Proceedings of Symposia in Applied Mathematics
Volume 36, 1986

POLYNOMIAL AND RATIONAL APPROXIMATION IN THE COMPLEX DOMAIN

E. B. SAFF[1]

ABSTRACT. Approximation theory in the complex variable setting has its roots in classical function theory, but is rich in modern applications. Moreover, it is a subject that lends much insight into real approximation problems. Starting with the example of Taylor series, we describe methods (such as Faber series and interpolation) for generating good polynomial approximants to a function analytic on a compact set in the plane. We also discuss characterizations for polynomials of best uniform approximation and the "near circularity property." An introduction is given to the theory of Padé approximants, which are rational function analogues of the Taylor sections. We conclude by discussing some contrasts between the theories of polynomial and rational approximation.

1. TAYLOR SECTIONS.

The properties of the Taylor sections for an analytic function are a convenient starting point for approximation and interpolation in the complex z-plane (denoted by $\mathbb{C}$). This is because Taylor sections are least squares polynomial approximants as well as interpolating polynomials. Indeed, if f is analytic at $z = 0$, then the Taylor sections

$$(1.1) \qquad s_n(z) = s_n(f;z) := \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} z^k$$

satisfy the interpolation conditions

$$(1.2) \qquad s_n^{(j)}(0) = f^{(j)}(0), \qquad j = 0,1,\ldots,n.$$

Moreover, the polynomials $1, z, z^2,\ldots$ are orthogonal with respect to the inner product

$$(1.3) \qquad (g,h) := \frac{1}{2\pi r} \int_{C_r} g(z)\overline{h(z)}|dz|, \qquad C_r : |z| = r,$$

and, if f is analytic on $|z| \le r$,

1980 Mathematics Subject Classification 41A10, 41A20, 41A21.
[1]Supported by the National Science Foundation.

$$(1.4)\qquad (f,z^k) = \frac{1}{2\pi r}\int_{C_r} f(z)\bar{z}^k|dz| = \frac{1}{2\pi r}\int_{C_r}\frac{f(z)r^{2k}}{z^k}\frac{rdz}{iz}$$

$$= r^{2k}\frac{1}{2\pi i}\int_{C_r}\frac{f(z)}{z^{k+1}}dz = r^{2k}\frac{f^{(k)}(0)}{k!}.$$

Thus, the least squares (best L^2) polynomial approximation to f out of Π_n on the circle C_r is

$$(1.5)\qquad \sum_{k=0}^{n}\frac{(f,z^k)}{(z^k,z^k)}z^k = \sum_{k=0}^{n}\frac{r^{2k}f^{(k)}(0)}{r^{2k}k!}z^k = s_n(z).$$

Here and below, Π_n denotes the collection of all algebraic polynomials (with complex coefficients) of degree at most n.

Another significant property of Taylor sections is that they provide minimal projections onto Π_n with respect to the <u>sup norm</u>.

<u>Definition</u> 1.1. Let $A(\Delta)$ denote the collection of functions f that are analytic in the open disk $|z| < 1$ and continuous on the closed disk $\Delta : |z| \le 1$. A <u>projection</u> $P : A(\Delta) \to \Pi_n$ is a bounded linear operator such that $P^2 = P$ and $P = I$ on Π_n.

Endowing $A(\Delta)$ and Π_n with the sup norm

$$(1.6)\qquad \|f\| := \sup\{|f(z)|: z \in \Delta\},$$

we expect to find "near best" polynomial approximants to f on Δ by utilizing a projection with smallest possible norm. It was shown by Geddes and Mason [21] that this minimal projection is the <u>Taylor projection</u>:

$$(1.7)\qquad (S_nf)(z) := s_n(f;z) = \sum_{k=0}^{n}\frac{f^{(k)}(0)}{k!}z^k.$$

Namely, they proved

<u>Theorem</u> 1.2. <u>Let</u> P <u>be any projection of the space</u> $A(\Delta)$ <u>onto the subspace</u> Π_n. <u>Then, for the operator norm induced by the sup norm over</u> Δ, <u>we have</u>

$$(1.8)\qquad \|S_n\| \le \|P\|,$$

<u>where</u> S_n <u>is the Taylor projection of</u> (1.7).

The proof of this theorem follows from the clever observation that for any projection P

$$(1.9)\qquad (S_nf)(z) = \frac{1}{2\pi i}\int_{|t|=1}(A_{\bar{t}}PA_tf)(z)\frac{dt}{t},$$

where A_t is the shift operator defined by $(A_tf)(z) := f(tz)$.

What can be said about the rate of convergence of the Taylor sections? The answer is intimately related to the familiar Cauchy-Hadamard formula for the radius of convergence ρ of a power series $\sum_{k=0}^{\infty} c_k z^k$. That is,

$$(1.10) \qquad \frac{1}{\rho} = \limsup_{k\to\infty} |c_k|^{1/k}.$$

The basic convergence result is the following.

Theorem 1.3. *Let f be analytic in an open set that contains the closed unit disk Δ. Then for the sup norm (1.6), the Taylor sections s_n satisfy*

$$(1.11) \qquad \limsup_{n\to\infty} \|f - s_n\|^{1/n} = 1/\rho < 1,$$

where ρ is the radius of the largest open disk centered at the origin throughout which f has a single-valued analytic continuation. Moreover, the sequence s_n converges to f for $|z| < \rho$.

The above theorem, which provides a model for more general results to be mentioned later, nicely illustrates the relationship between the degree of convergence and the maximal circular region of analyticity for f; that is, the larger this circular region, the faster the convergence. In particular, for entire functions f

$$(1.12) \qquad \lim_{n\to\infty} \|f - s_n\|^{1/n} = 0.$$

While the proof of Theorem 1.3 can be deduced via (1.10), it is more instructive to give an argument based on the interpolation property (1.2) of Taylor sections. For this purpose we appeal to the Hermite representation (cf. Walsh [62, §3.1]) for interpolating polynomials.

Lemma 1.4. *Suppose f is analytic inside and on the simple closed contour Γ that surrounds the $n+1$ points $z_0, z_1, \ldots, z_n$. If p is the unique polynomial in Π_n that interpolates f in these points, then*

$$(1.13) \qquad f(z) - p(z) = \frac{1}{2\pi i} \int_\Gamma \frac{w(z)f(t)}{w(t)(t-z)}\, dt, \qquad z \text{ inside } \Gamma,$$

where $w(z) := \prod_{k=0}^{n} (z - z_k)$.

Proof. Replacing $f(z)$ by its Cauchy integral representation

$$f(z) = \frac{1}{2\pi i} \int_\Gamma \frac{f(t)}{t - z}\, dt, \qquad z \text{ inside } \Gamma,$$

equation (1.13) becomes

$$(1.14) \qquad p(z) = \frac{1}{2\pi i}\int_{\Gamma} \frac{w(t) - w(z)}{w(t)(t - z)} f(t)dt, \qquad \text{all } z.$$

From (1.14) we see that p is indeed a polynomial in Π_n and from (1.13) that it interpolates f in the points z_k (the zeros of $w(z)$). □

It is important to keep in mind that (1.13) is valid even when the points z_k are not distinct; in such a case interpolation is meant in the Hermite sense. That is, if z_k is repeated ℓ times, then $p^{(j)}(z_k) = f^{(j)}(z_k)$ for $j=0,1,\ldots,\ell-1$. In particular, since the Taylor section s_n interpolates in the origin of multiplicity $n + 1$, equation (1.13) gives

$$(1.15) \qquad f(z) - s_n(z) = \frac{1}{2\pi i}\int_{|t|=r} \frac{z^{n+1}f(t)}{t^{n+1}(t-z)}\, dt, \qquad |z| < r,$$

for any r such that f is analytic on $|z| \le r$. With the assumptions of Theorem 1.3, we deduce from (1.15) that

$$(1.16) \qquad \limsup_{n\to\infty} \|f - s_n\|^{1/n} \le 1/\rho,$$

and that the sequence s_n converges to f in $|z| < \rho$. If strict inequality holds in (1.16), then

$$\limsup_{n\to\infty} |f^{(n)}(0)/n!|^{1/n} = \limsup_{n\to\infty} \|s_n - s_{n-1}\|^{1/n} < 1/\rho,$$

which implies that the sequence s_n (the Taylor expansion for f) converges to an analytic function in some disk $|z| < R$, with $R > \rho$. As this violates the definition of ρ, the equality of (1.11) follows.

As is often the case where n-th root asymptotics are concerned, results that hold for best L^2 polynomial approximants (such as the Taylor sections) are also valid for best L^p, $1 \le p \le \infty$, approximants. For example, Theorem 1.3 holds if the sections s_n are replaced by the polynomials p_n^* of best uniform approximation to f on Δ.

