JOURNAL OF APPROXIMATION THEORY 31, 344–367 (1981)

Near-Circularity of the Error Curve In Complex Chebyshev Approximation

LLOYD N. TREFETHEN

Computer Science Department, Stanford University, Stanford, California 94305

Communicated by Carl de Boor

Received September 5, 1979

Let $f(z)$ be analytic on the unit disk, and let $p^*(z)$ be the best (Chebyshev) polynomial approximation to $f(z)$ on the disk of degree at most n. It is observed that in typical problems the "error curve," the image of the unit circle under $(f-p^*)(z)$, often approximates to a startling degree a perfect circle with winding number $n+1$. This phenomenon is approached by consideration of related problems whose error curves are exactly circular, making use of a classical theorem of Carathéodory and Fejér. This leads to a technique for calculating approximations in one step that are roughly as close to best as the best approximation error curve is close to circular, and hence to strong theorems on near-circularity as the radius of the domain shrinks to 0 or as n increases to ∞. As a computational example, very tight bounds are given for approximation of e^z on the unit disk. The generality of the near-circularity phenomenon (more general domains, rational approximation) is discussed.

1. INTRODUCTION

Let Ω be a Jordan region in the complex plane bounded by a Jordan curve Γ, and let $f(z)$ be analytic in Ω and continuous on $\bar{\Omega} \equiv \Omega \cup \Gamma$. Let P_n be the set of polynomials in z of degree at most n. Let $\| \; \|_\Omega$ denote the supremum norm on Ω. The following is the *polynomial* (*Chebyshev*) *approximation problem*: given f, find a polynomial $p^*(z) \in P_n$ such that $\|f-p^*\|_\Omega = \inf_{p\in P_n} \|f-p\|_\Omega$. Such a function is a *best approximation* to f in Ω. A best polynomial approximation always exists for this problem, and it is unique [17].

From the maximum modulus principle it follows that we may dispense with the interior of $\bar{\Omega}$, attempting only to minimize the norm $\|f-p\|_\Gamma \equiv \sup_{z\in\Gamma} |f(z)-p(z)|$. Now for $p \in P_n$, the image $(f-p)(\Gamma)$ describes some curve in the complex plane, which we call the *error curve* corresponding to p. The approximation problem may be restated geometrically: What choice

0021-9045/81/040344-24$02.00/0

of p, if any, leads to an error curve which may be contained in a disk of minimal radius about the origin?

What motivates this paper is the discovery, based on numerical experiments, that *in many cases the error curve* $(f-p^*)(\Gamma)$ *has nearly the shape of a perfect circle with winding number* $n+1$. This effect is much stronger than can be fully explained by simple arguments. It has been observed in isolated cases before (see [3, 4, 9]), but not accounted for.

Section 2 is devoted to presenting numerical evidence of the near-circularity phenomenon. In Section 3 we show, using Rouché's theorem, that if $(f-p)(\Gamma)$ is perfectly circular with winding number $\geqslant n+1$ for some p then p must be a best approximation, but observe that this can only occur on the disk if the error function $f-p$ is a finite Blaschke product, which implies that f is itself rational. So our approach becomes the consideration of *nearby* problems with exactly circular error curves. In Section 4 we present our tool for finding such nearby problems: the so-called Carathéodory–Fejér problem, whose solution is always characterized by a perfectly circular error curve. The C–F theory is exploited in Section 5 to give a method for constructing a near-best approximation to f whose error curve is nearly circular. A posteriori estimates are then presented which show that such an approximation must be close to best. These results are applied in Section 6 to show that if f is approximated successively on disks of radius R shrinking to the origin, then as $R \to 0$ the best approximation error curves approach perfect circles in shape with a relative deviation of magnitude $O(R^{n+2})$. A similar result is proved for approximation on the unit disk by polynomials of degree n as $n \to \infty$.

The Carathéodory–Fejér method is fully constructive, and in Section 7 it is applied to a familiar test example, approximation of e^z on the unit disk, allowing us to compute nearly exact best approximations in one step by solving a matrix eigenvalue problem.

A consequence of a nearly circular error curve is that *Lawson's algorithm* for computing best approximations essentially stops converging once it gets close to the solution.

The near-circularity phenomenon is more general than the case concentrated on here: it appears also in rational approximation and in approximation over regions that are not disks. The generality of the phenomenon and of the techniques used in approaching it here are discussed in the final section. We find that the Carathéodory–Fejér strategy applies in principle on any Jordan region.

Two papers which helped to motivate this work are those of Motzkin and Walsh [11] and Saff [14]. These study zeros of the error function $f-p^*$ in the two asymptotic limits $R \to 0$ and $n \to \infty$, respectively, and their central results are reproduced here as Theorems 10i and 11i. The work most closely related to this in spirit is that of S. J. Poreda, who in a sequence of papers has been concerned with the winding number and modulus of error curves in various complex approximation problems. See, for example, [12, 13]; in the latter, Poreda also makes use of the Carathéodory–Fejér theorem. His main concern, however, has been the approximation of functions that are not analytic and therefore cannot be approximated arbitrarily closely by polynomials.

2. THE NEAR-CIRCULARITY PHENOMENON

Let us begin with our standard computational example: e^z on the unit disk. The simplest approximation from P_n we might construct would be the first $n+1$ terms of the Taylor series of e^z, which in fact is also the best approximation to e^z on the unit circle out of P_n in the least-squares sense. Error curves corresponding to this choice of approximation are shown in the left column of Fig. 1 for $n = 0$, 1, and 2. Each error curve has winding number $n+1$ but is not very circular in shape. In the right column of the same figure, error curves corresponding to the true Chebyshev approximations (computed numerically—see Section 7) are shown. These too wind $n+1$ times, but now for $n = 1$ and $n = 2$ the curves are very close to perfect circles. Let $E^*_{\min}$ and E^* denote the minimum and maximum values of $|(f - p^*)(z)|$ on the unit circle. Then in fact we find:

n	$E^*_{\min}$	E^*	$E^* - E^*_{\min}$
0	1.082	1.321	0.239
1	0.5559	0.5584	0.25×10^{-2}
2	0.177369	0.177376	0.7×10^{-5}
3	0.043369	0.043369	$< 0.2 \times 10^{-6}$

Plots like those of Fig. 1 are given in [16] for a number of other choices of $f(z)$ on the unit disk. For $n = 2$, say, the error curve in "most" cases deviates from a circle by no more than 1 or 2%. This is true, for example, for the following analytic functions $f(z)$: $\sqrt{z-2}$, $\ln(z-2)$, $\Gamma(z+3)$, e^{e^z}, $\arctan(z/2)$. It is not true, on the other hand, for certain other functions: e^{6z}, $z^4 + z^5$. Figure 2 shows best approximation error curves for the cases $n = 0, 1, 2, 3$ for the function $f(z) = 1/\Gamma(z)$. The behavior here is typical for a function that shows the near-circularity phenomenon poorly: for $n = 1$ or $n = 2$ the error curve is clearly not a circle, but as n increases it becomes closer to one, at least along most of its length.

Our computational experience for regions other than the disk is limited, but the same phenomenon apparently occurs to a weaker extent quite generally. Figures 3 and 4 give some indication of this. Figure 3 shows

