Numer. Math. 25, 307—322 (1976)

Geometric Convergence to e^{-z} by Rational Functions with Real Poles

E. B. Saff*, A. Schönhage, and R. S. Varga**

Received July 18, 1975

Summary. In this paper, we show that there exists a sequence of rational functions of the form $R_n(z)=p_{n-1}(z)/(1+z/n)^n$, $n=1, 2, \ldots$, with $\deg p_{n-1} \leqq n-1$, which converges geometrically to e^{-z} in the uniform norm on $[0, +\infty)$, as well as on some infinite sector symmetric about the positive real axis. We also discuss the usefulness of such rational functions in approximating the solutions of heat-conduction type problems.

1. Introduction

Because of its applications to the study of certain numerical methods used in computing solutions of systems of differential equations, a number of recent papers ([1, 2, 6, 8], among others) have been devoted to the topic of rational approximations to e^{-z} on bounded and unbounded sets in the complex plane $\mathbb{C}$. Perhaps the simplest of all such rational approximants is the sequence

$$s_n(z) := \frac{1}{\left(1+\frac{z}{n}\right)^n}, \qquad n=1, 2, \ldots \tag{1.1}$$

which enjoys the following readily verified properties:

(i) the sequence converges uniformly to e^{-z} on every bounded subset of the plane;

(ii) the sequence converges uniformly to e^{-x} on the infinite real interval $[0, +\infty)$; and

(iii) the poles of the sequence are all real and negative.

Of course, all of the above advantages are outweighed by the fact that the sequence (1.1) converges too slowly to e^{-z} for use in any actual computations. It is therefore desirable to seek other rational approximations $\{R_n(z)\}_{n=1}^{\infty}$ which possess properties (i), (ii), and (iii), and furthermore, say, converge geometrically to e^{-x} in the uniform norm on $[0, +\infty)$, i.e.,

$$\overline{\lim_{n\to\infty}} \|e^{-x} - R_n(x)\|_{[0, +\infty)}^{1/n} < 1.$$

* Research supported in part by the Air Force Office of Scientific Research under Grant AFOSR-74-2688, and by the University of South Florida Research Council.

** Research supported in part by the Air Force Office of Scientific Research under Grant AFOSR-74-2729, and by the Energy Research and Development Administration (ERDA) under Grant E(11-1)-2075.

In the present paper, we prove that there exists a sequence of rational functions $\{R_n(z)\}_{n=2}^{\infty}$ of the form

$$R_n(z) = \frac{p_{n-1}(z)}{\left(1+\frac{z}{n}\right)^n}, \qquad \text{with } \deg p_{n-1} \leqq n-1, \tag{1.2}$$

such that $\{R_n(z)\}_{n=2}^{\infty}$ converges geometrically (in the uniform norm) to e^{-z} on every bounded subset of $\mathbb{C}$, as well as on the ray $[0, +\infty)$. Furthermore, we show that no sequence of reciprocals of polynomials $\{1/q_n(z)\}_{n=1}^{\infty}$, $\deg q_n \leqq n$, with the $q_n(z)$ having only negative real roots, can converge geometrically to e^{-x} on $[0, +\infty)$. These and related results are stated and discussed in Sec. 2, with their proofs being given in Sec. 3. Then, in Sec. 4 we comment on the usefulness of rational functions, of the form (1.2), in approximating the solutions of heat-conduction type problems, and we also give numerical results estimating the geometric convergence rate of such rational functions to e^{-x} on $[0, +\infty)$.

2. Statements of New Results

We now introduce some notation and state our main results, deferring their proofs to Sec. 3.