Many of the elegant properties of Taylor sections can be found in the book of Dienes [14]. We mention only one more fact concerning the behavior of the zeros of Taylor sections for the case when the radius of convergence ρ is finite and positive. Namely, Jentzsch proved [14, p.352] that every point of the circle of convergence $|z| = \rho$ is a limit point of the set of zeros of the sequence $\{s_n\}_0^\infty$. The zeros of the partial sums $s_n(z) = (z^{n+1} - 1)/(z - 1)$ of $f(z) = 1/(1 - z)$ provide a simple illustration of this theorem.

2. POLYNOMIAL APPROXIMATIONS FOR FUNCTIONS ANALYTIC ON E.

Given a compact set E in the z-plane and a function f analytic on E (i.e., f is analytic on an open set $G \supset E$), how do we generate good polynomial approximations to f on E? When E is a closed disk, we can use Taylor sections which are "good" in the sense of Theorem 1.3. For general sets E we need a procedure that likewise reflects the geometry of E.

First, we insist that E does not separate the plane; that is, $\mathbb{C}\backslash E$ is connected. This assumption is necessary if we expect to get uniform convergence (of polynomials) to an arbitrary function analytic on E. For example, the function $f(z) = 1/z$ is analytic on the circle $E : |z| = 1$, but is not the uniform limit on E of any sequence of polynomials because (by the maximum principle) uniform convergence on $|z| = 1$ implies convergence to an analytic function throughout $|z| \leq 1$.

The connectedness of $\mathbb{C}\backslash E$ is also a sufficient condition for polynomial approximation to functions analytic on E as is stated in the following version of the classical Runge's theorem (cf. [62, §1.10]).

Theorem 2.1. *If f is analytic on a compact set E that does not separate the plane, then there exists a sequence of polynomials that converges uniformly to f on E.*

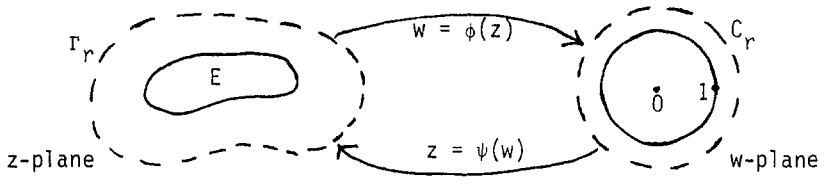
(The question of polynomial approximation to functions not analytic on E is much more delicate and will be addressed in the next section.)

To prove Theorem 2.1, Runge's approach was to first form Riemann sum approximations to the Cauchy integral representation for f. These Riemann sums are rational functions whose poles lie outside E. Through a process of "pole moving," the rational approximants are converted to polynomial approximants.

For reasonable sets E, we can generate polynomial approximants more directly by constructing an analogue of Taylor series. This was the fruitful approach taken by Faber [15]. To simplify the description of Faber's method we assume that E is a compact set (not a single point) whose complement $C^*\backslash E$ with respect to the extended plane is simply connected. The Riemann mapping theorem asserts that there exists a conformal mapping $w = \phi(z)$ of $C^*\backslash E$ onto the exterior of the unit circle in the w-plane (see Figure 2.1). We can insist that $\phi(\infty) = \infty$ and $\phi'(\infty) > 0$ so that, in a neighborhood of infinity,

$$\phi(z) = \frac{z}{c} + b_0 + \frac{b_1}{z} + \frac{b_2}{z^2} + \cdots, \qquad c > 0. \tag{2.1}$$

Figure 2.1



The polynomial basis $\{w^n\}_0^\infty$ for Taylor expansions in the w-plane now corresponds to the functions $\{\phi(z)^n\}_0^\infty$ in the z-plane. The obvious fly in the ointment is that the latter functions are not (in general) polynomials. However, $\phi(z)^n$ does have a polynomial part that will serve our purpose. Indeed, from (2.1), we get

$$(2.2) \qquad \phi(z)^n = \left(\frac{z^n}{c^n} + \cdots\right) + \frac{1}{z} M_n(z)$$
$$= F_n(z) + \frac{1}{z} M_n(z),$$

where $F_n(z) = z^n/c^n + \cdots \in \Pi_n$ and $M_n(z)$ is analytic at infinity. We call F_n the n-th degree <u>Faber polynomial for the set</u> E, but caution the reader that many authors reserve this terminology for its <u>monic</u> brother $c^n F_n(z)$.

For a function f analytic on E, our goal is to obtain an expansion of the form

$$(2.3) \qquad f(z) = a_0F_0(z) + a_1F_1(z) + a_2F_2(z) + \cdots .$$

For this purpose, it is convenient to introduce the inverse mapping of ϕ, denoted by $z = \psi(w)$, and the level curves

$$(2.4) \qquad \Gamma_r : |\phi(z)| = r \qquad (r > 1)$$

which are images under ψ of the circles $C_r : |w| = r$ (see Figure 2.1). Since F_n is the principal part of the Laurent series (2.2) for $\phi(z)^n$, we can write

$$F_n(z) = \frac{1}{2\pi i} \int_{\Gamma_r} \frac{\phi(t)^n}{t - z}\, dt \qquad (z \text{ inside } \Gamma_r),$$

and transforming to the w-plane we obtain

$$(2.5) \qquad F_n(z) = \frac{1}{2\pi i} \int_{C_r} \frac{s^n}{\psi(s) - z}\, \psi'(s) ds.$$

To derive the expansion (2.3) we begin with the Cauchy integral representation for f(z):

$$(2.6)\qquad f(z) = \frac{1}{2\pi i}\int_{\Gamma_r}\frac{f(t)}{t-z}\,dt = \frac{1}{2\pi i}\int_{C_r}\frac{f(\psi(s))}{\psi(s)-z}\,\psi'(s)ds.$$

Since $f(\psi(s))$ is analytic in an annulus of the form $1 < |s| < R$, we can expand this function in a Laurent series:

$$f(\psi(s)) = \sum_{n=-\infty}^{\infty} a_n s^n.$$

Substituting this series into (2.6) and recalling (2.5) we get

$$(2.7)\qquad f(z) = \sum_{n=-\infty}^{\infty}\frac{a_n}{2\pi i}\int_{C_r}\frac{s^n\psi'(s)ds}{\psi(s)-z} = \sum_{n=0}^{\infty}\frac{a_n}{2\pi i}\int_{C_r}\frac{s^n\psi'(s)ds}{\psi(s)-z} = \sum_{n=0}^{\infty} a_n F_n(z)$$

(the integrals with negative n vanish because the integrand is $O(1/s^2)$ near ∞).

To summarize, we obtain the Faber expansion for f by forming the Taylor series for the Cauchy integral of the composition $f\circ\psi$ and substituting F_n for w^n. The process is diagrammed below.

$$f(z)\longrightarrow(f\circ\psi)(w)\longrightarrow\frac{1}{2\pi i}\int_{C_r}\frac{(f\circ\psi)(s)}{s-w}\,ds = \sum_{0}^{\infty} a_n w^n\longrightarrow\sum_{0}^{\infty} a_n F_n(z)$$

Exploiting the relationship between Taylor and Faber series leads to the following analogue of Theorem 1.3.

<u>Theorem</u> 2.2. <u>Let</u> f <u>be analytic on</u> E <u>and let</u> $\rho(>1)$ <u>be the largest index such that</u> f <u>has a single-valued analytic continuation throughout the interior of the level curve</u> Γ_ρ. <u>Then the partial sums of the Faber series for</u> f <u>satisfy</u>

$$(2.8)\qquad \limsup_{n\to\infty}\left\|f - \sum_{k=0}^{n} a_k F_k\right\|_E^{1/n} = 1/\rho < 1,$$

<u>where</u> $\|\cdot\|_E$ <u>denotes the sup norm on</u> E. <u>Moreover, the Faber series converges to</u> f <u>throughout the interior of</u> Γ_ρ.

What does Theorem 2.2 say to realists <u>who</u> do approximation on an interval? If $E = [-1,1]$, then $\phi(z) = z+\sqrt{z^2-1}$ is just the <u>Joukowski transformation</u> with inverse

$$(2.9)\qquad \psi(w) = \frac{1}{2}(w + w^{-1}).$$

For $n \geq 1$, the polynomial part of $\phi(z)^n$ is the same as the polynomial part of

$$\phi(z)^n + \phi(z)^{-n} = w^n + w^{-n}$$

which reduces to $2\cos n\theta$ when $w = e^{i\theta}$. Thus the Faber polynomials are (apart from a multiplicative constant) the same as the classical Chebyshev polynomials T_n, and the Faber series reduces to the (orthogonal) Chebyshev expansion! For the Joukowski transformation, the level curve Γ_r is an ellipse with foci ± 1 and semi-major axis of length $(r + r^{-1})/2$. Hence Theorem 2.2 asserts that the Chebyshev expansion for a function f analytic on $[-1,1]$ will converge to f throughout the largest ellipse with foci ± 1 in which f is analytic.