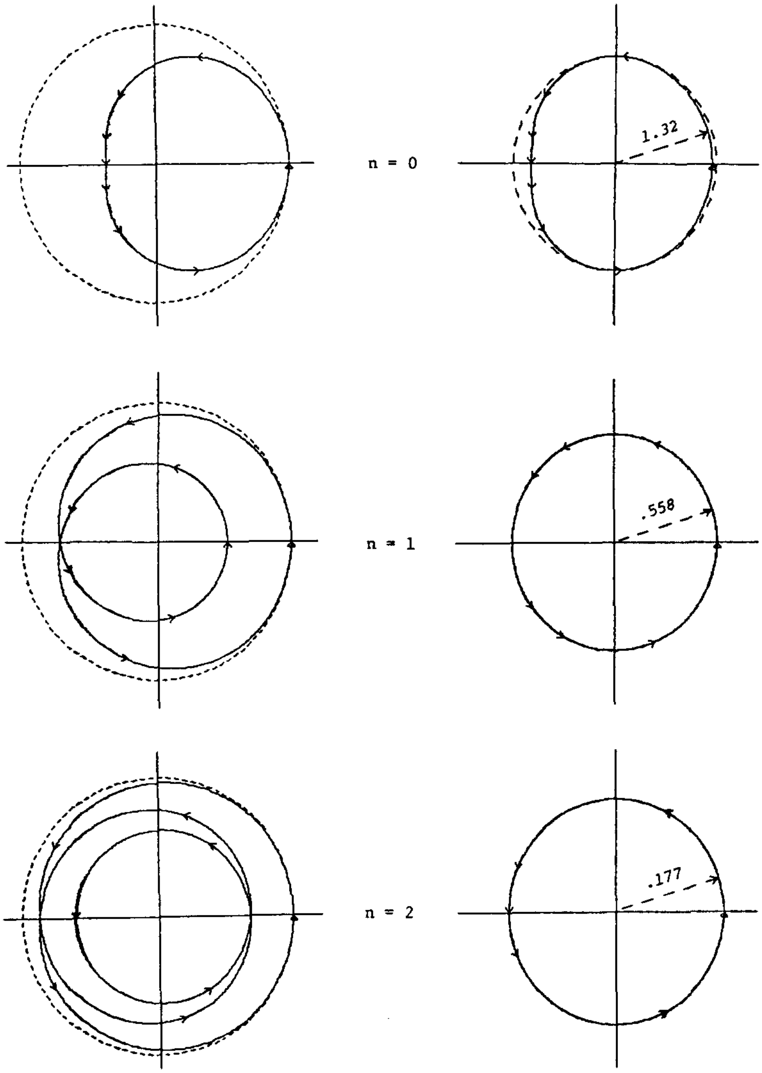

FIG. 1. Error curves for polynomial approximation to e^z over the unit disk. Left column shows partial sums of the Taylor series, right column shows Chebyshev approximations. Each pair is plotted on a single scale.

approximation to $1/(z-i)$ over an ellipse, and Fig. 4 best approximation to e^z over a square. In the latter case no polynomial can ever eliminate the four corners, since the error function is a conformal map, but even here the error curve seems to hew closer to a circle along most of its length as n increases.

This is the phenomenon which we would like to understand and exploit.

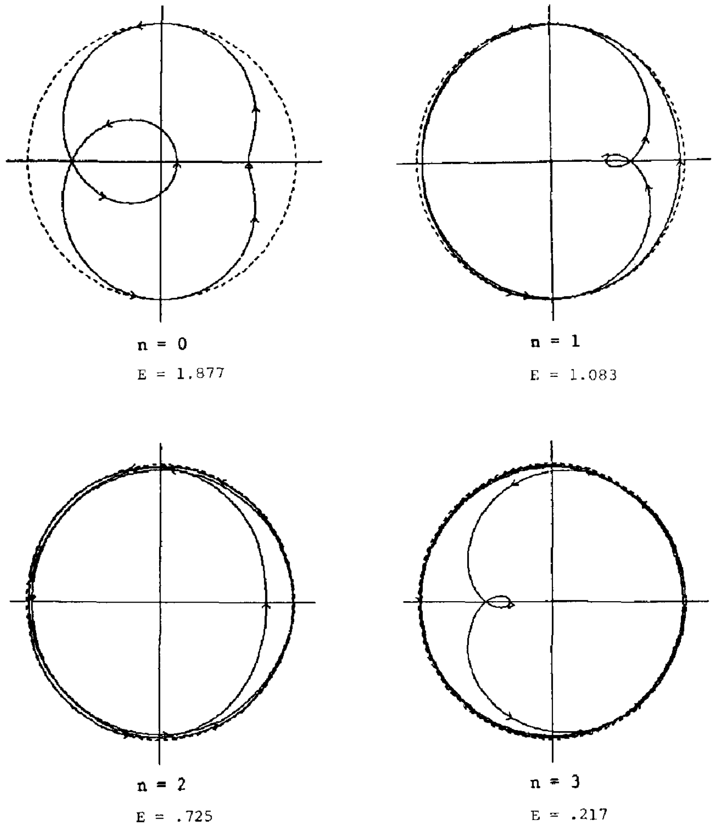

FIG. 2. Error curves for best polynomial approximation to $f(z) = 1/\Gamma(z)$ over the unit disk.

3. PERFECTLY CIRCULAR ERROR CURVES AND BLASCHKE PRODUCTS

If the error curve for some $p(z)$ happens to be a perfect circle with sufficiently large winding number, we find in an elementary fashion that p must be p^*. This is analogous to one half of the equioscillation theorem for real Chebyshev approximation. This theorem appeared in [8, 10, 16]. The same proof works for rational and polynomial approximation, so we state it in this generality. A rational function of *type* (n, m) is one which may be expressed as a quotient p/q with $p \in P_n$, $q \in P_m$.

THEOREM 1. *Let $f(z)$ be analytic in a Jordan region Ω bounded by Γ and continuous on $\bar{\Omega}$. Suppose that $r(z)$ is a rational approximation to f of*

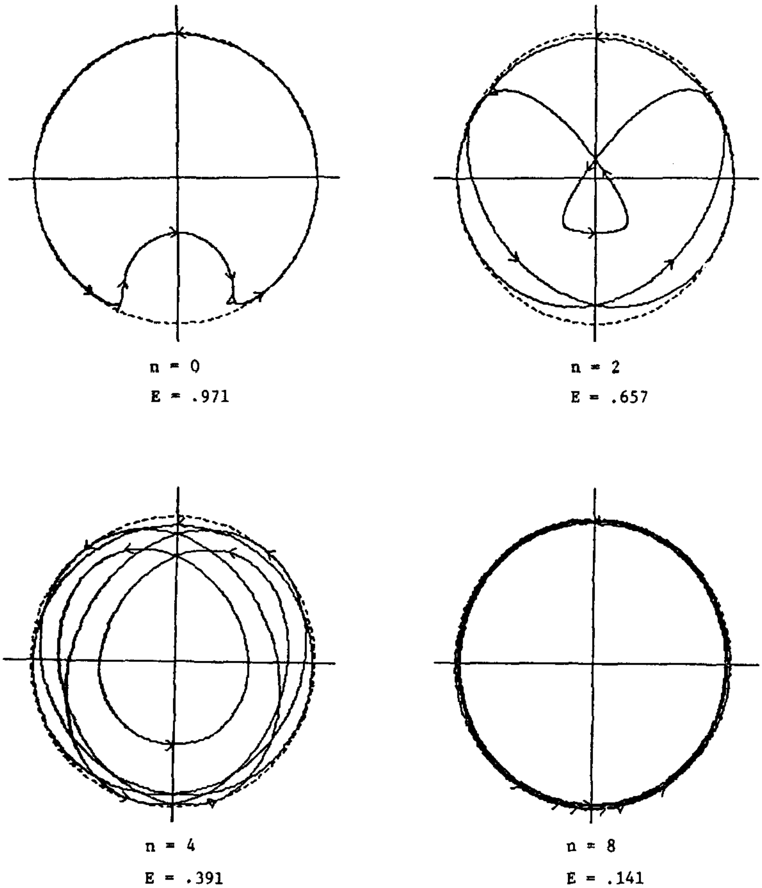

FIG. 3. Error curves for best polynomial approximation to $f(z) = 1/(z-i)$ over the ellipse passing through the points ± 2, $\pm \frac{1}{2}i$.

type (n, m), analytic in Ω, such that the error function $(f-r)(z)$ maps Γ onto a perfect circle about the origin with positive winding number $\geqslant n+m+1$. Then r is a best approximation to f over Ω among rational functions of type (n, m) which are analytic in Ω.

Proof. Suppose on the contrary that there exists some rational function $\tilde{r}$ of type (n, m) such that $\|f-\tilde{r}\| < \|f-r\|$ over Ω, hence the same over Γ. Since $(f-r)(\Gamma)$ is a circle, it follows that $|(f-\tilde{r})(z)| < |(f-r)(z)|$ for every $z \in \Gamma$. Therefore, by Rouché's theorem, $r-\tilde{r}$ must have the same number of zeros interior to Γ as $f-r$, which must be at least $m+n+1$ by the argument principle. This is impossible since $r-\tilde{r}$ is rational of type $(n+m, 2m)$. ∎

