x) = \frac{\varepsilon \prod_{j=1}^{m} \left[(-1 + (2j - 1)\varepsilon) - x\right]}{[x + (1 + \varepsilon)]^{m+1} \left[\frac{1}{i\varepsilon} - x\right]} \left[\frac{1}{i\varepsilon} - x\right]
\]

and as the function in \(C^r\) to be approximated take \(f(x) = \text{Re} \, \phi(x)\). We will show that \(f\) has the following two properties:

(a) \(\|f - \phi\| = \|\text{Im} \, \phi\| = O(\varepsilon)\) as \(\varepsilon \to 0\).
(b) There exists a constant \(C > 0\) such that for all sufficiently small \(\varepsilon\),

\[
(-1)^j f(-1 + 2\varepsilon j) \geq C, \quad 0 \leq j \leq m,
\]

and

\[
(-1)^{m+1} f(1) \geq C.
\]

Condition (b) states that the error function for the zero approximation to \(f\) approximately equioscillates at \(m + 2\) points, and by the de la Vallée Poussin theorem for real rational approximation \([8, \text{Theorem 98}]\), this implies \(E^r \geq C\). (For the purposes of this theorem \(r \equiv 0\) has rational type \((\mu, \nu) = (-\infty, 0)\), so the "defect" \(d = \min(m - \mu, n - \nu)\) is \(n\), which means one needs approximate equioscillation at \(m + n + 2 - d = m + 2\) points.) On the other hand if \(n \geq m + 3\), then \(\phi \in R_{mn}\), so (a) implies \(E^c = O(\varepsilon)\). Thus since \(\varepsilon\) can be arbitrarily small, the theorem will be proved once (a) and (b) are established.
Proof of (a). Let us write \( \phi \) as a product of three functions \( \phi_1, \phi_2, \phi_3 \) corresponding to the poles and zeros near \(-1, 0, \) and \(1\), respectively. Of these functions only \( \phi_2 \) has a nonzero imaginary part on \( I \), and we bring this into the numerator. The factor \( \phi_2 \) gets the constant \( \epsilon \) from (3):

\[
\phi(x) = \phi_1(x) \phi_2(x) \phi_3(x)
\]

\[
= \left( \frac{\prod_{j=1}^{\infty} \left( \frac{-1 + (2j - 1) \epsilon}{x + (1 + \epsilon)} \right)^{m+1}}{x^2 + \epsilon} \right) \left( \frac{-i\sqrt{\epsilon} - x}{(1 + \epsilon) - x} \right).
\]

Since \( (f - \phi)(x) = -i \Im \phi(x) \), we compute

\[
(f - \phi)(x) = -i \phi_1(x) \Im \phi_2(x) \phi_3(x) = \phi_1(x) \frac{-i\sqrt{\epsilon}}{x^2 + \epsilon} \phi_3(x).
\]

It is not hard to see that on \([-1, -\frac{1}{2}]\) these factors have magnitude \( O(1), O(\sqrt{\epsilon}) \), and \( O(1) \), so their product is \( O(\sqrt{\epsilon}) \). Similarly in \([-\frac{1}{2}, \frac{1}{2}]\) one has \( O(\epsilon) O(1/\sqrt{\epsilon}) O(1) = O(\sqrt{\epsilon}) \), and in \([\frac{1}{2}, 1]\), \( O(\epsilon) O(1/\sqrt{\epsilon}) O(1/\epsilon) = O(\sqrt{\epsilon}) \). Together these estimates give \( (f - \phi)(x) = O(\sqrt{\epsilon}) \) for all \( x \in I \), as claimed.