A more in-depth discussion of Faber series and Faber transforms is given in [13], [17], [22], [49], and [2]. The reader will find the subject rich in applications to geometric function theory.

Polynomial approximants can also be constructed via interpolation. As we observed, the Taylor section $s_n(f,z)$ of (1.1) interpolates f in the origin or, more precisely, in the zeros of the polynomial $w_n(z) = z^{n+1}$. Since the mapping function ϕ for the disk $\Delta_c : |z| \le c$ is just $\phi(z) = z/c$, the $w_n(z)$ trivially satisfy

$$\lim_{n\to\infty} |w_n(z)|^{1/n} = c|\phi(z)|. \tag{2.10}$$

In fact, this asymptotic relation, when used to estimate the integral in the Hermite interpolation formula (1.15), is all that is needed to prove the convergence assertions of Theorem 1.3. For more general compact sets E (with $\mathbb{C}^*\backslash E$ simply connected) this suggests that we determine a triangular scheme of points for E

$$\begin{array}{l} \beta_0^{(0)} \\ \beta_0^{(1)}, \quad \beta_1^{(1)} \\ \cdots\cdots\cdots\cdots \\ \beta_0^{(n)}, \quad \beta_1^{(n)}, \ldots, \beta_n^{(n)} \\ \cdots\cdots\cdots\cdots\cdots\cdots \end{array} \tag{2.11}$$

such that $w_n(z) := \prod_{k=0}^{n} (z - \beta_k^{(n)})$ satisfies (2.10) uniformly on compact subsets of $C\backslash E$, where $\phi(z) = z/c + \cdots$ is now the mapping function of (2.1). Coupled with the Hermite interpolation formula this leads to the following result (cf. [62, §7.2]).

<u>Theorem</u> 2.3. <u>If the scheme of points (2.11) of</u> E <u>satisfies (2.10) uniformly on compact subsets of</u> $\mathbb{C}\backslash E$, <u>then the assertions of Theorem 2.2 remain valid when the Faber sections are replaced by the sequence of polynomials</u> p_n <u>that interpolate</u> f <u>in the successive rows of (2.11).</u>

For the unit disk Δ: $|z| \le 1$, the zeros of $w_n(z) = z^{n+1} - 1$ (the roots of unity) provide "good points" of interpolation in the sense of (2.10).

When E is bounded by a smooth Jordan arc or curve, we obtain good points by taking the images under $z = \psi(w)$ of such equally spaced points on $|w| = 1$. For example, if $E = [-1,1]$, the images of the roots of $w^n + i = 0$ under the transformation (2.9) yield the zeros of the Chebyshev polynomial T_n.

There are good points of interpolation that can be determined without knowledge of the mapping function. These are the Fekete points (cf. [62,§7.8]).

Definition 2.4. Let $V_n(z_0,z_1,\ldots,z_n) := \prod_{i<j} (z_i - z_j)$ denote the Vandermonde determinant of order $n + 1$. The points $\beta_k^{(n)} = z_k \in E$ for which the maximum

$$\max \{|V_n(z_0,z_1,\ldots,z_n)|;\ z_k \in E,\ k=0,\ldots,n\}$$

is attained are called Fekete points for E.

The positive constant c that appears in the expansion (2.1) for the mapping function has great importance; it is called the transfinite diameter or logarithmic capacity of E and is denoted by $\mathrm{cap}(E)$. Such terminology arises from an electrostatics problem that we now describe.

For a compact set E (with $\mathbb{C}^*\backslash E$ simply connected) we distribute a unit charge over its boundary ∂E so that equilibrium is reached in the sense that the energy with respect to the logarithmic potential is minimized. This corresponds to the problem of finding the minimum of the energy integral

$$(2.12) \qquad I[\mu] := \int_{\partial E} \int_{\partial E} \log|\xi_1 - \xi_2|^{-1} d\mu(\xi_1) d\mu(\xi_2)$$

over all positive unit measures μ supported on ∂E. The unique measure μ_E that minimizes $I[\mu]$ gives the equilibrium charge distribution with potential

$$(2.13) \qquad U_E(z) := \int_{\partial E} \log|z - t|^{-1} d\mu_E(t).$$

Apart from a small exceptional set, this potential has the constant value $I[\mu_E]$ on the boundary of E. The capacity of E is defined as

$$(2.14) \qquad \mathrm{cap}(E) := \exp(-I[\mu_E]).$$

In this context, the essential criterion for (2.11) to be "good points" of interpolation is that the discrete measures

$$(2.15) \qquad \mu_n := \frac{1}{n+1} \sum_{k=0}^{n} \delta(\beta_k^{(n)}),$$

where $\delta(\beta_k^{(n)})$ denotes the unit measure supported at $\beta_k^{(n)}$, converge to the equilibrium measure μ_E (in the weak-star topology). Such convergence implies

that for $z \in \mathbb{C}\backslash E$,

$$(2.16) \qquad \int_{\partial E} \log|z - t|^{-1} d\mu_n(t) \longrightarrow U_E(z) \quad \text{as } n\to\infty,$$

which is equivalent to property (2.10). In this light, the fact that the Fekete points are good interpolation points seems reasonable since they are defined by minimizing the energy $\log(V_n^{-1})$ for $n + 1$ distinct point charges.

The above discussion applies to more general sets E and these aspects of potential theory can be found in Hille [27, §16.4], Tsuji [57], and Landkof [28]. We mention one further characterization of good interpolation points (cf. [62, §7.4]).

<u>Theorem</u> 2.5. <u>The points</u> $\beta_k^{(n)}$ <u>of</u> E <u>satisfy (2.10) if and only if</u>

$$\lim_{n\to\infty} \|w_n\|_E^{1/n} = \operatorname{cap}(E).$$

3. <u>POLYNOMIALS OF BEST UNIFORM APPROXIMATION</u>.

Let E be a compact set in the z-plane and f a function continuous on E. Since Π_n is finite dimensional, there exists a polynomial $p_n^* \in \Pi_n$ of best uniform approximation to f on E in the sense that

$$(3.1) \qquad \|f - p_n^*\|_E = \inf\{\|f - q\|_E : q \in \Pi_n\},$$

where $\| \ \|_E$ is the sup norm over E. Moreover, if E contains at least $n + 1$ points, then Π_n is a Chebyshev subspace and hence p_n^* is unique (see §2 of DeVore's notes). A fundamental characterization of best approximation in the complex variable setting is the <u>Kolmogoroff criterion</u>:

<u>Theorem</u> 3.1. <u>A polynomial</u> $p \in \Pi_n$ <u>is a best uniform approximation to</u> f <u>on</u> E <u>if and only if</u>

$$(3.2) \qquad \min_{z\in M} \operatorname{Re}\{\overline{(f(z) - p(z))}\, q(z)\} \le 0$$

<u>holds for every</u> $q \in \Pi_n$, <u>where</u> M <u>is the set of extremal points for</u> $f(z) - p(z)$; <u>that is</u>,

$$(3.3) \qquad M := \{z \in E : |f(z) - p(z)| = \|f - p\|_E\}.$$

A proof of Theorem 3.1 is given in Meinardus [34, p.15].

For real functions, condition (3.2) asserts that there is <u>no</u> polynomial in Π_n that has the same sign as the error $f - p_n^*$ on its extremal point

set. This is the essential fact that is used to prove the Chebyshev Equioscillation theorem. For complex functions, an analogue of the alternating-sign patterns was developed by Rivlin and Shapiro (cf. [48, §2.6]) and is called the extremal signature.

Let's turn to the geometric aspects of best approximation. We let $A(E)$ denote the collection of functions f that are analytic in the interior of E and continuous on E. If $f \in A(E)$ and E is bounded by a Jordan curve Γ, then best polynomial approximation to f on E reduces to best approximation on Γ; that is, by the maximum principle,

$$\|f - p\|_\Gamma = \|f - p\|_E, \quad \text{all} \quad p \in \Pi_n.$$

The image of Γ under $f - p$ is a curve in the w-plane which we denote by $(f - p)(\Gamma)$ and call an error curve. In this context, the problem of best uniform approximation to f is equivalent to finding an error curve that is contained in a disk of minimal radius about $w = 0$.

It had been observed by some authors and crystallized by Trefethen [53] that the minimal error curve $(f - p_n^*)(\Gamma)$ often has a near circularity property in the sense that it winds around the origin $n + 1$ times and is close to being a perfect circle. Before proceeding with a discussion of this phenomenon we give a consequence of perfect circularity.