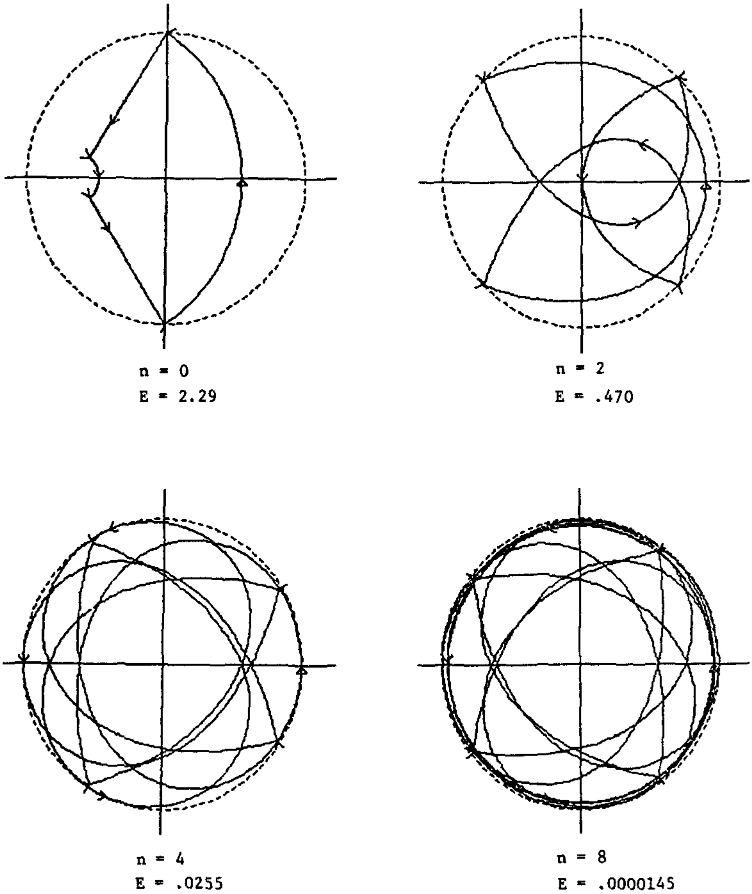

FIG. 4. Error curves for best polynomial approximation to $f(z)=e^z$ over the square with corners $\pm 1 \pm i$.

However, as one might expect, a perfectly circular error curve cannot, in general, be achieved:

THEOREM 2. *Let Ω be a disk, Γ the boundary circle, and r any rational function. Let f be analytic in Ω and continuous on $\bar{\Omega}$, and suppose that $f-r$ maps Γ onto a perfect circle about the origin. Then f is rational.*

Proof. By the reflection principle, $f-r$ can be extended to a function meromorphic in the plane whose poles and zeros are symmetric with respect to Γ. As $f-r$ is analytic at the origin, this extension has at worst a pole at infinity. Thus it is in fact meromorphic in the extended plane, hence rational. Therefore f is rational also. ∎

We can be more concrete. It is well known that the set of rational functions which are analytic in the unit disk and map the unit disk onto itself are the *finite Blaschke products*

$$\lambda \frac{\prod_{k=1}^{K} (z - \zeta_k)}{\prod_{k=1}^{K} (\bar{\zeta}_k z - 1)},$$

where K is finite, each root ζ_k satisfies $|\zeta_k| < 1$, and λ is a constant of modulus 1. If we let λ take on any complex value, we obtain the set of maps which map the unit circle onto any circle about the origin. If we fix $K = n + 1$ as suggested by Theorem 1 we see that the space of $n + 1$ – winding Blaschke products has complex dimension $n + 2$. In other words, there are quite a lot of them—which will be the basis of our success in finding that the problem of approximating a given function $f(z)$ is typically very near to a related problem in which an exactly circular error curve can be achieved. This approach is an extension of that in [1], where Blaschke products are considered that have one free zero anywhere in the unit disk but the remaining n at the origin.

4. The Carathéodory–Fejér Theorem

We may restate the *polynomial approximation problem* on the disk as follows: Given the tail of a power series, $f(z) = c_{n+1} z^{n+1} + c_{n+2} z^{n+1} + \cdots$, analytic in the unit disk, how can $f(z)$ be extended backwards by adding a polynomial $c_0 + \cdots + c_n z^n$ in such a way that $\| \sum_{k=0}^{\infty} c_k z^k \|$ is minimized? The reverse of this problem is the *Carathéodory–Fejér approximation problem*: Given a polynomial $p(z) = c_0 + \cdots + c_n z^n$, how can $p(z)$ be extended to a power series, analytic in the unit disk, of minimal norm?

In polynomial approximation we have only $n + 1$ coefficients to choose, but in C–F approximation we have an infinite number. So perhaps it is not surprising that in the C–F case, the appearance of a circular error curve is not only a sufficient but also a necessary condition for a best approximation:

Theorem 3 (Carathéodory and Fejér). *Given $\nu \geqslant 0$ and a polynomial $p(z) = c_0 + \cdots + c_\nu z^\nu$, there exists a unique power series extension $q(z) = c_0 + \cdots + c_\nu z^\nu + c_{\nu+1} z^{\nu+1} + \cdots$ which is analytic in the unit disk and minimizes $\|q\|$ among all such analytic extensions. Moreover, $q(z)$ is a finite Blaschke product with at most ν zeros in the disk, and it is the only extension of $p(z)$ to a finite Blaschke product with at most ν zeros in the disk that is analytic in the disk.*

Proof. See Ref. [5, pp. 497–506] or [6, pp. 154–163]. ∎

The similarities between the polynomial approximation problem and the C–F approximation problem are many and close. In both cases *existence* and *uniqueness* of a best approximation are assured. As another example, Rouché's theorem may be used to prove that "*circular implies best*" in the C–F problem as it was in the polynomial case (previous section). We argue: Suppose that $q(z)$ is a power series extension of $p(z)$ which maps the unit circle onto a circle of winding number at most ν. Suppose $\tilde{q}(z)$ is another extension of $p(z)$ such that $\|\tilde{q}\| \leqslant \|q\|$ on the disk. Then by Rouché's theorem, $q - \tilde{q}$ has at most ν zeros in the disk. This contradicts the fact that its Taylor series begins with a term of order at least $\nu + 1$.

The advantage of the C–F problem is that its solution is relatively simple, and in fact it can be computed easily. Following [5, p. 497 ff.], suppose we are given $p(z) = c_0 + \cdots + c_\nu z^\nu$ and extend it to a Blaschke product

$$q(z) = \lambda \frac{\bar{a}_\nu + \cdots + \bar{a}_0 z^\nu}{a_0 + \cdots + a_\nu z^\nu},$$

where now λ is real and positive. Multiplying through formally by the denominator, we obtain in succession

$$\begin{aligned}
\lambda \bar{a}_\nu &= c_0 a_0, \\
\lambda \bar{a}_{\nu-1} &= c_0 a_1 + c_1 a_0, \\
&\vdots \\
\lambda \bar{a}_0 &= c_0 a_\nu + \cdots + c_\nu a_0.
\end{aligned}$$

If each c_k is real, then the solution of these equations reduces to the eigenvalue problem for the $(n+1) \times (n+1)$ Hankel matrix

$$A \equiv \begin{bmatrix} & & & c_0 \\ 0 & & c_0 & c_1 \\ & \cdot^{\cdot^{\cdot}} & & \vdots \\ c_0 & c_1 & \cdots & c_\nu \end{bmatrix}.$$

Again using Rouché's theorem, the desired λ can be shown to be the largest of the absolute values of the eigenvalues of this matrix, and coefficients a_k are given by any corresponding eigenvector $(a_\nu, \ldots, a_0)^T$. (The non-maximum eigenvalues correspond to Blaschke products which have poles inside the unit disk.)

If the coefficients c_k are not necessarily real, it can be shown that the correct λ is now the largest singular value of the same Hankel matrix A, and that any corresponding right singular vector in a singular value decomposition $A = \bar{U} \Sigma U^H$ provides a suitable set of coefficients $\{a_k\}$.

5. The CF Near-Best Approximation Method; A Posteriori Estimates

Given $f(z)$ analytic on the unit disk, we now apply the Carathéodory–Fejér theorem to construct an approximation $p^{cf}(z)$ whose error curve is very close to circular. From the near-circularity it will be possible to show by a posteriori arguments that p^{cf} is close to p^*, and therefore that p^* also has a nearly circular error curve.

We depend upon two transformations that leave any Chebyshev approximation problem on the unit circle essentially unchanged: multiplication of all terms by a fixed power z^N for any integer N (positive or negative), and replacement of z by $1/z$. Combining these two, we see that the problem of extending $c_{n+1}z^{n+1} + c_{n+2}z^{n+2} + \cdots$ backwards to degree 0 with minimal norm on the unit circle is equivalent to the problem of extending $\cdots + c_{n+3}z^{N-(n+3)} + c_{n+2}z^{N-(n+2)} + c_{n+1}z^{N-(n+1)}$, a series extending back towards degree $-\infty$, forwards to degree N with minimal norm (on the circle, not the disk). The latter problem may be solved approximately by truncating the terms of negative degree, applying the Carathéodory–Fejér theorem with $\nu = N-(n+1)$, and then truncating terms of degree greater than N. This is our central computational technique, which we shall call the *Carathéodory–Fejér (CF) method.*

Here is a more precise statement of the method. It is nothing more than a repetition of the C–F theory of the last section with the subscripts changed to reflect the multiplication by z^N and inversion $z \mapsto 1/z$.