Let π_m denote the collection of all polynomials in one variable having degree at most m, and $\pi_{\nu,n}$ be the collection of all complex rational functions of the form

$$\frac{p(z)}{q(z)}, \quad \text{where } p \in \pi_\nu, \quad q \in \pi_n.$$

Next, for an arbitrary set A in the complex plane, we denote by $\|\cdot\|_A$ the supremum norm on A; i.e., for f defined on A

$$\|f\|_A := \sup\{|f(z)| : z \in A\}.$$

Our primary result is the following:

Theorem 2.1. There exists a sequence of rational functions $\{R_n(z)\}_{n=1}^{\infty}$ of the form

$$R_n(z) = \frac{p_{n-1}(z)}{\left(1+\frac{z}{n}\right)^n}, \qquad \text{with} \quad p_{n-1} \in \pi_{n-1} \quad \text{for all} \quad n \geqq 1, \tag{2.1}$$

which satisfies

$$\|e^{-x} - R_n(x)\|_{[0,+\infty)} = \mathcal{O}\left(\frac{n}{2^n}\right) \quad \text{as} \quad n \to \infty. \tag{2.2}$$

Consequently, for these rational functions, there exists a $q \geqq 2$ such that

$$\overline{\lim_{n\to\infty}} \|e^{-x} - R_n(x)\|_{[0,+\infty)}^{1/n} = \frac{1}{q} \leqq \frac{1}{2}. \tag{2.3}$$

Regarding Theorem 2.1, it should first be noted that Newman [3] has shown that for any sequence of rational functions $\{r_n(z)\}_{n=1}^{\infty}$, with $r_n \in \pi_{n,n}$ for all $n \geqq 2$, there holds

$$\|e^{-x} - r_{n-1}(x)\|_{[0,+\infty)}^{1/n} \geqq \frac{1}{1280}. \tag{2.4}$$

Errata in

"Geometric convergence to e^{-z} by rational functions with real poles"

by

E. B. Saff, A. Schönhage, and R. S. Varga

(Numer. Math. 25(1976), 307-322).

p. 308, eq. (2.2). Read "$O(\frac{n}{2^n})$" for "$O(\frac{n}{2!})$"

p. 310, first line of eq. (3.2). Read "$\ldots \frac{p(cx)}{(1+x)^n}$" for "$\ldots \frac{p(cx)}{(1+x)_n}$";

second line of eq. (3.2). Read "$\ldots e^{-c(\frac{1+t}{1-t})}$" for "$\ldots e_{-c}(\frac{1+t}{1-t})$"

p. 311, last line. Insert a right bracket after e^{-y}.

p. 312, line +4. Read "$\gamma_k = \frac{\cdots}{2^{k-\frac{1}{2}}}$" for "$\gamma_k = \frac{\cdots}{2^{k-\frac{1}{2}}}$"

p. 316, line +2. Raise up "tan" in display.

p. 318, eq. (3.41). Read "$\sum_{k=1}^{n} \frac{1}{a_k^2}$" for "$\sum_{k=1}^{n} \frac{1}{2a_k}$".

The same mistake appears again two lines below.

Hence a faster than geometric convergence rate to e^{-x} on $[0, +\infty)$ is not possible. We further remark that, in [6], Schönhage has proved that if

$$\lambda_{0,n} := \inf\{\|e^{-x} - \frac{1}{q_n(x)}\|_{[0,+\infty)} : q_n \in \pi_n\},$$

then

$$\lim_{n\to\infty} \lambda_{0,n}^{1/n} = \frac{1}{3}.$$

Thus, a geometric convergence rate faster than that indicated in (2.3) can be attained by reciprocals of polynomials; however, as a consequence of Theorem 2.5 below, rational functions of this type cannot have all their poles on the real axis.

Theorem 2.2. Let $\{R_n(z)\}_{n=1}^{\infty}$ be any sequence of rational functions of the form (2.1) which satisfies (2.3) for some $q \geqq 2$. Then, on every bounded subset K of the complex plane, there holds

$$\overline{\lim_{n\to\infty}} \|e^{-z} - R_n(z)\|_K^{1/n} \leqq \frac{1}{q} \leqq \frac{1}{2}. \tag{2.5}$$

Since the poles of the $R_n(z)$ are real, it follows immediately from Theorems 2.1, 2.2, and the results of Saff and Varga [4], that the sequence $\{R_n(z)\}_{n=1}^{\infty}$ must in fact converge geometrically to e^{-z} in an infinite sector symmetric about the positive x-axis. To precisely state this result, we introduce the following set notation:

$$S(\theta) := \{z \in \mathbb{C} : |\arg z| < \theta\}, \qquad 0 < \theta < \pi. \tag{2.6}$$