Lemma 3.2. *Suppose E is bounded by a Jordan curve Γ, $f \in A(E)$, and $p \in \Pi_n$. If the error curve $(f - p)(\Gamma)$ is a perfect circle with center at the origin and winding number $\geq n + 1$, then p is the polynomial of best uniform approximation to f on E out of Π_n.*

Proof. If, to the contrary, there exists $q \in \Pi_n$ such that $\|f - q\|_E < \|f - p\|_E$, then

$$|(f - p)(z) - (q - p)(z)| = |(f - q)(z)| < \|f - p\|_E = |(f - p)(z)|$$

for all z on Γ. By Rouché's theorem, this means that $q - p$ and $f - p$ have the same number of zeros interior to Γ. But since this number is at least $n + 1$ and $q - p \in \Pi_n$, we arrive at a contradiction.□

As a simple application of Lemma 3.2, consider the problem of finding the polynomial in Π_n that is of best uniform approximation to $f(z) = z^{n+1}$ on $\Delta : |z| \leq 1$. Since f itself has the perfect circularity property, then $p_n^* \equiv 0$. In other words, the Chebyshev polynomials for the disk Δ are just the powers of z.

Using finite Blaschke products we can produce other examples of per-

fectly circular error curves, but only for certain rational functions f. (The reader is invited to determine the polynomials of best approximation on Δ to $f(z) = 1/(z - \alpha)$, $|\alpha| > 1$.)

While near circularity is a property that can be made precise in an asymptotic sense (cf. Trefethen [53]), its practical importance is in the construction of very accurate polynomial approximations. The starting point for this algorithm is an elegant theorem due to Carathéodory and Fejér (cf. [22, p. 497]).

Theorem 3.3. *Given a polynomial* $p(z) = \sum_{k=0}^{\nu} c_k z^k$, *there exists a unique power series extension* $B(z) = p(z) + \sum_{k=\nu+1}^{\infty} c_k^* z^k$ *analytic in the unit disk* Δ *that minimizes* $\|B\|_\Delta$ *among all such extensions. Moreover,* $B(z)$ *is a finite Blaschke product with at most* ν *zeros in the disk.*

The solution $B(z)$ to the minimal extension problem of Theorem 3.3 can be computed quite easily. We know that it has the form

$$(3.4) \qquad B(z) = \lambda \frac{\bar{b}_\nu + \cdots + \bar{b}_0 z^\nu}{b_0 + \cdots + b_\nu z^\nu}, \qquad \lambda > 0,$$

and that it extends $p(z)$ in the sense that $B^{(k)}(0)/k! = c_k$ for $k = 0,\ldots,\nu$. When the c_k's are real, this system reduces to an eigenvalue problem for a $(\nu + 1) \times (\nu + 1)$ Hankel matrix formed from the c_k's. It turns out that the constant λ (which equals $\|B\|_\Delta$) in (3.4) is the largest of the absolute values of the eigenvalues of this matrix and that the coefficients b_k are determined by a corresponding eigenvector. For complex coefficients c_k, the procedure is modified by working instead with the largest singular value of the same Hankel matrix.

How does the CF extremal problem of Theorem 3.3 relate to the problem of best polynomial approximation on Δ? Finding the minimal error $f - p_n^*$ for $f(z) = \sum_0^\infty a_k z^k$ is equivalent to minimizing

$$(3.5) \qquad \Big\|\sum_{k=0}^{n} c_k z^k + \sum_{k=n+1}^{\infty} a_k z^k\Big\|_C, \qquad C : |z| = 1,$$

over all $(n + 1)$-tuples $(c_0,\ldots,c_n)$, which is the *converse* of the CF problem. Nonetheless, we can utilize Theorem 3.3 by performing a truncation and an inversion $z \to 1/z$. Following Trefethen [53], we truncate the given series for f at $k = N$ so that $\sum_{N+1}^{\infty} a_k z^k$ is negligible; that is, we work with $\sum_{n+1}^{N} a_k z^k =: z^{n+1} q(z)$ instead of $\sum_{n+1}^{\infty} a_k z^k$ in (3.5).

Next we solve the CF problem for the inverse polynomial

$$(3.6) \qquad p(z) := z^{N-n-1}q(1/z) \in \Pi_{N-n-1},$$

to obtain the minimal extension Blaschke product

$$B(z) = p(z) + \sum_{k=N-n}^{\infty} c_k^* z^k.$$

Since

$$(3.7) \qquad \|B\|_C = \|z^N B(1/z)\|_C = \|z^{n+1}q(z) + \sum_{k=0}^{n} c_{N-k}^* z^k + \sum_{k=N+1}^{\infty} c_k^* z^{N-k}\|_C,$$

then discarding the terms involving negative powers of z (which have small coefficients), we see that the choice

$$c_k = c_{N-k}^*, \qquad k = 0,1, \ldots ,n,$$

in (3.5) gives an error curve with a near circularity property.

The polynomial approximants obtained via this CF method are often much better <u>in the sup norm sense</u> than the Taylor sections. Moreover the technique can be extended to find near best <u>rational</u> approximants (cf. [54], [56]). The theoretical underpinnings of the CF method are contained in a paper of Adamjan, Arov, and Krein [1] who generalized the results of Carathéodory, Fejér, Schur, and Takagi.

Let's now turn to the question of <u>convergence</u> of approximating polynomials. We naturally ask, what is the extension of the <u>Weierstrass theorem</u> to the complex setting? Runge's theorem (Theorem 2.1) is not a true generalization because it assumes far more than continuity - it requires f to be analytic in an open set containing E. Only in 1951 did the Russian mathematician Mergelyan confirm the suspicions of many who had worked on the problem by proving that the assumption on f in Runge's theorem could be weakened.

<u>Theorem</u> 3.4 (Mergelyan [35]). <u>Let</u> E <u>be a compact set that does not separate the plane</u>. <u>If</u> $f \in A(E)$ (<u>that is,</u> f <u>is analytic in the interior of</u> E <u>and continuous on</u> E), <u>then there exists a sequence of polynomials that converges uniformly to</u> f <u>on</u> E.

The proof of Mergelyan's theorem (cf. [17], [41]) is a *tour de force* that utilizes the Tietze Extension theorem as well as Koebe's 1/4-theorem. Observe that the Weierstrass theorem is a special case of Theorem 3.4 because an interval has an empty interior and so $A(E)$ reduces to the collection of functions continuous on E.

As an application of Theorem 3.4 we mention the following

generalization of the Cauchy integral formula: If Γ is a rectifiable Jordan curve and f is analytic in the interior G of Γ and continuous on $G \cup \Gamma$, then

$$(3.8) \qquad f(z) = \frac{1}{2\pi i} \int_\Gamma \frac{f(t)}{t - z}\, dt, \qquad z \in G.$$

To prove (3.8) we take a sequence of polynomials p_n that converges uniformly to f on $G \cup \Gamma$ (the special case of Mergelyan's theorem used here was proved in 1926 by Walsh [61]). Since the Cauchy integral representation holds for polynomials, we have for $z \in G$

$$f(z) = \lim_{n\to\infty} p_n(z) = \lim_{n\to\infty} \frac{1}{2\pi i} \int_\Gamma \frac{p_n(t)}{t - z}\, dt = \frac{1}{2\pi i} \int_\Gamma \frac{f(t)}{t - z}\, dt,$$

as claimed in (3.8).

Results on the <u>rate</u> of polynomial convergence require special assumptions on the smoothness of the boundary E as well as on the modulus of continuity of f. For some extensions of the Jackson type theorems, see Sewell [47].

As the reader might suspect from the results of §2, <u>geometric rates</u> of convergence characterize the functions that are analytic <u>on</u> E. Before making this precise we present a useful lemma dealing with the growth of polynomials.

<u>Lemma</u> 3.5 (Bernstein-Walsh [62, §4.6]). <u>Suppose that</u> E <u>is a compact set (not a single point), whose complement</u> $C^*\backslash E$ <u>is simply connected. If</u> $p \in \Pi_n$ <u>satisfies</u> $|p(z)| \le M$ <u>for</u> z <u>on</u> E, <u>then</u>

$$(3.9) \qquad |p(z)| \le Mr^n, \qquad z \text{ on } \Gamma_r \ (r > 1),$$

<u>where</u> Γ_r <u>is the level curve defined in (2.4).</u>

<u>Proof</u>. We apply the maximum principle to $g(z) := p(z)/\phi(z)^n$, where $\phi(z)$ is the mapping function of (2.1). Observe that since $p \in \Pi_n$ and ϕ has a simple pole at ∞, then $g(z)$ is analytic exterior to E, even at ∞. As z approaches the boundary of E from the outside, $|g(z)| \le M$; hence $|g(z)| \le M$ for all z outside E. For $z \in \Gamma_r$, the last inequality gives (3.9). □

We can now prove

<u>Theorem</u> 3.6 (Walsh [62, §4.7]). <u>Let</u> E <u>be as in Lemma 3.5,</u> f <u>a function continuous on</u> E, <u>and set</u>

$$(3.10) \qquad E_n(f) := \|f - p_n^*\|_E,$$

where p_n^* is the polynomial in Π_n of best uniform approximation to f on E. Then f is analytic on E if and only if

$$(3.11) \qquad \limsup_{n\to\infty} E_n(f)^{1/n} < 1.$$

Proof. In one direction the proof is trivial. Namely, if f is analytic on E, then Theorem 2.2 asserts that the Faber sections and, *a fortiori*, the polynomials of best approximation converge geometrically.