(1) Given $n \geqslant 0$ and $f(z) = \sum_{k=n+1}^{\infty} c_k z^k$, pick $N > n$ large enough so that $\sum_{k=N+1}^{\infty} c_k z^k$ may be considered negligible. Construct the $(N-n) \times (N-n)$ Hankel matrix

$$A = \begin{bmatrix} & & & c_N \\ 0 & & \cdot^{\cdot^{\cdot}} & c_{N-1} \\ & \cdot^{\cdot^{\cdot}} & & \vdots \\ c_N & c_{N-1} & \cdots & c_{n+1} \end{bmatrix}.$$

(2) Find the largest singular value λ of A and corresponding right singular vector $\mathbf{a} = (a_1, \ldots, a_{N-n})^T$ in any singular value decomposition of A of the form $A = \bar{U}\Sigma U^H$. If all c_k are real it suffices to take λ as the largest eigenvalue of A in absolute value and $\mathbf{a}$ as a corresponding real eigenvector.

(3) The Blaschke product

$$\sum_{k=-\infty}^{N} c_k z^k = \lambda z^N \frac{\bar{a}_{N-n} + \cdots + \bar{a}_1 z^{N-n-1}}{a_1 + \cdots + a_{N-n} z^{N-n-1}}$$

now represents an extension of f to degree $-\infty$ with winding number at least $n+1$ on $|z|=1$. The CF near-best approximation is defined by

$$p^{cf}(z) = -\sum_{k=0}^{n} c_k z^k.$$

If $a_{N-n} \neq 0$, the coefficients c_k can be computed one by one by the recurrence formula

$$c_k = \frac{-1}{a^{N-n}}(c_{k+1}a_{N-n-1} + \cdots + c_{k+N-n-1}a_1),$$

$k = n, n-1, \ldots, 0$, which is based on matching coefficients of z^k in the formula above.

(4) Further applications of the same recurrence relation give the coefficients $c_{-1}, c_{-2}, \ldots$. Computing a few of these can give an estimate of $\sum_{k=-\infty}^{-1} |c_k|$, hence an approximate bound on how far $f - p^{cf}$ deviates from a circle on $|z|=1$.

Once p^{cf} is computed for a given f and n, $f - p^{cf}$ will typically be nearly circular on $|z|=1$ with winding number $\geqslant n+1$. In this case it follows quickly by Rouché's theorem that the minimum error attained by p^{cf} is a lower bound for $\|f - p^*\|$. This is a direct extension of Theorem 1, and analogous to the bound of de la Vallée Poussin in real Cebyshev approximation:

THEOREM 4. *Let $f(z)$ be analytic in a Jordan region Ω bounded by Γ and continuous on $\bar{\Omega}$, and let $p(z)$ be any polynomial of degree $\leqslant n$. If $f - p$ has winding number $\geqslant n+1$ on Γ, then*

$$\min_{z \in \Gamma} |(f-p)(z)| \leqslant \|f - p^*\| \leqslant \|f - p\|,$$

where p^* is the best approximation to f of degree $\leqslant n$ over Ω.

Proof. The right-hand inequality comes from the definition of p^*, and the left-hand inequality follows from Rouché's theorem as in Theorem 1. ∎

Theorem 4 relates $\|f - p^{cf}\|$ to $\|f - p^*\|$; we would like to complement it by relating p^{cf} to p^* directly. For this a lemma on the image of the unit disk under a polynomial is needed:

LEMMA 5. *Let $p(z) = a_1 z + \cdots + a_\nu z^\nu$, $a_k \in \mathbb{C}$. Then the image of the unit disk under p covers every point of the disk about 0 of radius $2^{-\nu}\|p\|$.*

Proof. Suppose $p(z)$ fails to achieve, say, the value 1 in the unit disk.

Then $1 - p(z)$ can be written $\prod_{k=1}^{\nu}(1 - z/z_k)$ with $|z_k| > 1$ (possibly ∞) for each k. At any point z in the disk, therefore, $-p(z) = \prod_{k=1}^{\nu}(1 + \alpha_k) - 1$, with $|\alpha_k| < 1$. Expanding this product in terms of symmetric polynomials in the α_k, it follows that

$$\|p(z)\| < \binom{\nu}{1} + \binom{\nu}{2} + \cdots + \binom{\nu}{\nu} = 2^{\nu} - 1.$$

The lemma is a consequence of this inequality. ∎

A bound on $\|p^{cf} - p^*\|$ now follows from Lemma 5.

THEOREM 6. *Let f be analytic on the unit disk and let p be a polynomial of degree $\leqslant n$. Suppose $\|f - p - cz^{n+1}\| \leqslant \alpha|c|$ for some $\alpha < 1$ and $c \in \mathbb{C}$, and suppose $\|f - p\| - \min_{|z|=1}|(f - p)(z)| \leqslant \delta$ for some $\delta > 0$. Then*

$$\|p - p^*\| \leqslant 2^{n+1}\delta/\sqrt{1 - \alpha^2},$$

where p^* is the best approximation to f of degree n on the unit disk.

Proof. From the first assumption it is not hard to see that the function $(f(z) - p(z))/cz^{n+1}$ maps the unit disk onto a region in the right half plane all of whose points have complex argument at most $\sin^{-1}\alpha$ in amplitude. If a real number $\eta > 0$ is added to any such point, its modulus must increase by at least $\eta\cos(\sin^{-1}\alpha) = \eta/\sqrt{1 - \alpha^2}$. Now since p^* is the best approximation to f, adding $p - p^*$ to $f - p$ cannot increase the maximum modulus of $f - p$, hence cannot increase the modulus $|(f - p)(z)|$ by more than δ at any point with $|z| = 1$. Equivalently, adding $(p - p^*)/cz^{n+1}$ to $(f - p)/cz^{n+1}$ cannot increase the latter's modulus by more than $\delta/|c|$ at any such point. Putting these facts together, we must have that whenever $(p - p^*)/cz^{n+1}$ is real and positive on $|z| = 1$, it is no greater than $\delta/|c|\sqrt{1 - \alpha^2}$. By Lemma 5, after an inversion $z \mapsto 1/z$, this implies $\|p - p^*\| \leqslant 2^{n+1}\delta/\sqrt{1 - \alpha^2}$. ∎

Theorems 4 and 6 are fully practical estimates, useful for the analysis both of asymptotic error curve behavior (next section) and of particular computed examples (Section 7).

6. NEAR-CIRCULARITY AS $R \to 0$ AND $n \to \infty$

Here we analyze p^{cf} in two asymptotic limits: for fixed degree n on a disk whose radius R shrinks to 0, and for $n \to \infty$ on the fixed unit disk. In each case p^{cf} is assumed to be constructed with $N = 2n + 2$, but the results hold equally if p^{cf} is based on any larger N.

We begin with a key lemma on the dependence of the CF extension (Theorem 3) on the given polynomial coefficients.

LEMMA 7. *Let* $R > 0$ *and* $\nu \geqslant 1$ *be fixed, and let* $p(z) = R^\nu c_0 + R^{\nu-1}c_1 z + \cdots + Rc_{\nu-1}z^{\nu-1} + z^\nu$ *have coefficients satisfying* $|c_k| \leqslant 1$ *for* $0 \leqslant k \leqslant \nu - 1$. *Let p be extended according to the C–F theorem to a power series*

$$q(z) = R^\nu c_0 + \cdots + z^\nu + d_{\nu+1}z^{\nu+1} + d_{\nu+2}z^{\nu+2} + \cdots$$

analytic in the unit disk with minimal norm $\|q\|$. *Then if* $R \leqslant 1/13$, *we have* $|d_{\nu+k}| < 3^\nu(6R)^k$ *for all* $k \geqslant 1$, *and q has winding number* ν *on the unit circle.*

Proof. Let us adopt the following notation:

$$S_R: \text{ circle of radius } R \text{ about the origin,}$$

$$D_R: \text{ closed disk bounded by } S_R,$$

$$\|\phi\|_R: \sup_{z \in D_R} |\phi(z)|,$$

$$S \equiv S_1; \qquad D \equiv D_1; \qquad \| \ \| \equiv \| \ \|_1.$$

Now q has minimal norm of all analytic extensions of p to a power series. Considering the trivial extension $d_{\nu+k} = 0$ $(\forall k \geqslant 1)$ therefore yields the bound $\|q\| \leqslant \|p\|$, hence $\|q - p\| \leqslant 2\|p\| < 2/(1-R)$. Dividing $q - p$ by $z^{\nu+1}$ and applying the maximum modulus principle extends this to $\|q - p\|_{\alpha R} < (\alpha R)^{\nu+1}\, 2/(1-R)$ on the disk $D_{\alpha R}$, where $1 < \alpha \leqslant 1/R$. On $D_{\alpha R}$ we also have $\|p - z^\nu\|_{\alpha R} \leqslant R^\nu + \alpha R^\nu + \cdots + \alpha^{\nu-1}R^\nu < R^\nu \alpha^{\nu-1}/(1 - 1/\alpha) = (\alpha R)^\nu/(\alpha - 1)$. Adding these bounds together gives $\|q - z^\nu\|_{\alpha R} < (\alpha R)^\nu\,[2\alpha R/(1-R) + 1/(\alpha-1)]$. On the other hand $\|z^\nu\|_{\alpha R} = (\alpha R)^\nu$. Thus if $2\alpha R/(1-R) + 1/(\alpha - 1) \leqslant 1$, $\|q - z^\nu\|_{\alpha R} < \|z^\nu\|_{\alpha R}$. The condition $R \leqslant 1/13$ is enough to ensure this with $\alpha = 3$. Applying Rouché's theorem, it follows that $q(z)$ has ν zeros in the disk D_{3R}.