Theorem 2.3. Let $\{R_n(z)\}_{n=1}^{\infty}$ be any sequence of rational functions of the form (2.1) which satisfies (2.3) for some $q \geqq 2$. Then for each fixed θ, with

$$0 < \theta < 4\tan^{-1}\left(\frac{\sqrt{q}-1}{\sqrt{q}+1}\right), \tag{2.7}$$

where

$$\pi \geqq 4\tan^{-1}\left(\frac{\sqrt{q}-1}{\sqrt{q}+1}\right) \geqq 4\tan^{-1}\left(\frac{\sqrt{2}-1}{\sqrt{2}+1}\right) \doteqdot 38.942°,$$

the sequence $\{R_n(z)\}_{n=1}^{\infty}$ converges geometrically to e^{-z} on the closure $\bar{S}(\theta)$ of the infinite sector defined in (2.6). Moreover,

$$\overline{\lim_{n\to\infty}} \|e^{-z} - R_n(z)\|_{\bar{S}(\theta)}^{1/n} \leqq \frac{1}{q}\left(\frac{1+\sin(\theta/2)}{1-\sin(\theta/2)}\right) < 1. \tag{2.8}$$

We remark that the above property of geometric overconvergence to e^{-z} in infinite sectors is also shared by certain sequences of Padé approximants to e^{-z}, as was shown by Saff, Varga, and Ni in [5]. These Padé approximants, however, have poles off of the real axis.

As for a lower bound for the error, an estimate much sharper than Newman's result of (2.4) can be obtained for the case when the rational approximations have only real poles. This result is stated in

Theorem 2.4. If $\{r_n(z)\}_{n=1}^{\infty}$, with $r_n \in \pi_{n,n}$ for all $n \geqq 1$, is any sequence of rational functions having only real poles (not necessarily coincident), then

$$\overline{\lim_{n\to\infty}} \|e^{-x} - r_n(x)\|_{[0,+\infty)}^{1/n} \geqq 3 - 2\sqrt{2} \doteqdot \frac{1}{5.828}. \tag{2.9}$$

In particular, for the sequence $\{R_n(z)\}_{n=1}^{\infty}$ of Theorem 2.1, we have

$$3-2\sqrt{2} \leqq \overline{\lim_{n\to\infty}} \|e^{-x} - R_n(x)\|_{[0,+\infty)}^{1/n} \leqq \frac{1}{2}. \tag{2.10}$$

As our final result, we shall prove the impossibility of geometric convergence to e^{-x} on $[0, +\infty)$ by reciprocals of polynomials having only real zeros.

Theorem 2.5. For every polynomial $q_n \in \pi_n$, $n \geqq 1$, having only real zeros, there holds

$$\left\|e^{-x} - \frac{1}{q_n(x)}\right\|_{[0,+\infty)} \geqq C\left(\frac{\ln n}{n}\right)^2, \tag{2.11}$$

where C is a positive constant independent of n. Consequently, we have

$$\underline{\lim_{n\to\infty}} \left\|e^{-x} - \frac{1}{q_n(x)}\right\|_{[0,+\infty)}^{1/n} \geqq 1,$$

for every sequence of polynomials $q_n \in \pi_n$, having only real zeros.

3. Proofs of New Results

We now justify the results stated in Sec. 2.

Proof of Theorem 2.1. For each $n \geqq 1$, it is convenient to work with rational functions of the general form

$$r(x) = \frac{p(x)}{\left(1+\frac{x}{c}\right)^n}, \qquad p \in \pi_{n-1}, \qquad c > 0.$$

As we shall see, the choice $c \doteq n$ is optimal for our error estimates.