On the other hand, if (3.11) holds, then

$$(3.12) \qquad \limsup_{n\to\infty} \|p_{n+1}^* - p_n^*\|_E^{1/n} < 1.$$

Appealing to Lemma 3.5, we deduce that, for some $r > 1$,

$$\limsup_{n\to\infty} \|p_{n+1}^* - p_n^*\|_{\Gamma_r}^{1/n} < 1.$$

But this means that the sequence $\{p_n^*\}_0^\infty$ converges in the interior of Γ_r, necessarily to an analytic extension of f. □

As with several of the theorems presented, Theorem 3.6 is not stated in its full generality - the assumption on E can be considerably weakened.

4. PADÉ APPROXIMANTS.

Polynomials have the advantage of being easy to evaluate. But the same is true of rational functions. Moreover, rational functions have poles which can imitate the singularities of a function to be approximated. In this section we introduce a class of interpolating rational functions called Padé approximants. These rationals provide a natural extension of the Taylor sections. (Standard references are [39], [4], [5,6]; for a historical treatment, see Brezinski [11].)

Given a formal power series

$$(4.1) \qquad f(z) = \sum_{k=0}^{\infty} a_k z^k,$$

we wish to construct a rational function of a certain type whose Taylor coefficients match those of f as far as possible. To be precise, let

$$(4.2) \qquad \Pi_{m,n} := \{R(z) = P(z)/Q(z) : P \in \Pi_m,\ Q \in \Pi_n,\ Q \not\equiv 0\}.$$

Then the matching condition can be stated as follows: For a fixed pair (m,n), find an $R \in \Pi_{m,n}$ such that

$$(4.3) \qquad (f - R)(z) = O(z^\ell),$$

where ℓ is as large as possible. (Here and below, $O(z^{\ell})$ denotes a power series with lowest order term z^{ℓ}.) What is a realistic value for ℓ? Since there are $m+1$ free parameters in the choice for the numerator P, and $n+1$ in the choice for the denominator Q, there are $m+n+1$ parameters available in the ratio P/Q (one parameter is lost in the division process). Thus we expect to have $\ell \geq m+n+1$ or, equivalently, to match the first $m+n+1$ terms of (4.1). Unfortunately this is not always possible (try $m=0$, $n=1$, and $f(z) = z$). To circumvent this difficulty we work, instead, with the following linearized version of (4.3).

Given (m,n), select $P_{mn} \in \Pi_m$ and $Q_{mn}(\not\equiv 0) \in \Pi_n$ so that

$$(4.4) \qquad (Q_{mn}f - P_{mn})(z) = O(z^{m+n+1}).$$

If f is $(m+n)$-times differentiable at $z=0$, then (4.4) is equivalent to

$$(Q_{mn}f - P_{mn})^{(k)}(0) = 0, \qquad k=0,1,\ldots,m+n.$$

Notice that (4.4) represents a homogeneous system of $m+n+1$ equations in $m+n+2$ unknowns (the coefficients of P_{mn} and Q_{mn}). Hence this system has a nontrivial solution, necessarily with $Q_{mn} \not\equiv 0$. With this observation we give

<u>Definition</u> 4.1. The <u>Padé approximant</u> (PA) of type (m,n) to f is the rational

$$(4.5) \qquad [m/n](z) := P_{mn}(z)/Q_{mn}(z),$$

where $P_{mn} \in \Pi_m$ and $Q_{mn}(\not\equiv 0) \in \Pi_n$ satisfy (4.4).

Notice that for $n=0$, the PA reduces to a Taylor section of (4.1):

$$(4.6) \qquad [m/0](z) = \sum_{k=0}^{m} a_k z^k.$$

Tacit in Definition 4.1 is the fact that a PA is <u>unique</u>. To prove this, suppose that

$$(4.7) \qquad (Q_1 f - P_1)(z) = O(z^{m+n+1}) \quad \text{and} \quad (Q_2 f - P_2)(z) = O(z^{m+n+1}),$$

where $P_1, P_2 \in \Pi_m$ and $Q_1, Q_2 \in \Pi_n$. On multiplying the first equation in (4.7) by Q_2 and the second by Q_1, we deduce on subtracting that

$$(4.8) \qquad Q_1P_2 - Q_2P_1 = O(z^{m+n+1}).$$

But the left-hand side of (4.8) is a polynomial of degree $\leq m+n$. Hence $Q_1P_2 - Q_2P_1 \equiv 0$ or $P_1/Q_1 \equiv P_2/Q_2$.

The Padé numerators and denominators are rich in algebraic properties such as the 3-term recurrence relations found by Frobenius (see [4, 24, 39] for a detailed discussion of these properties). Here we pause only to mention a representation for Q_{mn} that illustrates the important role played by the <u>Toeplitz determinants</u>

$$(4.9)\qquad D(m/n) := \begin{vmatrix} a_m & a_{m+1} & \cdots & a_{m+n-1} \\ a_{m-1} & a_m & \cdots & a_{m+n-2} \\ \vdots & \vdots & & \vdots \\ a_{m-n+1} & a_{m-n+2} & \cdots & a_m \end{vmatrix} \qquad (a_k := 0 \text{ if } k < 0)$$

formed from the coefficients of f.

<u>Theorem</u> 4.2. (Jacobi). <u>If</u> $D(m/n) \neq 0$, <u>then</u>

$$(4.10)\qquad f(z) - [m/n](z) = O(z^{m+n+1})$$

<u>and the Padé denominator</u> Q_{mn} <u>normalized by</u> $Q_{mn}(0) = 1$ <u>is</u>

$$(4.11)\qquad Q_{mn}(z) = \frac{1}{D(m/n)} \begin{vmatrix} a_m & a_{m+1} & \cdots & a_{m+n} \\ a_{m-1} & a_m & \cdots & a_{m+n-1} \\ \vdots & \vdots & & \vdots \\ a_{m-n+1} & a_{m-n+2} & \cdots & a_{m+1} \\ z^n & z^{n-1} & \cdots & 1 \end{vmatrix} .$$

A fast numerical method (based on the Euclidean algorithm) for solving Toeplitz systems and computing PAs is described in [10].

The PAs for (4.1) are typically displayed in a doubly infinite array known as the <u>Padé table</u>:

$$\begin{array}{|cccc} [0/0] & [1/0] & [2/0] & \cdots \\ [0/1] & [1/1] & [2/1] & \cdots \\ [0/2] & [1/2] & [2/2] & \cdots \\ \vdots & \vdots & \vdots & \end{array}$$

Here the first row lists Taylor sections; the 2nd row consists of PAs with at most one pole; the 3rd row consists of PAs with at most two poles; etc. The structure of this table was the subject of the 1892 thesis of E. Padé. He showed that the table breaks up into square blocks of identical entries, with the common entry not appearing elsewhere in the table. When all blocks are of size one, i.e., no entry is repeated, the table is said to be normal. Normal tables arise when all Toeplitz determinants $D(m/n)$ are nonzero. It is possible for a Padé table to contain an infinite block of identical entries, but (as shown by Kronecker) such a table arises only for the power series of a rational function.

Of special interest in the Padé table are the diagonal entries, for these represent continued fraction expansions. Indeed, if

$$(4.12) \qquad f(z) = \sum_0^\infty a_k z^k = d_0 + \cfrac{d_1 z}{1 + \cfrac{d_2 z}{1 + \ddots}},$$

then an inductive argument shows that the successive truncations d_0, $d_0 + d_1 z$, $d_0 + d_1 z/(1 + d_2 z)$, etc. are rational functions that have maximal contact with f at the origin. In other words, these truncations give the PAs

$$[0/0]\ ,\ [1/0]\ ,\ [1/1]\ ,\ldots,\ [n/n]\ ,\ [(n+1)/n]\ ,\ldots,$$

which form a staircase of entries in the Padé table (the main diagonal and first superdiagonal). For many of the classical special functions (such as e^z), the $[n/n]$ approximant in continued fraction form provides an accurate, computationally stable approximation that is considerably better than using the $2n$-th degree Taylor section. Of course, continued fraction expansions of real numbers have played an important role in number theory and, in this respect, PAs provide their function theoretic analogues. (For further discussion of the continued fraction aspects of PAs, see [39], [59]).

The PAs for a function f have poles that can be used to predict the positions of the poles as well as other singularities of f. For example, the qd-algorithm (cf. [26, §7.6]) for computing the zeros of a given polynomial p is based on the fact that the poles in certain rows of the Padé table for $f = 1/p$ tend to poles of f (zeros of p). The basic row convergence theorem involved is the following.