Thus $q(z)$ is a finite Blaschke product

$$q(z) = \lambda\,\frac{\prod_{k=1}^{\nu}(z - \zeta_k)}{\prod_{k=1}^{\nu}(\bar{\zeta}_k z - 1)}$$

with all its zeros inside D_{3R} and poles outside $D_{1/3R}$. This proves the winding number claim. Moreover, we can bound $|d_{\nu+k}|$ by applying Cauchy's estimate using the circle $S_{1/6R}$. We find

$$\begin{aligned} |d_{\nu+k}| &< |\lambda|\,\frac{(1/6R + 3R)^\nu}{(1/2)^\nu} \times \frac{1}{(1/6R)^{\nu+k}} \\ &< \frac{2^\nu}{1-R}(1 + 18R^2)^\nu\,(6R)^k \\ &< 3^\nu(6R)^k. \quad \blacksquare \end{aligned}$$

From Lemma 7 it can now be shown that $f - p^{cf}$ has a nearly circular error curve.

LEMMA 8. *Let $R \leqslant 1/36$ be fixed and positive and suppose $f(z) = z^{n+1} + Rc_{n+2}z^{n+2} + R^2c_{n+3}z^{n+3} + \cdots$ with $|c_k| \leqslant 1$ for $k \geqslant n+2$. Let $p^{cf}(z)$ be the CF approximation to f described in the previous section, with $N = 2n+2$. Then*

$$\|f - p^{cf}\| - \min_{|z|=1} |(f - p^{cf})(z)| < (18R)^{n+2},$$

and $f - p^{cf}$ has winding number $n+1$ on $|z| = 1$.

Proof. The first step of the CF method is to truncate f to form $f^T(z) \equiv z^{n+1} + \cdots + c_{2n+2}R^{n+1}z^{2n+2}$. We have immediately $\|f - f^T\| \leqslant R^{n+2}/(1-R)$. Now applying the C–F theorem by way of the substitution $z \mapsto 1/z$ and multiplication by z^{2n+2}, as described in the last section, leads to a minimal-norm extension of f^T to degree $-\infty$: $\tilde{f}^T(z) = \cdots + c_{-1}z^{-1} + c_0 + \cdots + c_{2n+2}R^{n+1}z^{2n+2}$. p^{cf} is the negative of the portion of $\tilde{f}^T$ from degree 0 to n. From Lemma 7 now follows the bound $|c_{n+1-k}| < 3^{n+1}(6R)^k$ for $k > 0$. Thus $|c_{-1}| < 3^{n+1}(6R)^{n+2}$, $|c_{-2}| < 3^{n+1}(6R)^{n+3}$, etc., and therefore

$$\|\tilde{f}^T - (f^T - p^{cf})\| = \left\| \sum_{k=-\infty}^{-1} c_k z^k \right\| < 3^{n+1} \frac{(6R)^{n+2}}{1-6R} = \frac{(18R)^{n+2}}{3(1-6R)}.$$

Combining the two inequalities yields

$$\begin{aligned}\|\tilde{f}^T - (f - p^{cf})\| &\leqslant \|\tilde{f}^T - (f^T - p^{cf})\| + \|f - f^T\| \\ &< \frac{(18R)^{n+2}}{3(1-6R)} + \frac{R^{n+2}}{1-R}.\end{aligned}$$

Since $n \geqslant 0$ and $R \leqslant 1/36$ it follows readily that

$$\|\tilde{f}^T - (f - p^{cf})\| < \tfrac{1}{2}(18R)^{n+2}.$$

The first assertion now follows from the fact that $\tilde{f}^T$ has constant modulus on $|z| = 1$.

$\|\tilde{f}^T\|$ must satisfy $1 - R/(1-R) \leqslant \|\tilde{f}^T\| = |\tilde{f}^T(z)|$ (for $|z| = 1$), for otherwise it would be possible to approximate z^{n+1} by terms of degree $-\infty$ through n with maximum error less than 1. The winding number statement of Lemma 7 implies that $\tilde{f}^T$ has winding number $-[n+1-(2n+2)] = n+1$. The winding number assertion now follows from these two facts by Rouché's theorem, since we have $\|\tilde{f}^T - (f - p^{cf})\| < \tfrac{1}{2}(18R)^{n+2} < 1 - R/(1-R) \leqslant \min_{|z|=1} |\tilde{f}^T(z)|$. ∎

Thus $f - p^{cf}$ has a nearly circular error curve. Combining this with Theorem 6 will show that $\|p^{cf} - p^*\| = O(R^{n+2})$:

THEOREM 9. *Let f and p be as in Lemma 8, and now assume $R \leqslant 1/36$ and $R \leqslant 1/4\sqrt{n+1}$. Then $\|p^{cf} - p^*\| < (36R)^{n+2}$, where p^* is the best approximation to f of degree $\leqslant n$ on the unit disk.*

Proof. Combining Theorem 6 and Lemma 8, it is clear that if $\|f - p^{cf} - z^{n+1}\| \leqslant \alpha$ for some $\alpha < 1$, then

$$\|p^{cf} - p^*\| < 2^{n+1}(18R)^{n+2}/\sqrt{1-\alpha^2} = \frac{1}{2\sqrt{1-\alpha^2}}(36R)^{n+2}.$$

$1/2\sqrt{1-\alpha^2}$ equals 1 when $\alpha = \sqrt{3}/2$, so it remains to show $\|f - p^{cf} - z^{n+1}\| \leqslant \sqrt{3}/2$.

$\tilde{f}^T$ is the extension of f^T towards degree $-\infty$ of minimal norm. In particular, we must have

$$\begin{aligned}
\|\tilde{f}^T\| &\leqslant \|f^T - \bar{c}_{n+2}Rz^n\| \\
&= \|-\bar{c}_{n+2}Rz^n + z^{n+1} + \cdots + c_{2n+2}R^{n+1}z^{2n+2}\| \\
&= \|-\bar{c}_{n+2}Rz^{-1} + 1 + c_{n+2}Rz + \cdots + c_{2n+2}R^{n+1}z^{n+1}\| \\
&\leqslant \|-\bar{c}_{n+2}Rz^{-1} + 1 + c_{n+2}Rz\| + \sum_{k=3}^{n+2} |c_{n+k}|\, R^{k-1} \\
&\leqslant \|1 + 2iR\,\mathrm{Im}(c_{n+2}z)\| + R^2/(1-R) \\
&\leqslant \sqrt{1+(2R)^2} + R^2/(1-R).
\end{aligned}$$

(This computation follows Saff [14].) For $R \leqslant 1/36$, therefore, $\|\tilde{f}^T\|^2 < 1 + 8R^2$.

Let $\|\cdot\|_2$ be the mean-square norm: $\|\phi\|_2^2 \equiv (1/2\pi)\int |\phi(e^{i\theta})|^2\, d\theta$. Then $\|\tilde{f}^T\|_2^2 \leqslant \|\tilde{f}^T\|^2$, and also $\|\tilde{f}^T\|_2^2$ is the sum of the squares of the magnitudes of the Laurent coefficients of $\tilde{f}^T$. Therefore $\sum_{k=0}^n |c_k|^2 < 8R^2$, from which

$$\left(\sum_{k=0}^{n} |c_k|\right)^2 < (n+1)\, 8R^2$$

follows and hence

$$\|p^{cf}\| < \sqrt{8}\, R\, \sqrt{n+1}.$$

Adding $\|f - z^{n+1}\|$ to this gives, finally,

$$\|f - p^{cf} - z^{n+1}\| \leqslant \sqrt{8}\, R\, \sqrt{n+1} + \frac{R}{1-R} < 3R\sqrt{n+1}$$

and with $R \leqslant 1/4\sqrt{n+1}$ this is less than $\sqrt{3}/2$. ∎

Theorem 9, combined with Lemma 8, implies directly that under appropriate hypotheses p^* interpolates f at exactly $n+1$ points in the domain disk. This must be true, for example, if any function f analytic at 0 is approximated on a disk of radius R sufficiently small. This result was proved a number of years ago by Motzkin and Walsh [11]. Similarly, it must be true on the unit disk for large enough n if $f(z)=\sum_{k=0}^{\infty} c_k z^k$ is entire and satisfies $\lim_{k\to\infty} \sqrt{k}\,|c_{k+1}/c_k|=0$. This result is also known, and due to Saff [14].