Proof of (b). Again we use the factorization \( \phi = \phi_1\phi_2\phi_3 \) of (6). Let \( \{x_j\}_{j=0}^{\infty} \) be the set of points \( x_j = -1 + 2j \epsilon \) that appear in condition (4). At each \( x_j \), \( \phi_1 \) evidently takes the form \( a_j e^{\alpha x_j} + b_j e^{\beta x_j} \) for some constants \( a_j \) and \( b_j \), and thus \( \phi_1(x_j) \) is independent of \( \epsilon \). Moreover these quantities obviously alternate in sign, i.e.

\[
\phi_1(x_0) = \tau_0 > 0, -\phi_1(x_1) = \tau_1 > 0, \ldots, (-1)^m \phi_1(x_m) = \tau_m > 0,
\]

with \( \tau_j \) independent of \( \epsilon \). In addition since all of the points \( x_j \) are contained in \([-1, -1 + 2m \epsilon]\), we have \( \phi_2(x_j) = 1 + O(\sqrt{\epsilon}), \phi_3(x_j) = \frac{1}{2} + O(\epsilon) \) on \( \{x_j\} \). Together these facts establish (4) for some \( C = C_1 > 0 \).

For condition (5) we compute

\[
\phi(1) = \phi_1(1) \phi_2(1) \phi_3(1)
\]

\[
= \left( \frac{\epsilon}{2} (-1)^m (1 + O(\epsilon)) \right) \left( -1 + O(\sqrt{\epsilon}) \right) \frac{1}{\epsilon} = \frac{1}{2} (-1)^{m+1} + O(\sqrt{\epsilon}),
\]

which implies that (5) holds for \( C = C_2 \) with any \( C_2 < \frac{1}{2} \). Taking \( C = \min(C_1, C_2) \) now yields (b). \( \square \)

Remark on an Argument of Ellacott. As alluded to in the Introduction, Ellacott has observed that one can conclude from the CF method [13, 4] that if \( p \) is a polynomial of degree \( m + 1 \), then

\[
E'(p)/E'(p) \geq \frac{1}{2}
\]

for \( n \leq m \) [3]. This is one of his arguments for suggesting that \( \gamma_{mn} = \frac{1}{2} \) or at least \( \gamma_{mn} > 0 \) may hold for \( n \leq m \). However we claim that (7) is valid in fact for all \( n \leq 2m + 1 \), which by Theorem 1 means that it holds even in many cases with \( \gamma_{mn} = 0 \). Therefore although Ellacott's conjecture is plausible, it appears that (7) does not provide very strong support for it.
To demonstrate that (7) holds for \( n \leq 2m + 1 \), let \( p \) be transplanted to the unit circle by defining a function \( \hat{p} \) for \( z \in \mathbb{C} \) as follows:

\[
x = \frac{1}{2}(z + z^{-1}), \quad \hat{p}(z) = p(x) = p\left(\frac{1}{2}z + \frac{1}{2}z^{-1}\right) = \sum_{k=-m-1}^{m+1} \alpha_k z^k.
\]

For \( n \leq 2m + 1 \), the BA to \( p \) in \( R_{mn}^{'} \) on \( I \) was obtained explicitly by Talbot [12, 5], and its deviation from \( p \) is

\[
E'(p) = 2\sigma_n,
\]

where \( \sigma_n \) is the smallest singular value of the \((n + 1) \times (n + 1)\) Hankel matrix \((\alpha_{m-n+1+i+j})_{i,j=0}^{n} \). On the other hand if \( r \in R_{mn}^{'} \) is any complex approximation to \( p \) on \( I \), consider the transplanted function \( \hat{r} \) defined by \( \hat{r}(z) = r(x) \). It is readily verified that \( \hat{r} \) has \( \nu \leq n \) poles in \( 1 <|z|< \infty \) and is of order \( O(z^{m-r}) \) at \( \infty \). Therefore \( \hat{r} \) lies in the space \( \hat{R}_{mn}^{'} \) defined in [13, 4], and by the theory given there this implies

\[
\sigma_n \leq \sup_{|z|=1} |(\hat{p} - \hat{r})(z)| = \sup_{|x|=1} |(p - r)(x)|.
\]

Thus

\[
E'(p) \geq \sigma_n,
\]

which together with (8), establishes (7).