Theorem 4.3. (de Montessus de Ballore [4, p.139]). *Let f be analytic in the disk $D : |z| < R$ $(0 < R \leq \infty)$ except for poles of total multiplicity ν, none of which occurs at $z=0$. Then, as $m \to \infty$, the sequence of Padé approximants $[m/\nu](z)$ converges to $f(z)$ uniformly on every compact subset of*

$D\setminus\{$ poles of $f\}$. Furthermore, as $m\to\infty$, the poles of $[m/\nu](z)$ tend, respectively, to the ν poles of f in D.

For example, suppose that f is a meromorphic function in the plane whose poles are simple and occur at the points ξ_k, where

$$0 < |\xi_1| < |\xi_2| < \cdots .$$

Then Theorem 4.3 asserts that the poles of $[m/1](z)$ tend to ξ_1; the two poles of $[m/2](z)$ tend to ξ_1, ξ_2; etc.

The proof of Theorem 4.3 is based on the following simple observation (cf. [46]). Since

$$(Q_{m\nu}f - P_{m\nu})(z) = O(z^{m+\nu+1}),$$

then for any $Q \in \Pi_\nu$, the product $QP_{m\nu} \in \Pi_{m+\nu}$ satisfies

$$(Q_{m\nu}Qf - QP_{m\nu})(z) = O(z^{m+\nu+1}),$$

and so $QP_{m\nu}$ is the $(m+\nu)$-th Taylor section of $Q_{m\nu}Qf$. Consequently, we can use the Hermite formula (1.15) to write

$$(4.13)\qquad (Q_{m\nu}Qf - QP_{m\nu})(z) = \frac{1}{2\pi i}\int_{|t|=r} \frac{z^{m+\nu+1}(Q_{m\nu}Qf)(t)}{t^{m+\nu+1}(t-z)}\,dt, \qquad |z| < r,$$

provided $Q_{m\nu}Qf$ is analytic on $|t| \le r$. If Q is chosen to be the monic polynomial whose zeros are the poles of f, then r can be taken arbitrarily close to R. On suitably normalizing the Padé denominators $Q_{m\nu}$ we find that the right-hand side of (4.13) tends to zero in D. In particular, at a zero ξ of Q, we have $(Q_{m\nu}Qf)(\xi) \to 0$ and so $Q_{m\nu}(\xi) \to 0$ because $(Qf)(\xi) \neq 0$. This means that every limit polynomial of the $Q_{m\nu}$'s has zeros at the poles of f (the zeros of Q), which establishes the last assertion of Theorem 4.3. (This same argument can be applied to rational functions that interpolate in the "good points" discussed in §2; see [43].)

In proving convergence theorems for PAs, the essential question is: Where (asymptotically) are the poles of the PAs? In Theorem 4.3, the ν poles of f serve as "attractors" for all the available poles of the $[m/\nu]$ approximants. However, if f has fewer than ν poles, then only a subset of the poles of $[m/\nu](z)$ "know where to go," and the remaining poles may wander aimlessly, destroying convergence. The following simple example illustrates this point.

Consider a sequence of nonzero coefficients a_m for which there is a large discrepancy between the root test and the ratio test:

(4.14) $\lim_{m\to\infty} |a_m|^{1/m} = 0$ and $\limsup_{m\to\infty} |a_{m+1}/a_m| = \infty$.

As shown by Perron [39, §78], it is possible to construct such a_m's so that the sequence $\{a_m/a_{m+1}\}_0^\infty$ has limit points that are dense in the plane. But, from (4.11), we see that a_m/a_{m+1} is the zero of the Padé denominator $Q_{m1}(z)$ for $f(z) = \sum_0^\infty a_k z^k$ (which is an entire function). Hence the 2nd row of the Padé table for this f has poles everywhere dense in the plane.

Even more startling is the following result due to Wallin [60] concerning the diagonal of the Padé table.

Theorem 4.4. *There exists an entire function* f *such that the sequence of diagonal PAs* $\{[n/n](z)\}_0^\infty$ *for* f *is unbounded at every point in the plane except* $z = 0$.

In light of these anomalies, results on the convergence of PAs usually pursue one of three directions:

(i) Proving uniform convergence for *special classes* of functions;

(ii) Replacing uniform convergence by a weaker condition, such as *convergence in measure or in capacity*;

(iii) Extracting *subsequences* of PAs that do have the desired uniform convergence properties.

An early step in the first direction was taken by Padé, who studied the table for the exponential function. He showed that whenever $m+n\to\infty$, the approximants $[m/n](z)$ for e^z converge to e^z uniformly on compact subsets of the plane. Precise asymptotic results for the location of the zeros and poles of these PAs were obtained by Saff and Varga [45]. The approximants for the exponential have several important applications. For example, proving the stability of certain numerical schemes for solving differential equations boils down to showing that certain of these approximants are bounded by 1 in the left-half plane.

A substantial extension of Padé's results for the exponential function was obtained by Arms and Edrei [3]. They proved convergence of the approximants for the class of functions generated by totally positive sequences (also called *Pólya frequency series*).

The PAs for the class of *Stieltjes functions* have particularly elegant properties. Here we discuss Stieltjes functions that can be written in the form

$$(4.15) \qquad f(z) = \int_0^b \frac{d\mu(t)}{1 + zt},$$

where μ is a finite positive measure on $[0,b]$, with $0 < b < \infty$. Such a function is analytic in the cut plane $C^*\backslash(-\infty, -1/b]$ and has the power series

expansion $f(z) = \sum_0^\infty (-1)^k c_k z^k$, where the c_k's are the moments

$$(4.16) \qquad c_k := \int_0^b t^k d\mu(t), \qquad k=0,1,\ldots .$$

As we now show, the Padé denominators $Q_{n-1,n}$ for f are related to the polynomials that are orthogonal with respect to $d\mu$. Starting with the defining property

$$(Q_{n-1,n}f - P_{n-1,n})(z) = O(z^{2n}),$$

we replace z by $-1/z$ and multiply by z^n to obtain

$$(4.17) \qquad q_n(z) \int_0^b \frac{z d\mu(t)}{z - t} - z p_{n-1}(z) = O(1/z^n),$$

where $q_n(z) := z^n Q_{n-1,n}(-1/z) \in \Pi_n$ and $p_{n-1}(z) := z^{n-1} P_{n-1,n}(-1/z) \in \Pi_{n-1}$. Then for $j=0,1,\ldots,$ we have

$$(4.18) \qquad q_n(z) \int_0^b \frac{z^j}{z - t} d\mu(t) - z^j p_{n-1}(z) = O(z^j/z^{n+1}).$$

Next, we integrate with respect to z around a simple closed contour containing $[0,b]$ in its interior. Using the Cauchy formula, we find that

$$\int_0^b q_n(t) t^j d\mu(t) = 0, \quad \text{for } j=0,1,\ldots, n-1;$$

that is,

$$(4.19) \qquad q_n(z) = z^n Q_{n-1,n}(-1/z)$$

<u>is the n-th degree orthogonal polynomial for</u> $d\mu$. One consequence of this relation is that the zeros of $Q_{n-1,n}(z)$ must be simple and lie on the cut $(-\infty, -1/b)$. On writing the approximant $[(n-1)/n]$ in the form

$$(4.20) \qquad [(n-1)/n](z) = \frac{P_{n-1,n}(z)}{Q_{n-1,n}(z)} = \sum_{j=1}^n \frac{A_{nj}}{1 + z t_{nj}},$$

where the t_{nj}'s are zeros of $q_n(t)$, we deduce in a similar manner from (4.17) that

$$(4.21) \qquad \int_0^b P(t) d\mu(t) = \sum_{j=1}^n A_{nj} P(t_{nj})$$

for any polynomial $P \in \Pi_{2n-1}$. Hence, the constants A_{nj} are the Christoffel

numbers for Gaussian quadrature (cf. [52]). Since Christoffel numbers are positive, we see from (4.20) that the approximant $[(n-1)/n](z)$ is itself a Stieltjes function of the form (4.15) corresponding to a discrete measure $d\mu_n$. In particular, the zeros and poles of this approximant are interlaced along the cut $(-\infty, -1/b)$. With all these facts in hand, a simple normal families argument can be used to prove that $[(n-1)/n](z) \to f(z)$ in $\mathbb{C}^* \setminus (-\infty, -1/b]$. This is a classical result due to Markoff [33], which has been further extended by Stahl [50].

The example of Stieltjes functions shows that the Padé theory provides a natural setting for generalizing the classical theory of orthogonal polynomials. In this regard, convergence results for PAs to functions of the form (4.15) with μ a complex measure, were obtained by Magnus [32], Nuttall and Wherry [38], and Stahl [51].