Much more, however, follows from Theorem 9 and its predecessors. First, consider approximation on disks whose radii shrink to 0.

THEOREM 10. *Let $n\geqslant 0$ be fixed, let $f(z)$ be analytic at $z=0$ with $f^{(n+1)}(0)/(n+1)!=c_{n+1}\neq 0$, and let p_R^* and p_R^{cf} be the best and CF approximation of degree $\leqslant n$ to f, respectively, on the disk $|z|=R$. Then as $R\to 0$, for all sufficiently small R,*

(i) *$f-p_R^*$ has winding number exactly $n+1$ on $|z|=R$,*

(ii) $\|f-p_R^*\|_R=[1+O(R^{n+2})]\,\|f-p_R^{cf}\|_R=[1+O(R)]\,|c_{n+1}|\,R^{n+1}$,

(iii) $\|p_R^{cf}-p_R^*\|_R=O(R^{n+2})\,\|f-p_R^*\|_R$,

(iv) $\|f-p_R^*\|_R-\min_{|z|=R}|(f-p_R^*)(z)|=O(R^{n+2})\,\|f-p_R^*\|_R$.

Proof. For each radius R, we first rescale the problem to the unit disk by replacing z by Rz, then divide by $c_{n+1}R^{n+1}$ to make the results of this section directly applicable. Clearly $\|f-p_R^*\|=[1+O(R)]\,|c_{n+1}|\,R^{n+1}$. Now (iii) follows from Theorem 9, and the remainder of (ii) follows from (iii). Finally (i) and (iv) follow from (ii), (iii), and Lemma 8. ∎

It is easy to derive a similar theorem for $n\to\infty$ on the fixed unit disk:

THEOREM 11. *Let f be an entire function $\sum_{k=0}^{\infty} c_k z^k$ such that $\sqrt{k}\,|c_{k+1}/c_k|\to 0$ as $k\to\infty$. Let $R>0$ be* arbitrary. *Let p_n^* and p_n^{cf} be the best and CF approximations to f of degree $\leqslant n$, respectively, on $|z|=1$. Then as $n\to\infty$, for all sufficiently large n,*

(i) *$f-p_n^*$ has winding number exactly $n+1$ on $|z|=1$,*

(ii) $\|f-p_n^*\|=[1+O(R^{n+2})]\,\|f-p_n^{cf}\|=[1+O(R)]\,|c_{n+1}|$,

(iii) $\|p_n^{cf}-p_n^*\|=O(R^{n+2})\,\|f-p_n^*\|$,

(iv) $\|f-p_n^*\|-\min_{|z|=1}|(f-p_n^*)(z)|=O(R^{n+2})\,\|f-p_n^*\|$.

Statement (ii) *still holds if the factor $\sqrt{k}$ is removed from the hypothesis.*

Proof. The assumption on the coefficients implies that for all sufficiently large n the conditions of Theorem 9 are satisfied. From here the proof is just as in Theorem 10. For (ii) it suffices to apply Theorem 4 to Lemma 8. ∎

Note that by Cauchy's estimate, the bounds on $\|p^{cf} - p^*\|$ in Theorems 10(iii) and 11(iii) imply the same bounds for each coefficient of $p^{cf} - p^*$ (in the former case, scaled by powers of R).

The estimates of this section have too much slippage to be of much use except for asymptotic arguments. In practice, however, the asymptotic behavior indicated by Theorems 10 and 11 is typically observed in numerical experiments even when R is not very small and n is not very large.

Related results may be found in [1, 14, 11]; the latter also uses the same construction as that used in our proof of Lemma 7. However, the strength of the estimates in these papers is essentially $O(R)$ or $O(R^2)$ rather than $O(R^{n+2})$.

7. Computational Example: e^z on the Unit Disk

We will now consider in detail the approximation of e^z on the unit disk. For this problem the true best approximation error curves are exceedingly close to circles, and the CF method described in Section 5 yields correspondingly excellent results. Compare the bounds given in [1, 14].

For reference, we of course want to know the coefficients of the true best approximations p_n^*. Computing these accurately is not an easy matter. Most of our work has used Lawson's algorithm (see [3, 4, 9]), a procedure which finds the best Chebyshev approximation by solving a sequence of weighted least-squares approximation problems. After step k, the current weight function $w^{(k)}$ is updated according to the formula

$$w^{(k+1)}(\theta) := \frac{w^{(k)}(\theta)\,|(f - p^{(k)})(e^{i\theta})|}{(1/2\pi)\int_0^{2\pi} w^{(k)}(\theta')\,|(f - p^{(k)})(e^{i\theta'})|\,d\theta'}.$$

But from this formula it is apparent that to the extent that the error curve $(f - p^{(k)})(\Gamma)$ is circular, no update takes place at all. Thus we see that although it is suitable for getting close to a best approximation, *for many problems Lawson's algorithm converges asymptotically at an unacceptably slow rate.* Indeed, for approximation of $f(z) = e^z$ with $n = 2$, each additional digit of accuracy would require on the order of 10^6 iterations. This problem is presented in detail in [16, pp. 21–34].

A more promising approach is to formulate the Chebyshev approximation problem as a problem of minimizing a function, namely, $\|f - p\|$, of $n + 1$ complex variables, namely, the coefficients c_k. This problem may be solved numerically by any sufficiently robust multidimensional minimization program, but the convergence will normally be poor because $\|f - p\|$ is necessarily nondifferentiable at the optimal point $\{c_k\}$. It is better to use a method designed for this particular problem (see, e.g., [7]).

640/31/4-5

TABLE I

Properties of Various Approximations to e^z on the Unit Disk for $n = 0, 1, 2$

	(1) TRUE BEST APPROXIMATION				
$n=0$	$p^*(z) = -0.4495742$		$E_{\min} = 1.082$	$E = 1.321$	$E - E_{\min} = 0.239$
$n=1$	$p^*(z) = 0.0149 - 0.1756z$		$E_{\min} = 0.5559$	$E = 0.5584$	$E - E_{\min} = 0.0025$
$n=2$	$p^*(z) = 0.0001813 + 0.0021708z - 0.0432585z^2$		$E_{\min} = 0.177369$	$E = 0.177376$	$E - E_{\min} = 0.7 \times 10^{-5}$
	(2) PARTIAL SUMS OF TAYLOR SERIES				
$n=0$	$p(z) = 0$		$E_{\min} = 0.632$	$E = 1.718$	$E - E_{\min} = 1.086$
$n=1$	$p(z) = 0 + 0z$		$E_{\min} = 0.368$	$E = 0.718$	$E - E_{\min} = 0.350$
$n=2$	$p(z) = 0 + 0z + 0z^2$		$E_{\min} = 0.132$	$E = 0.218$	$E - E_{\min} = 0.0862$
	(3) CF, DROPPING HIGHER-ORDER TERMS IN R				
$n=0$	$p(z) = -0.5$		$E_{\min} = 1.132$	$E = 1.329$	$E - E_{\min} = 0.196$
$n=1$	$p(z) = 0.01389 - 0.16667z$		$E_{\min} = 0.5480$	$E = 0.5655$	$E - E_{\min} = 0.0175$
$n=2$	$p(z) = 0.00017361 + 0.0020833z - 0.041667z^2$		$E_{\min} = 0.1757$	$E = 0.1789$	$E - E_{\min} = 0.0032$
	(4) CF BASED ON TAYLOR COEFFICIENTS $n+1$ THROUGH $2n+2$ $(N = 2n+2)$				
$n=0$	$p(z) = -0.414214$	$[\lambda = 1.2071]$	$E_{\min} = 1.046$	$E = 1.326$	$E - E_{\min} = 0.280$
$n=1$	$p(z) = 0.0189255 - 0.1741693z$	$[\lambda = 0.55702]$	$E_{\min} = 0.5522$	$E = 0.5630$	$E - E_{\min} = 0.0108$
$n=2$	$p(z) = 0.000267511 + 0.00219986z - 0.04325227z^2$	$[\lambda = 0.17737239]$	$E_{\min} = 0.177271$	$E = 0.177497$	$E - E_{\min} = 0.23 \times 10^{-3}$
	(5) CF BASED ON TAYLOR COEFFICIENTS $n+1$ THROUGH ∞ $(N \gg n)$				
$n=0$	$p(z) = -0.54475$	$[\lambda = 1.2584]$	$E_{\min} = 1.173$	$E = 1.344$	$E - E_{\min} = 0.171$
$n=1$	$p(z) = 0.0145209 - 0.1761860z$	$[\lambda = 0.55752907]$	$E_{\min} = 0.5565$	$E = 0.5586$	$E - E_{\min} = 0.0021$
$n=2$	$p(z) = 0.000180862 + 0.0021712208z - 0.04325992458z^2$	$[\lambda = 0.177373815]$	$E_{\min} = 0.1773708$	$E = 0.1773767$	$E - E_{\min} = 0.58 \times 10^{-5}$