By applying [4, Lemma 5.1 in Part II] (7) can be seen to hold even for some rational functions \( f \), namely for those of exact type \((M, N)\) where either \( M \leq m + 1 \), \( N = n + 1 \), \( n \leq m \) or \( M = m + 1 \), \( N = n + 1 \), \( n \leq 2m + 1 - N \); details will be given in [5].

3. \( \gamma_{01} > 0 \).

**Theorem 2.** \( \gamma_{01} > 0 \).

**Proof.** Let \( f \in C^\prime \) be arbitrary, and let \( c^* \) be a BA to \( f \) in \( R_{mn}^{'} \). Then for any \( r \in R_{mn}^{'} \) one has \( ||\text{Im} c^*|| \leq ||f - c^*|| = E'^*(f) \) and \( E'(f) \leq E'^*(f) + ||c^* - r|| \), and therefore

\[
E'(f) \leq E'^*(f) + ||\text{Im} c^*|| \frac{||c^* - r||}{||\text{Im} c^*||} \leq E'^*(f) \left(1 + \frac{||c^* - r||}{||\text{Im} c^*||}\right).
\]

Now suppose that for any \( c \in R_{mn} \setminus R_{mn}^{'} \) with no poles on \( I \), one can find \( r(c) \in R_{mn}^{'} \) such that

\[
||c - r(c)||/||\text{Im} c|| \leq M
\]

for some fixed \( M \). Then \( r(c) \) can be inserted in (10), independent of \( f \), and one obtains \( \gamma_{mn} \geq 1/(1 + M) \). Our proof of \( \gamma_{01} > 0 \) consists of exhibiting a mapping \( c \mapsto r(c) \) for the case \((m, n) = (0, 1)\) that satisfies (11).

Thus let \( c(z) = a/(1 - z/z_0) \) be given, where \( z_0 \) lies in the region \( C^0 = \mathbb{C} \cup \{\infty\} \setminus I \). Let \( \theta \in (0, \pi/2) \) and \( \rho \in (1, \infty) \) be arbitrary fixed constants (say,
\( \theta = \pi/4, \rho = 2 \). Our choice of \( r^{(c)} \) depends on which of four domains \( A^+, A^-, B, C \) the pole lies in:

\[
A^+ = \{ z \in \mathbb{C} : |\arg(-1 \pm z)| < \theta \}, \\
B = \{ z \in \mathbb{C} - A^+ - A^- : |z| \leq \rho \}, \\
C = \mathbb{C}^0 - A^+ - A^- - B.
\]

The configuration is indicated in Figure 2.

![Figure 2](image)

We define \( r^{(c)} \) as follows:

For \( z_0 \in A^+ \):
\[
\frac{1}{|z_0|} \left( 1 - \frac{1}{|z_0|} \right) \Re c(\pm 1).
\]

For \( z_0 \in B \):
\[
r^{(c)} \equiv 0.
\]

For \( z_0 \in C \):
\[
r^{(c)} \equiv \Re a.
\]

The proof can now be completed by showing that there exist constants \( M_A, M_B, M_C \) such that (11) holds for \( z_0 \) restricted to each domain \( A^+ \cup A^-, B, C \). The global constant \( M \) can then be taken as \( M = \max\{M_A, M_B, M_C\} \). The algebra involved is unfortunately quite tedious, so we will omit these verifications. However, details of a similar argument for the case of approximation on certain Jordan regions in \( \mathbb{C} \) are given in [17].

4. \( \gamma_{0n}^\Delta = 0 \) for \( n \geq 4 \).