For fixed $c > 0$, $n \geqq 1$, set

$$\mu_n(c) := \inf_{p\in\pi_{n-1}} \left\| e^{-x} - \frac{p(x)}{\left(1+\frac{x}{c}\right)^n} \right\|_{[0,+\infty)}. \tag{3.1}$$

Then, on replacing x by cx and on replacing x by $(1+t)/(1-t)$, it follows that

$$\begin{aligned}\mu_n(c) &= \inf_{p\in\pi_{n-1}} \left\| e^{-cx} - \frac{p(cx)}{(1+x)_n} \right\|_{[0,+\infty)} \\ &= \inf_{p\in\pi_{n-1}} \left\| e_{-c}\left(\frac{1+t}{1-t}\right) - \frac{(1-t)^n}{2^n}\, p\left(c\left(\frac{1+t}{1-t}\right)\right) \right\|_{[-1,1)} \\ &= \inf_{\substack{q\in\pi_n \\ q(1)=0}} \left\| e^{-c\left(\frac{1+t}{1-t}\right)} - q(t) \right\|_{[-1,1)} \\ &= \inf_{T\in\pi_{n-1}} \left\| \int_t^1 [f'(s) - T(s)]\, ds \right\|_{[-1,1)},\end{aligned} \tag{3.2}$$

where

$$f(s) := -e^{-c\left(\frac{1+s}{1-s}\right)}, \qquad (f(1) = 0). \tag{3.3}$$

Now, by Schwarz's inequality, we have for $t\in[-1,1]$ that

$$\left|\int_t^1 [f'(s)-T(s)]\,ds\right| \leqq \int_{-1}^1 |f'(s)-T(s)|\,ds \leqq \sqrt{2}\left(\int_{-1}^1 |f'(s)-T(s)|^2\,ds\right)^{1/2}.$$

Hence, if $\varrho_{n-1}(c)$ denotes the least squares error

$$\varrho_{n-1}(c):=\inf_{T\in\pi_{n-1}}\left(\int_{-1}^1 |f'(s)-T(s)|^2\,ds\right)^{1/2},$$

we deduce from (3.2) that

$$\mu_n(c)\leqq\sqrt{2}\,\varrho_{n-1}(c). \tag{3.4}$$

Using the Legendre polynomials [7]

$$P_k(s):=\frac{1}{2^k(k!)}\frac{d^k}{ds^k}[(s^2-1)^k] \tag{3.5}$$

as an orthogonal basis on $[-1,1]$, we have

$$\varrho_{n-1}(c)=\left(\sum_{k=n}^{\infty}\gamma_k^2\right)^{1/2}, \tag{3.6}$$

where

$$\gamma_k:=\frac{\sqrt{2k+1}}{\sqrt{2}}\int_{-1}^1 f'(s)\,P_k(s)\,ds. \tag{3.7}$$

Next, we show that γ_k can be expressed in the equivalent form

$$\gamma_k=\frac{c\sqrt{2k+1}}{\sqrt{2}}\int_0^\infty e^{-cx}\left(\frac{x}{1+x}\right)^k L_k^{(1)}[c(1+x)]\,dx, \tag{3.8}$$

where $L_k^{(1)}(u)$ denotes the Laguerre polynomials [7] belonging to the weight function ue^{-u} on $[0,+\infty)$, i.e.,

$$L_k^{(1)}(u):=\frac{1}{k!\,u\,e^{-u}}\frac{d^k}{du^k}[u^{k+1}e^{-u}]. \tag{3.9}$$

To justify (3.8), we first use (3.5) and (3.7) to write

$$\gamma_k=\frac{\sqrt{2k+1}}{2^k(k!)\sqrt{2}}\int_{-1}^1 f'(s)\frac{d^k}{ds^k}[(s^2-1)^k]\,ds,$$

and then we integrate by parts k times to obtain

$$\gamma_k=\frac{\sqrt{2k+1}\,(-1)^k}{2^k(k!)\sqrt{2}}\int_{-1}^1 (s^2-1)^k\frac{d^{k+1}f(s)}{ds^{k+1}}\,ds. \tag{3.10}$$

Using (3.3) and mathematical induction, it can be easily verified that, with $v=2c/(1-s)$,

$$e^{-c}(1-s)^{k+1}\frac{d^{k+1}f(s)}{ds^{k+1}}=\frac{d^k}{dv^k}[v^{k+1}e^{-v}, \qquad k=0,1,2,\ldots,$$

and so from (3.9), we have

$$(1-s)^{k+1}\frac{d^{k+1}f(s)}{ds^{k+1}}=\frac{e^c k!\,2c\,e^{-2c/(1-s)}}{(1-s)}\,L_k^{(1)}\left(\frac{2c}{1-s}\right).$$

Making this substitution, equation (3.10) becomes

$$\gamma_k=\frac{\sqrt{2k+1}\,e^c c}{2^{k-\frac{1}{2}}}\int\limits_{-1}^{1}(1+s)^k e^{-2c/(1-s)}\,L_k^{(1)}\left(\frac{2c}{1-s}\right)\frac{ds}{(1-s)^2},$$

and a final change of variables $x=(1+s)/(1-s)$ yields (3.8).