Table I lists some properties of five different sets of approximations to e^z on the unit disk. To begin with, set (1) presents actual best approximations for $n = 0, 1, 2$. Inconsistently, throughout this table $p(z)$ represents not an approximation to e^z, but an extension of the tail of the Taylor series for e^z from degree $n+1$ back to degree 0. All the approximations in Table I have error curves with winding number $n+1$, so Theorem 4 is always applicable. The assumptions of Theorem 6 are also satisfied for each approximation with $n = 1$ or $n = 2$ (taking $c = \frac{1}{2}$ and $\frac{1}{6}$, respectively).

Next, set (2) shows the corresponding numbers for approximation by partial sums of the Taylor series.

As a first step using the CF approach, we can compute with pencil and paper an approximation that is much better than the truncated Taylor series, by pretending that the radius $R = 1$ of the unit disk is small and dropping terms of higher order in R. Thus, for example, suppose ($n = 1$) we seek $\{c_k\}$ such that

$$h(z) = \cdots + c_{-1}(zR)^{-1} + c_0 + c_1(zR) + \frac{1}{2!}(zR)^2 + \cdots$$

maps the unit circle to a $1 + 1 = 2$-winding circle. Then we must have, for $|z| = 1$, $h(z)\overline{h(z)} = \lambda^2 \doteq (\frac{1}{2}R^2)^2 + O(R^5)$. Enforcing this condition on the product $h\bar{h}$ termwise, and using the fact that $\bar{z} = 1/z$, we derive, after a little manipulation,

$$c_0 = \frac{1}{72}R^4 + O(R^5),$$

$$c_1 = \frac{-1}{6}R^2 + O(R^3).$$

Dropping higher-order terms in R and setting $R := 1$ leads to the approximations given in Table I in (3). Note that each of these coefficients is 10 times closer to correct than the corresponding coefficient in (2), and that $E - E_{\min}$ for $n = 2$ has been reduced by a factor of 27. Applying Theorem 4, we see that we are already within 2% of a best approximation. If we choose c_n by this method and set $c_0 = c_1 = \cdots = c_{n-1} = 0$, we derive, from Theorem 4, a bound on E^* given by Saff in [14, p. 113].

To apply the CF approach more seriously, we must solve an eigenvalue problem (see Section 5) and thus abandon closed-form solutions. One way to proceed is to follow Section 6, dropping terms in e^z of degree higher than $N = 2n + 2$. This means solving an $(n+2) \times (n+2)$ Hankel eigenvalue problem, which would be a singular value problem if the coefficients of e^z were not real. The results are given as (4) in Table I. $E - E_{\min}$ has dropped by an additional factor of 14 for $n = 2$.

Still, we can do better if we eliminate the arbitrary cutoff after term $2n+2$. Suppose we solve progressively larger and larger eigenvalue problems, taking $N = 2n+3$, $2n+4$, and so on. After five or six such steps the coefficients have clearly converged to the limits given in (5). $E - E_{\min}$ for $n = 2$ has dropped by still another factor of 40, to 0.6×10^{-5}. At this point it follows from Theorem 6 that each coefficient is accurate to within about 0.6×10^{-4} (evidently a conservative bound). In fact, we have now constructed error curves which are slightly more circular than those for the true best approximations: compare the last columns in (1) and (5). For a justification of the existence of this limit Blaschke product, see [2].

Theorem 4 applies to the results in (5), but we can do about twice as well by taking not $E_{\min}$ but λ as a lower bound for E^* here. The justification for this is that λ is the norm of the minimal extension of f back to degree $-\infty$; the extension given by e^* must have norm at least this large. Thus from (5) we have the following tight bounds of the form $\lambda \leqslant E^* \leqslant E$:

n	E^*	
0	$1.30 \pm 0.5 \times 10^{-1}$	(true value: 1.32...)
1	$0.5581 \pm 0.6 \times 10^{-3}$	(true value: 0.5584...)
2	$0.177375 \pm 0.2 \times 10^{-5}$	(true value: 0.177376...)
3	$0.04336898 \pm 0.6 \times 10^{-7}$	
4	$0.0085686585 \pm 0.6 \times 10^{-9}$	
5	$0.001417607269 \pm 0.4 \times 10^{-11}$	
	$\vdots$	

To compute these numbers we have not bothered to search the unit circle for E so that the bound $\lambda \leqslant E^* \leqslant E$ can be applied directly. Almost as tight bounds can be derived by estimating E with the triangle inequality by considering the size of the neglected coefficients $c_{-1}, c_{-2},\ldots$ in the CF technique.

For $n = 3$ the most accurate value for E^* we have computed directly is 0.0433689, which is no more accurate than the bound given above and took a great deal more work. For $N \geqslant 4$ the CF bounds are tighter than values we have computed directly. Indeed, these bounds appear to be somewhat narrower in breadth than the quantities $E^* - E^*_{\min}$ for the true best approximations given in (1). Thus the CF approach yields estimates for the best approximation error E^* which are typically at least as accurate as the corresponding error curve $(f - p^*)(\Gamma)$ is circular. In similar experiments applied to the more troublesome function $1/\Gamma(z)$ (Fig. 2), the CF method yields curves which are considerably more circular in the $E - E_{\min}$ measure than those of the best approximation: for $n = 3$ $(E - E_{\min})/E \approx 0.04$ (CF), $(E^* - E^*_{\min})/E^* \approx 1$. In such a case only E is particularly close to correct, however, not the coefficients of the CF approximation $p(z)$.

8. Generality of the Phenomenon

For Chebyshev approximation on a more general region, and/or approximation by rational functions, much of what we have done still applies. Theorem 1—circular implies best—was already stated in this generality. Theorem 4—nearly circular implies nearly best—is an extension of this result and also generalizes immediately. Theorem 2—circular implies rational—was stated for rational approximation on the disk, but for regions other than the disk we derive readily: if $f-r$ maps a Jordan curve Γ onto a perfect circle, then $f-r$ is the conformal transplant to the region bounded by Γ of a finite Blaschke product. It follows that for approximation by rational functions of type (n, m) on any Jordan region, the space of possible $m+n+1$-winding circular error curves has dimension $m+n+2$. Thus there is good reason to expect best rational approximations on general domains to exhibit nearly circular error curves.

These results show that nearly circular error curves are worth looking for; the rest of our work has been a strategy for finding them. Let us consider approximation by polynomials on a Jordan region Ω with boundary Γ containing the origin. With the aid of a conformal map from Ω to the unit disk it is easy to extend the Carathéodory–Fejér theorem as stated to the region Ω, and in fact (without the winding number claim) to an arbitrary simply connected region not equal to all of $\mathbb{C}$. For the CF approximation technique, however, we need a somewhat different generalization of the C–F theorem, as follows. This might be called a "reverse Carathéodory–Fejér theorem" for any Jordan region. Its proof is an extension of the method used in Section 5 for the unit disk.

THEOREM 12. *Let Ω be a Jordan region containing the origin with boundary Γ. Let $p(z) \in P_N$ have the form $p(z) = c_{n+1}z^{n+1} + \cdots + c_N z^N$. Then there exists a unique extension of $p(z)$ backwards to a Laurent series*

$$q(z) = \cdots + c_{-2}z^{-2} + c_{-1}z^{-1} + \cdots + c_{n+1}z^{n+1} + \cdots + c_N z^N$$

which converges in a neighborhood of $z = \infty$ (except at ∞) to a function which is analytic in the exterior of $\bar{\Omega}$ except at ∞ and continuous on Γ, and which has minimal norm on Γ of all such extensions. $q(z)$ maps Γ to a perfect circle with winding number $\geqslant n+1$ about the origin, and it is the only extension of p in the class described which does this.