Let \( \Delta \) be the closed unit disk \( \{ z \in \mathbb{C} : |z| \leq 1 \} \), and let \( f \) be continuous in \( \Delta \) and analytic in the interior and satisfy \( f(\tilde{z}) = \tilde{f}(z) \). Let \( \|f\|_\Delta \) denote \( \sup_{z \in \Delta} |f(z)| \), and define \( E^\prime(f; \Delta), E^\prime(\tilde{f}; \Delta) \), and \( \gamma_{mn}^\Delta \) as in (1) and (2). Until recently it was not even known whether \( \gamma_{mn}^\Delta < 1 \) is possible, but in a separate paper we show that this inequality holds at least for all pairs \( (m, n) \) with \( m = 0, n \geq 1 \) or \( m \geq 0, n = 1 \) [6].

By a variation of the argument of §2, we will now prove

**Theorem 3.** \( \gamma_{0n}^\Delta = 0 \) for \( n \geq 4 \).
Proof. Let \( \xi = e^{i \theta} \) for some fixed \( \theta \in (0, \pi) \), and for any \( \epsilon > 0 \), define
\[
\phi(z) = \frac{\epsilon(1 - \xi)^2}{[z + (1 + \epsilon)][(1 + \epsilon) - z][z - (1 + \epsilon/3)\xi]^2}
\]
and
\[
f(z) = \frac{1}{2} \left( \phi(z) + \overline{\phi(z)} \right).
\]
In analogy to the proof of Theorem 1, \( \gamma_{0n} = 0 \) for \( n \geq 4 \) will follow from the properties
\begin{enumerate}
\item[(a)] \( \|f - \phi\|_\Delta = O(\epsilon^{1/3}) \);
\item[(b)] there exists a constant \( C > 0 \) such that for all sufficiently small \( \epsilon, f(-1) < -C, f(1) > C \).
\end{enumerate}
Both (a) and (b) can be readily derived by observing that the term
\[
(1 - \xi)^2/[z - (1 + \epsilon/3)\xi]^2
\]
behaves like \( 1 + O(\epsilon^{1/3}) \) near \( z = 1 \) and like \(-[(1 - \xi)/(1 + \xi)]^2 + O(\epsilon^{1/3}) \) near \( z = -1 \). We omit the details. \( \square \)

This argument can be extended to show \( \gamma_{0n} = 0 \) for \( n \geq 4 \) for approximation on any Jordan region \( \Omega \) with \( \Omega = \overline{\Omega} \), provided \( \partial \Omega \) is differentiable at its two points of intersection with \( \mathbb{R} \), say \( z_1 \) and \( z_2 \), hence forms a right angle to \( \mathbb{R} \) at these points. Again one introduces a complex double pole, slightly above the point \( z_1 \) (analogous to taking \( \xi = e^{i \theta} \) with \( \theta \) small above), and this generates an approximate sign change between \( \phi(z_1) \) and \( \phi(z_2) \).

One can also prove \( \gamma_{01} > 0 \) for the same class of regions \( \Omega \). See [17].

Note added in proof. After studying the present paper, E. Saff has pointed out to us that the existence of arbitrarily small numbers \( \gamma_{mn} \) is implied by a result of Walsh in 1934 [19, Theorem IV], although this consequence was never recognized. Walsh showed that for any \( m \geq 0 \), the family \( \bigcup_{n=0}^{\infty} R_{mn} \) is dense in \( C[1] \) (or indeed in the space of continuous functions on any Jordan arc in \( \mathbb{C} \)), so that \( \lim_{n \to \infty} E_{mn}(f) = 0 \) for \( f \in C[1] \). On the other hand, as we have seen, if \( f \) has \( m + 1 \) zeros, then it cannot be approximated arbitrarily closely in \( \bigcup_{n=0}^{\infty} R_{mn}^* \), i.e. \( \lim_{n \to \infty} E_{mn}^*(f) > 0 \). It follows that for any \( m \geq 0 \), \( \lim_{n \to \infty} \gamma_{mn} = 0 \).

References


COURANT INSTITUTE OF MATHEMATICAL SCIENCES, NEW YORK UNIVERSITY, NEW YORK, NEW YORK 10012

SEMINAR FÜR ANGEWANDTE MATHEMATIK, ETH, ZÜRICH, SWITZERLAND