Now, since (see [7, p. 100])

$$\int\limits_0^\infty u\,e^{-u}[L_k^{(1)}(u)]^2\,du=k+1,$$

it follows by applying Schwarz's inequality to (3.8) that

$$\gamma_k^2=c^2\left(\frac{2k+1}{2}\right)\left[\int\limits_0^\infty(1+x)^{1/2}e^{-cx/2}L_k^{(1)}[c(1+x)]\cdot(1+x)^{-1/2}e^{-cx/2}\left(\frac{x}{1+x}\right)^k dx\right]^2$$

$$\leqq c^2\left(\frac{2k+1}{2}\right)\int\limits_0^\infty(1+x)\,e^{-cx}(L_k^{(1)}[c(1+x)])^2\,dx\cdot\int\limits_0^\infty e^{-cx}\frac{x^{2k}}{(1+x)^{2k+1}}\,dx$$

$$=e^c\left(\frac{2k+1}{2}\right)\int\limits_c^\infty u\,e^{-u}[L_k^{(1)}(u)]^2\,du\cdot\int\limits_0^\infty e^{-cx}\frac{x^{2k}}{(1+x)^{2k+1}}\,dx$$

$$<\frac{e^c(2k+1)\,(k+1)}{2}\int\limits_0^\infty e^{-cx}\frac{x^{2k}}{(1+x)^{2k+1}}\,dx,$$

with strict inequality holding because $c>0$.

Hence, from (3.6), we have

$$\varrho_{n-1}^2(c)<\frac{e^c}{2}\int\limits_0^\infty e^{-cx}g_n(x)\,dx, \tag{3.11}$$

where

$$g_n(x):=\sum_{k=n}^{\infty}(2k+1)\,(k+1)\frac{x^{2k}}{(1+x)^{2k+1}}. \tag{3.12}$$

We can obtain a closed form expression for $g_n(x)$ by setting $z=x/(1+x)$ and writing

$$g_n(x)=\frac{1}{2(1+x)}\sum_{k=n}^{\infty}(2k+1)\,(2k+2)\,z^{2k}=\frac{1}{2(1+x)}\frac{d^2}{dz^2}\left(\sum_{k=n}^{\infty}z^{2k+2}\right)$$

$$=\frac{1}{2(1+x)}\frac{d^2}{dz^2}\left(\frac{z^{2n+2}}{1-z^2}\right).$$

After performing the differentiation and observing that $(1-z^2)^{-1}=(1+x)^2/(1+2x)$, we find that

$$g_n(x)=\left(\frac{x}{1+x}\right)^{2n-2}\omega_n(x),$$

where

$$\omega_n(x)=\frac{(n+1)(2n+1)x^2}{(1+2x)(1+x)}+\frac{(4n+5)x^4}{(1+2x)^2(1+x)}+\frac{4x^6}{(1+2x)^3(1+x)}. \quad (3.13)$$

Consequently, inequality (3.11) becomes

$$\varrho_{n-1}^2(c)<\frac{e^c}{2}\int_0^\infty e^{-cx}\left(\frac{x}{1+x}\right)^{2n-2}\omega_n(x)\,dx=\frac{1}{2}\int_0^\infty e^{\varphi_n(x;c)}\cdot e^{-x}\omega_n(x)\,dx, \quad (3.14)$$

where

$$\varphi_n(x;c):=c-cx+x-(2n-2)\ln\left(1+\frac{1}{x}\right). \quad (3.15)$$

Finally, observe from (3.13) that

$$\int_0^\infty e^{-x}\omega_n(x)\,dx=\mathcal{O}(n^2), \qquad n\to\infty, \quad (3.16)$$

and so we endeavor to minimize the maximum value of $\varphi_n(x;c)$ on $[0,+\infty)$. For $c=n$, it is easily verified that