Many commonly occurring functions have "smooth" Taylor coefficients in the sense that $a_{k-1}a_{k+1}/a_k^2$ has a limit as $k \to \infty$. Convergence properties of the PAs for such functions were investigated by Lubinsky [29], [30].

Space limitations preclude a discussion of results concerning the convergence of subsequences of the rows, columns, or diagonals of the Padé table. We also leave it for the reader to delve into results (such as the Nuttall-Pommerenke theorem [5, §6.5]) that deal with the convergence in capacity of near diagonal PAs.

We do wish to emphasize that various generalizations of PAs exist that are quite useful; e.g. multipoint Padé approximants (rational functions found by interpolation in distinct points), Faber-Padé approximants (rational functions whose Faber series matches the Faber expansion of f as far as possible), multivariate Padé approximants; etc. (A dictionary [31] of these generalizations is available, upon request, from this author.)

5. RATIONAL VERSUS POLYNOMIAL APPROXIMATION (WHAT A DIFFERENCE A DIVISION MAKES!)

We now discuss some essential differences between polynomial and rational approximation in the complex variable setting. Some of the contrasts are rooted in function theoretic properties, while others are more typical of linear vs. nonlinear approximation theory (cf. the forthcoming book of Braess [9]).

Possibility of Convergence. For rational approximation, Runge's classical theorem asserts that if f is analytic on a compact set E, then f is the uniform limit on E of a sequence of rational functions. Unlike its polynomial version (Theorem 2.1), the hypothesis that $\mathbb{C}\setminus E$ be connected is not needed; it is compensated for by choosing rational approximants that have poles in the components of $\mathbb{C}\setminus E$. For example, a function analytic on the annulus

$E : r_1 \le |z| \le r_2$ is the uniform limit of rational functions that have poles at $z = 0$ and $z = \infty$ (think of its Laurent series!).

To describe the more delicate problem of approximating functions in $A(E)$, we let $\overline{\Pi(E)}$ denote the uniform limits on E of polynomials, and $\overline{R(E)}$ denote the uniform limits on E of rational functions whose poles lie outside E. Then the theorem of Mergelyan (Theorem 3.4) states that $A(E) = \overline{\Pi(E)}$ if and only if $C\backslash E$ is connected. In contrast, the compact sets E for which $A(E) = \overline{R(E)}$ cannot be characterized topologically; that is, this property is not invariant under a homeomorphism of the plane (cf. [20]). The most popular (and most tasteful) example of a compact set E for which $A(E) \ne \overline{R(E)}$ is the Swiss cheese of A. Roth (cf. [17]), which she manufactured by removing a countable number of disjoint open disks from the closed unit disk. For further discussion of the possibility of rational approximation see Gamelin [18].

Existence of Best Approximants. For an arbitrary compact set E, the existence of best polynomial approximants from Π_m is a simple compactness argument. However, for best rational approximants from $\Pi_{m,n}$ $(n > 0)$, this argument must be modified to handle the possibility of poles tending to the boundary of E. Using normal families, Walsh [62, §12.2] proved that best rational approximants exist provided E contains no isolated points.

Uniqueness of Best Approximants. If $f \in C[a,b]$ is real-valued, then Chebyshev showed that the best uniform approximation to f on $[a,b]$ out of

$$\Pi^r_{m,n} := \{R \in \Pi_{m,n} : R \text{ has real coefficients}\} \tag{5.1}$$

is unique (cf. [34, §9.2]). Surprisingly, this is no longer true if approximation to a real-valued f is done from $\Pi_{m,n}$; that is, if we allow rational approximants with complex coefficients. Indeed, as was shown by Saff and Varga [44], the function $f(x) = x^2$ has no unique best uniform approximation on $[-1,1]$ out of $\Pi_{1,1}$ (any such best rational r_{11} has complex coefficients, so that $\bar{r}_{11}(\ne r_{11})$ is also best). Further examples of this type, as well as non-uniqueness results for approximation on a disk can be found in [25], [42].

Given $f \in A(E)$ we can nonetheless construct a table of best uniform rational approximants to f on E by making a specific choice for each pair (m,n). This analogue of the Padé table is called the Walsh array. The convergence theory for this array closely parallels the theory for the Padé table (e.g. Walsh [64] proved an analogue of Theorem 4.3). Moreover, the Padé table can be viewed as a limiting version of Walsh arrays where best approximation is done on disks $E_\varepsilon : |z| \le \varepsilon$ with $\varepsilon \to 0$ (cf. [55], [63]).

Degree of Convergence of Best Approximants. For $f \in A(E)$, we set

$$E_n(f) := \inf\{\|f - p\|_E : p \in \Pi_n\}, \qquad e_n(f) := \inf\{\|f - R\|_E : R \in \Pi_{n,n}\}.$$

Clearly, $e_n(f) \le E_n(f)$ for all n, and so the essential question is: Can $e_n(f)$ tend to zero substantially faster than $E_n(f)$? (Let's rule out the trivial situation where f itself is rational.) The now famous example of Newman [36] answered this affirmatively for $f(x) = |x|$ on $E : [-1,1]$, where $E_n(f) \approx 1/n$, while $e_n(f) \approx e^{-\pi\sqrt{n}}$ (cf. [12], [58]). Another example of the contrast is readily accessible to the reader. Using a simple calculus argument, one shows that for the partial sums $s_n(x) := \sum_0^n x^k/k!$ of e^x, there holds for the sup norm on $[0,\infty)$,

$$\limsup_{n\to\infty} \|e^{-x} - 1/s_n(x)\|_{[0,\infty)}^{1/n} \le 1/2.$$

Replacing x by $(1+x)/(1-x)$ we see that for $f(x) := \exp[-(1+x)/(1-x)]$ and $E : [-1,1]$,

$$\limsup_{n\to\infty} e_n(f)^{1/n} \le 1/2.$$

On the other hand, Theorem 3.6 asserts that

$$\limsup_{n\to\infty} E_n(f)^{1/n} = 1$$

because f is not analytic at $x = 1$.

At present, there is no simple characterization of the functions f for which $e_n(f) \ll E_n(f)$. However, some special classes of functions have been investigated in this direction. For example, Gončar [23] has obtained the precise geometric rate of convergence for rational approximation on an interval to Stieltjes functions of the form (4.15). Several important results for classes of real functions were obtained by Popov [40], Freud [16], and others. See also the survey articles of Ganelius [19] and Newman [37] on the subject.

Analytic Continuation. A simple but important observation concerning the class $\Pi_{n,n}$ of rational functions, is that it is invariant under a bilinear transformation. Thus, unlike polynomials, rational functions can provide analytic continuations of functions to unbounded regions of the plane. The convergence properties of the diagonal Padé approximants to Stieltjes functions illustrates this point. Another example of the contrast is for Newman's example. It can be shown that the sequence of polynomials $\{p_n^*\}_0^\infty$ of best uniform approximation to $f(x) = |x|$ on $[-1,1]$ diverges on every continuum in $\mathbb{C}\setminus[-1,1]$; moreover (analogous to the Jentzsch theorem of §1), every point of $[-1,1]$ is a limit point of zeros of the p_n^* (cf. [7]). On the other hand, the best (real) rational approximants

R_n^* to $|x|$ out of $\Pi_{n,n}$ have all their zeros and poles on the imaginary axis and satisfy (cf. [8])

$$\lim_{n\to\infty} R_n^*(z) = \begin{cases} z & \text{for } \operatorname{Re} z > 0 \\ -z & \text{for } \operatorname{Re} z < 0. \end{cases}$$

REFERENCES

1. V. M. Adamjan, D. Z. Arov, and M. G. Krein, "Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem", Math. USSR Sbornik, 15 (1971), 31-73.

2. J. M. Anderson, "The Faber operator", In: Rational Approximation and Interpolation (P.R. Graves-Morris, E. B. Saff, and R. S. Varga, eds.), Lecture Notes in Math., Vol. 1105, Springer-Verlag, Berlin (1984), 1-10.

3. R. J. Arms and A. Edrei, "The Padé tables and continued fractions generated by totally positive sequences", In: Mathematical Essays Dedicated to A. J. Macintyre, Ohio University Press, Athens, Ohio (1970), 1-21.

4. G. A. Baker, Jr., Essentials of Padé Approximants, Academic Press, New York (1975).

5. G. A. Baker, Jr. and P. R. Graves-Morris, Padé Approximants Part I: Basic Theory, Encyl. of Math., Vol. 13, Cambridge Univ. Press, Cambridge (1981).

6. G. A. Baker, Jr. and P. R. Graves-Morris, Padé Approximants Part II: Extensions and Applications, Encyl. of Math., Vol. 14, Cambridge Univ. Press, Cambridge (1981).

7. H.-P. Blatt and E. B. Saff, "Behavior of zeros of polynomials of near best approximation", J. Approx. Theory, 46 (1986).

8. H.-P. Blatt, A. Iserles, and E. B. Saff, "Remarks on the behavior of zeros of best approximating polynomials and rational functions", (to appear).