Proof. If E is a closed set in the extended plane $\mathbb{C}^*$, let H_E denote the set of functions analytic in the interior of E and continuous at the boundary. Define $\Gamma' \equiv \{z\colon 1/z \in \Gamma\}$, and let Ω' be the Jordan region enclosed by Γ': $\Omega' \equiv \{z\colon 1/z \notin \bar{\Omega}\}$. We are given the problem of Chebyshev approximation of $c_{n+1}z^{n+1} + \cdots + c_N z^N$ over Γ out of $z^n H_{\mathbb{C}^* - \Omega}$.

Step 1: Considering the map $z \mapsto 1/z$, we see that this is equivalent to the problem of approximating $c_N z^{-N} + \cdots + c_{n+1} z^{-n-1}$ over Γ' out of $z^{-n} H_{\bar{\Omega}'}$; i.e., there is a one-to-one correspondence between approximation functions in the two problems which preserves the norm of the error of the approximation.

Step 2: Let ϕ be a conformal map of Ω' onto the unit disk D such that $\phi(0) = 0$. By the Osgood–Carathéodory theorem, ϕ extends to a continuous homeomorphism of $\bar{\Omega}'$ onto $\bar{D}$ with winding number 1. Thus approximating $c_N z^{-N} + \cdots + c_{n+1} z^{-n-1}$ over Γ' out of $z^{-n} H_{\bar{\Omega}'}$ is equivalent in the above sense to approximating $c_N[\phi^{-1}(z)]^{-N} + \cdots + c_{n+1}[\phi^{-1}(z)]^{-n-1}$ over the unit circle S out of $z^{-n} H_{\bar{D}}$. We have used the fact that a pole at the origin in Ω' remains a pole of the same order at the origin in D when transplanted by ϕ^{-1}, and likewise in the other direction.

Step 3: This problem is, in turn, equivalent to the problem of approximating $z^N\{c_N[\phi^{-1}(z)]^{-N} + \cdots + c_{n+1}[\phi^{-1}(z)]^{-n-1}\}$ over S out of $z^{N-n} H_{\bar{D}}$. Now the Carathéodory–Fejér theorem (Theorem 3) applies, yielding a Blaschke product with winding number at most $N - (n+1)$.

Unwinding the three equivalences completes the proof. ∎

Thus the CF method applies in principle on any Jordan region: Given p on Ω, compute CF coefficients $c_n, c_{n-1}, \ldots, c_0, c_{-1}, \ldots$ and then truncate those of negative degree. If the truncated terms are small what remains is a $\geqslant n+1$-winding nearly circular error function. In practice, implementing this strategy requires knowledge of a conformal map from Ω' to the disk.

Theorem 12 may also be extended to cover the limiting case $N \to \infty$, as used in Section 7, Table I, section (5). According to [2], the conclusions of the C–F theorem hold for approximation of any Dini-continuous function f on the unit circle (i.e., $\int_0^\delta |(f(\theta+\eta) + f(\theta-\eta) - 2f(\theta))/\eta|\, d\eta < \infty$ for each θ) by the tail of a power series. For Theorem 12 to hold with $N = \infty$ it is therefore enough to require that p be Dini-continuous (say, differentiable) and that the Jordan curve Γ be analytic, which ensures that Dini continuity is preserved under transplantation to the disk.

The drawback of the CF method on a general region is that the terms $c_n, \ldots, c_{-1}, \ldots$ may not tail off as quickly as for approximation on a disk. Lemma 7 does not extend to a general region, and so the obvious extensions of Theorems 10 and 11 are not valid. All of this is easy to see: since a monomial $c_{n+1} z^{n+1}$ does not, in general, map Γ to a perfect circle, it does no good for f to approach such a monomial through $R \to 0$ or $n \to \infty$. The more irregular Ω is, the more the CF method is likely to face this problem.

In any case, Theorems 10 and 11 are not as strong as we would have liked, and their generalizations, if valid, would also have been incomplete. If f is not entire, and perhaps does not even have a Taylor series converging throughout Ω, we would still like to know what happens as $n \to \infty$. We have not answered this question. Until we do, a full generalization to non-circular regions cannot be expected.

For the case of rational approximation, on the other hand, the analysis is tougher but the near-circularity phenomenon is apparently quite strong. A report on the CF method in rational approximation is in preparation.

Note added in proof. The CF method turns out to have been considered before, notably by G. H. Elliott [18] and M. Hollenhorst [21], though it has not been applied before to the study of error curves. For a summary and further references see [19].

As speculated here in Section 8, both the CF method and the near-circularity phenomenon extend very strongly to rational approximation on the disk [22]. Surprisingly, the method also extends strongly to real approximation on a real interval [19, 20]. Extensions to regions other than a disk are presently being studied by Stephen Ellacott of the Department of Mathematics at Brighton Polytechnic in Great Britain.

ACKNOWLEDGMENTS

I am indebted to Dr. Martin Gutknecht and Professor Peter Henrici of the Eidgenössische Technische Hochschule in Zürich for many valuable suggestions, and to the latter for inviting me to Zürich for the summer during which most of this work was completed. Mr. Cornelio Rusca also spotted an error in the proof of Theorem 6. This work was supported in part by a National Science Foundation Graduate Fellowship.

REFERENCES

1. H.-P. BLATT AND V. KLOTZ, Zur Anzahl der Interpolations-punkte polynomialer Tschebyscheff-Approximationen im Einheitskreis, *in* "Numerische Methoden der Approximationstheorie" (Collatz *et al.*, Eds.), Vol. 4, Birkhäuser, Basel, 1978.
2. L. CARLESON AND S. JACOBS, Best uniform approximation by analytic functions, *Acta Math.* **109** (1963), 219–229.
3. A. CLINE, "Uniform Approximation as a Limit of L_2 Approximations," Ph.D. dissertation, University of Michigan, 1970.
4. S. ELLACOTT AND J. WILLIAMS, Linear Chebyshev approximation in the complex plane using Lawson's algorithm, *Math. Comp.* **30** (1976), 35–44.
5. G. M. GOLUZIN, Geometric theory of functions of a complex variable, *Amer. Math. Soc. Transl.* **26** (1969).
6. U. GRENANDER AND G. SZEGŐ, "Toeplitz Forms and Their Applications," Univ. of California Press, Berkeley/Los Angeles, 1958.
7. M. GUTKNECHT, Ein Abstiegsverfahren fuer nicht-diskrete Tschebyscheff-Approximationsprobleme, *in* "Numerische Methoden der Approximationstheorie" (Collatz *et al.*, Eds.), Vol. 4, Birkhäuser, Basel, 1978.
8. V. KLOTZ, Gewisse rationale Tschebyscheff-Approximationen in der komplexen Ebene, *J. Approx. Theory* **19** (1977), 51–60.
9. C. L. LAWSON, "Contributions to the Theory of Linear Least Maximum Approximations," Ph. D. dissertation, UCLA, 1961.

10. A. L. LEVIN AND V. M. TIHOMIROV, Approximation of analytic functions by rational functions, *Soviet Math. Dokl.* **8** (1967), 622–626.
11. T. S. MOTZKIN AND J. L. WALSH, Zeros of the error function for Tschebycheff approximation in a small region, *Proc. London Math. Soc.* **13** (1963), 90–98.
12. S. J. POREDA, A characterization of badly approximable functions, *Trans. Amer. Math. Soc.* **169** (1972), 249–256.
13. S. J. POREDA, On the convergence of best uniform deviations, *Trans. Amer. Math. Soc.* **179** (1973), 49–59.
14. E. B. SAFF, On the zeros of the error function for Tchebycheff approximation on a disk, *J. Approx. Theory* **9** (1973), 112–117.
15. D. SARASON, "Function Theory on the Unit Circle," Dept. of Math., Virginia Polytechnic Institute and State University, Blacksburg, Va., 1978.
16. L. N. TREFETHEN, "Chebyshev Approximation by Polynomials in the Complex Plane," Undergraduate thesis, Applied Mathematics Committee, Harvard College, May 1977.
17. J. L. WALSH, "Interpolation and Approximation by Rational Functions in the Complex Domain," 2nd ed., Colloquium Publications Vol. 20, Amer. Math. Soc., Providence, R.I., 1956.
18. G. H. ELLIOTT, "The Construction of Chebyshev Approximations in the Complex Plane," Ph.D. dissertation, University of London, 1978.
19. M. H. GUTKNECHT, Rational Carathéodory–Fejér approximation on a disk, a circle, and an interval, in preparation.
20. M. H. GUTKNECHT AND L. N. TREFETHEN, Real polynomial Chebyshev approximation by the Carathéodory–Fejér method, *SIAM J. Numer. Anal.*, in press.
21. M. HOLLENHORST, "Nichtlineare Verfahren bei der Polynomapproximation," Dissertation Universität Erlangen–Nürnberg, 1976.
22. L. N. TREFETHEN, Rational Chebyshev approximation on the unit disk, *Numer. Math.*, in press.