$$\max\{\varphi_n(x;n):x\geqq 0\}=\varphi_n(1;n)=1-(2n-2)\ln 2,$$

and, furthermore, $\varphi_n(1;c)=1-(2n-2)\ln 2$ for all $c>0$. Hence we choose $c=n$, which gives the estimate (see (3.14), (3.16))

$$\varrho_{n-1}^2(n)<\frac{e}{2^{2n-1}}\int_0^\infty e^{-x}\omega_n(x)\,dx=\mathcal{O}\left(\frac{n^2}{4^n}\right), \qquad n\to\infty.$$

Consequently, from (3.4) we arrive at the inequality

$$\mu_n(n)\leqq\sqrt{2}\,\varrho_{n-1}(n)=\mathcal{O}\left(\frac{n}{2^n}\right), \qquad n\to\infty, \quad (3.17)$$

which proves Theorem 2.1. ∎

We remark that the choice $c=1\cdot n$ in the above proof (as opposed to $c\cdot=\lambda\cdot n$) results in the best geometric decay in the upper bound of (3.17).

For the proof of Theorem 2.2, it is convenient to first state a special case of a result due to Walsh [9].

Lemma 3.1. For $\varrho>0$ and $s>1$, let $\varepsilon(\varrho,s)$ denote the ellipse in the complex z-plane with foci at $x=0$ and $x=\varrho$ and semimajor and semiminor axes a and b such that $b/a=(s^2-1)/(s^2+1)$, i.e.,

$$\varepsilon(\varrho,s):=\left\{z=x+iy\in\mathbb{C}:\frac{(x-\varrho/2)^2}{[(\varrho/4)(s+1/s)]^2}+\frac{y^2}{[(\varrho/4)(s-1/s)]^2}=1\right\}. \quad (3.18)$$

If $r_n\in\pi_{n,n}$ is a rational function which satisfies

$$\|r_n(x)\|_{[0,\varrho]}\leqq M<\infty \quad (3.19)$$

and all the poles of $r_n(z)$ lie on or exterior to $\varepsilon(\varrho, \lambda)$, then for $1 < \tau < \lambda$, there holds

$$\|r_n(z)\|_{\bar{\varepsilon}(\varrho,\tau)} \leqq M\left(\frac{\lambda\tau - 1}{\lambda - \tau}\right)^n, \tag{3.20}$$

where $\varepsilon(\varrho, \tau)$ denotes the closed interior of $\varepsilon(\varrho, \tau)$.

Proof of Theorem 2.2. More generally, let $\{r_n(z)\}_{n=1}^{\infty}$ be any sequence of rational functions with $r_n \in \pi_{n,n}$ which satisfies (2.3), and is such that the poles of r_n tend to infinity as $n \to \infty$. Next, let K be a fixed bounded set in the plane. Define

$$\tilde{r}_{2n}(z) := r_n(z) - r_{n-1}(z), \qquad \text{for } n = 2, 3, \ldots.$$

Clearly, $\tilde{r}_{2n} \in \pi_{2n,2n}$, and, as $n \to \infty$, all the poles of the $\tilde{r}_{2n}(z)$ tend to infinity. Also, from (2.3) and the triangle inequality, it follows that

$$\overline{\lim_{n\to\infty}} \|\tilde{r}_{2n}(x)\|_{[0,+\infty)}^{1/n} \leqq \frac{1}{q};$$

in particular,

$$\overline{\lim_{n\to\infty}} \|\tilde{r}_{2n}(x)\|_{[0,\varrho]}^{1/n} \leqq \frac{1}{q}, \qquad \text{for each} \quad \varrho > 0. \tag{3.21}$$

Now, let $\tau > 1$ be given and choose $\varrho^* > 0$ so large that $K \subset \bar{\varepsilon}(\varrho^*, \tau)$, where $\bar{\varepsilon}(\varrho^*, \tau)$ is as defined in Lemma 3.1. Furthermore, let $\lambda > \tau$. Then for n sufficiently large, all the poles of the $\tilde{r}_{2n}(z)$ lie outside $\varepsilon(\varrho^*, \lambda)$, and so from (3.21) and Lemma 3.1, it follows that

$$\overline{\lim_{n\to\infty}} \|\tilde{r}_{2n}(z)\|_{\bar{\varepsilon}(\varrho^*,\tau)}^{1/n} \leqq \frac{1}{q}\left(\frac{\lambda\tau - 1}{\lambda - \tau}\right)^2.$$