9. D. Braess, Nonlinear Approximation, Springer-Verlag, Berlin, (to appear).

10. R. P. Brent, F. G. Gustavson, and D. Y. Yun, "Fast solution of Toeplitz systems of equations and computation of Padé approximants", Journal of Algorithms, 1 (1980), 259-295.

11. C. Brezinski, "The long history of continued fractions and Padé approximants", In: Padé Approximations and Applications (M. G. de Bruin, H. van Rossum, eds.), Lecture Notes in Math., Vol. 888, Springer-Verlag, Berlin (1981), 1-27.

12. A. P. Bulanov, "The asymptotics of the maximum deviation of $|x|$ from rational functions", Mat. Sb. 76 (1968), 288-303.

13. J. H. Curtiss, "Faber polynomials and the Faber series", Amer. Math. Monthly, 78 (1971), 577-596.

14. P. Dienes, The Taylor Series, Dover, New York (1957).

15. G. Faber, "Über polynomische Entwicklungen", Math. Ann. 57 (1903), 398-408.

16. G. Freud, "Über die Approximation reller Funktionen durch rationale gebrochene Funktionen", Acta Math. Acad. Sci. Hungar, 17 (1966), 313-324.

17. D. Gaier, Vorlesungen über Approximation im Komplexen, Birkhäuser Verlag, Basel (1980).

18. T. W. Gamelin, Uniform Algebras, Prentice-Hall, Englewood Cliffs, N.J. (1969).

19. T. Ganelius, W. K. Hayman and D. J. Newman, Lectures on Approximation and Value Distribution, Seminaire de Mathematiques Superieures, Les Presses de l'Universite de Montreal, Montreal, Canada (1982).

20. P. M. Gauthier, "On the possibility of rational approximation", In: Padé and Rational Approximation (E. B. Saff and R. S. Varga, eds.), Academic Press, New York (1977), 261-264.

21. K. O. Geddes and J. C. Mason, "Polynomial approximation by projections on the unit circle", SIAM J. Numer. Anal., 12 (1975), 111-120.

22. G. M. Golusin, Geometric Theory of Functions of a Complex Variable, Amer. Math. Soc., Vol. 26, Providence, R.I. (1969).

23. A. A. Gončar, "On the speed of rational approximation of some analytic functions", Math. USSR Sbornik, 34 (1978), 131-145.

24. W. B. Gragg, "The Padé table and its relation to certain algorithms of numerical analysis", SIAM Rev., 14 (1972), 1-62.

25. M. H. Gutknecht and L. N. Trefethen, Nonuniqueness of best rational Chebyshev approximations on the unit disk," J. Approx. Theory, 39 (1983), 275-288.

26. P. Henrici, Applied and Computational Complex Analysis, Vol. I, John Wiley & Sons, New York (1974).

27. E. Hille, Analytic Function Theory (Introduction to Higher Mathematics, vol. II), Ginn and Co., Boston (1962).

28. N. S. Landkof, Foundations of Modern Potential Theory, Springer-Verlag, Berlin (1972).

29. D. S. Lubinsky, "Padé tables of entire functions of very slow and smooth growth", Constr. Approx. 1 (1985), 349-358.

30. D. S. Lubinsky, "Uniform convergence of rows of the Padé table for functions with smooth Maclaurin series coefficients", (to appear).

31. D. S. Lubinsky and E. B. Saff, "A dictionary of generalized Padé approximants", Institute for Constr. Math. Technical Report (1986), Univ. of South Fla.

32. A. P. Magnus, "Another theorem of convergence of complex weight Padé approximants", Padé Meeting, Luminy, France, 14-18 Oct. 1985.

33. A. Markoff, "Deux demonstrations de la convergence de certains fractions continues", Act. Math., 19 (1895), 93-104.

34. G. Meinardus, Approximation of Functions: Theory and Numerical Methods, Springer-Verlag, Berlin (1967).

35. S. N. Mergelyan, "On the representation of functions by series of polynomials on closed sets", (Russian) Dokl. Akad. Nauk SSSR, 78, 405-408; Translations Amer. Math. Soc. No. 85 (1953).

36. D. J. Newman, "Rational approximation to $|x|$ ", Michigan Math. J., 11 (1964), 11-14.

37. D. J. Newman, Approximation with Rational Functions, Regional Conference Series in Math, Vol. 41, Amer. Math. Soc., Providence, R.I. (1979).

38. J. Nuttall and C. J. Wherry, "Gaussian integration for complex weight Padé functions", J. Inst. Math. Appl., 21 (1978), 165-170.

39. O. Perron, Die Lehre von den Kettenbrüchen, Chelsea Pub. Co., New York (1929).

40. V. A. Popov, "Uniform rational approximation of the class V and its applications", Acta Math. Acad. Sci. Hungar, 29 (1977), 119-129.

41. W. Rudin, Real and Complex Analysis, McGraw-Hill, New York (1974).

42. A. Ruttan, "On the cardinality of a set of best complex rational approximations to a real function", In: Padé and Rational Approximations (E. B. Saff and R. S. Varga, eds.), Academic Press, New York (1977), 303-319.

43. E. B. Saff, "An extension of Montessus de Ballore's theorem on the convergence of interpolating rational functions", Journ. Approx. Theory, 6 (1972), 63-67.

44. E. B. Saff and R. S. Varga, "Nonuniqueness of best complex rational approximation to real functions on real intervals", J. Approx. Theory, 23 (1978), 78-85.

45. E. B. Saff and R. S. Varga, "On the zeros and poles of Padé approximants to e^z , III, Numer. Math., 30 (1978), 241-266.

46. E. B. Saff, "An introduction to the convergence theory of Padé approximants", In: Aspects of Contemporary Complex Analysis (D. A. Brannan, J. G. Clunie, eds.), Academic Press, New York (1980), 493-502.

47. W. E. Sewell, Degree of Approximation by Polynomials in the Complex Domain, Ann. of Math. Studies No. 9, Princeton Univ. Press, Princeton, N.J. (1942).

48. H. S. Shapiro, Topics in Approximation Theory, Lecture Notes in Math., Vol. 187, Springer-Verlag, Berlin (1971).

49. V. I. Smirnov and N. A. Lebedev, Functions of a Complex Variable: Constructive Theory, Iliffe Books Ltd., London (1968).

50. H. Stahl, Beitrage zum Problem der Konvergenz von Padé-approximierenden, Dissertation, Technischen Universitat Berlin (1976).

51. H. Stahl, "Orthogonal polynomials with complex valued weight function", I & II, Constr. Approx. (to appear).

52. G. Szegö, Orthogonal Polynomials, 3rd ed., Amer. Math. Soc. Colloq. Pub., Vol. 23, Amer. Math. Soc., Providence, R.I. (1967).

53. L. N. Trefethen, "Near-circularity of the error curve in complex Chebyshev approximation", J. Approx. Theory, 31 (1981), 344-367.

54. L. N. Trefethen, "Rational Chebyshev approximation on the unit disk", Numer. Math., 37 (1981), 297-320.

55. L. N. Trefethen and M. H. Gutknecht, "On convergence and degeneracy in rational Padé and Chebyshev approximation," SIAM J. Math. Anal., 16 (1985), 198-210.

56. L. N. Trefethen and M. H. Gutknecht, "The Carathéodory-Fejér method for real rational approximation, SIAM J. Numer. Anal., 20 (1983), 420-436.

57. M. Tsuji, Potential Theory in Modern Function Theory, Dover, New York (1959).

58. N. S. Vjaceslavov, "On uniform approximation of $|x|$ by rational functions", Soviet Math. Dokl.,16 (1975), 100-104.

59. H. S. Wall, Analytic Theory of Continued Fractions, Van Nostrand, Princeton, N.J. (1948).

60. H. Wallin, "On the convergence theory of Padé approximants", In: Linear Operators and Approximation, ISNM Vol. 20, Birkhäuser, Basel (1972), 461-469.

61. J. L. Walsh, "Über die Entwicklung einer analytischen Funktion nach Polynomen", Math. Ann., 96 (1926). 430-436.

62. J. L. Walsh, Interpolation and Approximation by Rational Functions in the Complex Domain. 3rd ed., Amer. Math. Soc. Colloq. Publ., Vol. 20, Amer. Math. Soc., Providence, R.I. (1960).

63. J. L. Walsh, "Padé approximants as limits of rational functions of best approximation", J. Math. Mech., 13 (1964), 305-312.

64. J. L. Walsh, "The convergence of sequences of rational functions of best approximation with some free poles", In: Proc. Sympos. Approximation of Functions (General Motors Res. Lab., 1964), Elsevier, Amsterdam (1965), 1-16.

E. B. Saff
Institute for Constructive Mathematics
Department of Mathematics
University of South Florida
Tampa, Florida 33620