Consequently,

$$\overline{\lim_{n\to\infty}} \|\tilde{r}_{2n}(z)\|_K^{1/n} \leqq \frac{1}{q}\left(\frac{\lambda\tau - 1}{\lambda - \tau}\right)^2. \tag{3.22}$$

Since (3.22) is valid for every λ with $\lambda > \tau$, then on letting $\lambda \to +\infty$ we obtain

$$\overline{\lim_{n\to\infty}} \|\tilde{r}_{2n}(z)\|_K^{1/n} \leqq \frac{1}{q}\tau^2.$$

Finally, as $\tau > 1$ was arbitrarily chosen, it follows that

$$\overline{\lim_{n\to\infty}} \|r_n(z) - r_{n-1}(z)\|_K^{1/n} = \overline{\lim_{n\to\infty}} \|\tilde{r}_{2n}(z)\|_K^{1/n} \leqq \frac{1}{q},$$

which implies by a standard M-test argument that the sequence $\{r_n(z)\}_{n=1}^{\infty}$ converges geometrically to e^{-z} in the uniform norm on K, and that

$$\overline{\lim_{n\to\infty}} \|e^{-z} - r_n(z)\|_K^{1/n} \leqq \frac{1}{q} \leqq \frac{1}{2}. \qquad \blacksquare$$

For the proofs of Theorems 2.3 and 2.4, we need the following result of Saff and Varga [4] concerning overconvergence in sectors.

Lemma 3.2. For $\theta > 0$ and $\mu \geqq 0$, let $S(\theta, \mu)$ denote generically the set

$$S(\theta, \mu) := \{z \in \mathbb{C} : |\arg z| < \theta, |z| > \mu\}. \tag{3.23}$$

Assume that for a function f defined and finite on $[0, +\infty)$, there exists a sequence of rational functions $\{r_n(z)\}_{n=1}^{\infty}$, with $r_n \in \pi_{n,n}$ for all $n \geqq 1$, and a real number $q' > 1$ such that

$$\overline{\lim_{n\to\infty}} \|f(x) - r_n(x)\|_{[0,+\infty)}^{1/n} < \frac{1}{q'} < 1. \tag{3.24}$$

Assume further that for some θ_0 and μ_0, with $0 < \theta_0 \leqq \pi$, $\mu_0 \geqq 0$, the region $S(\theta_0, \mu_0)$ contains no poles of the $r_n(z)$ for all n large. Then, for every θ satisfying the inequality

$$0 < \theta < 4 \tan^{-1}\left\{\left(\frac{\sqrt{q'}-1}{\sqrt{q'}+1}\right) \tan\left(\frac{\theta_0}{4}\right)\right\}, \tag{3.25}$$

there exists a $\mu = \mu(\theta) > 0$ and an analytic function $F(z)$ on the closure $\overline{S}(\theta, \mu)$ with $F(x) = f(x)$ for all real x in this set, such that $\{r_n(z)\}_{n=1}^{\infty}$ converges geometrically to $F(z)$ on $\overline{S}(\theta, \mu)$. Moreover

$$\overline{\lim_{n\to\infty}} \|F - r_n\|_{\overline{S}(\theta,\mu)}^{1/n} < \frac{1}{q'} \left\{\frac{\sin\left[\frac{1}{4}(\theta_0+\theta)\right]}{\sin\left[\frac{1}{4}(\theta_0-\theta)\right]}\right\}^2 < 1. \tag{3.26}$$

Proof of Theorem 2.3. For any sequence $\{R_n(z)\}_{n=1}^{\infty}$ of rational functions of the form (2.1) which satisfies (2.3), the hypotheses of Lemma 3.2 are fulfilled with $r_n(z) = R_n(z)$, $f(x) = e^{-x}$, $1 < q' < q$, $\theta_0 = \pi$, $\mu_0 = 0$. Now, fix θ satisfying (2.7) and choose q' with $1 < q' < q$, sufficiently close to q, so that inequality (3.25) holds. Then by Lemma 3.2, there exists a $\mu = \mu(\theta)$ such that

$$\overline{\lim_{n\to\infty}} \|e^{-z} - R_n(z)\|_{\overline{S}(\theta,\mu)}^{1/n} < \frac{1}{q'} \left\{\frac{\sin\left[\frac{1}{4}(\pi+\theta)\right]}{\sin\left[\frac{1}{4}(\pi-\theta)\right]}\right\}^2 = \frac{1}{q'}\left(\frac{1+\sin(\theta/2)}{1-\sin(\theta/2)}\right) < 1. \tag{3.27}$$

Furthermore, by Theorem 2.2, we have for the set $K = \{z : |z| \leqq \mu\}$

$$\overline{\lim_{n=\infty}} \|e^{-z} - R_n(z)\|_K^{1/n} \leqq \frac{1}{q} \leqq \frac{1}{q'},$$

and this inequality together with (3.27) implies that

$$\overline{\lim_{n\to\infty}} \|e^{-z} - R_n(z)\|_{\overline{S}(\theta)}^{1/n} < \frac{1}{q'}\left(\frac{1+\sin(\theta/2)}{1-\sin(\theta/2)}\right) < 1,$$

where $S(\theta)$ is defined in (2.6). Letting $q' \to q$ we obtain (2.8). ∎

Proof of Theorem 2.4. Assume to the contrary that there exists a sequence $\{r_n(z)\}_{n=1}^{\infty}$, with $r_n \in \pi_{n,n}$ for all $n \geqq 1$, such that the $r_n(z)$ have only real poles and satisfy

$$\overline{\lim_{n\to\infty}} \|e^{-x} - r_n(x)\|_{[0,+\infty)}^{1/n} < \frac{1}{q_0},$$

where

$$\frac{1}{q_0} = \left(\frac{1-\tan(\pi/8)}{1+\tan(\pi/8)}\right)^2 = 3 - 2\sqrt{2}.$$

Choose q' so that

$$\overline{\lim_{n\to\infty}} \|e^{-x} - r_n(x)\|_{[0,+\infty)}^{1/n} < \frac{1}{q'} < \frac{1}{q_0}.$$

Then since, as is readily verified,

$$4 \tan^{-1}\left\{\left(\frac{\sqrt{q_0}-1}{\sqrt{q_0}+1}\right) \tan\left(\frac{\pi}{4}\right)\right\} = \frac{\pi}{2},$$

we can find an angle θ which satisfies

$$\frac{\pi}{2} < \theta < 4\tan^{-1}\left\{\left(\frac{\sqrt{q'}-1}{\sqrt{q'}+1}\right)\tan\left(\frac{\pi}{4}\right)\right\}.$$

Hence, by Lemma 3.2 (with $\theta_0=\pi$), the $r_n(z)$ must converge geometrically to e^{-z} on the unbounded closed region $\bar{S}(\theta,\mu)$ (defined in (3.23)) for some $\mu>0$. But this in particular implies that $|e^{-z}|$ is bounded in $\bar{S}(\theta,\mu)$, which is absurd since $\theta>\pi/2$. ∎

Proof of Theorem 2.5. Let $q_n\in\pi_n$ be a polynomial having only real zeros. To establish the lower bound of (2.11) it suffices to assume that $q_n(x)$ has the normalization

$$q_n(x)=\prod_{k=1}^{n}\left(1+\frac{x}{a_k}\right), \quad \text{with} \quad a_k>0, \quad k=1,2,\ldots,n. \tag{3.28}$$

Furthermore, we may assume that $n\geqq n_0$ and that

$$\eta:=\left\|e^{-x}-\frac{1}{q_n(x)}\right\|_{[0,+\infty)}=e^{-r}/2, \quad \text{for some } r\geqq r_0>0, \tag{3.29}$$

where n_0 and r_0 are absolute constants which shall be specified later. The essence of the proof lies in abtaining upper and lower bounds for the positive quantity

$$A:=\sum_{k=1}^{n}\frac{1}{a_k}. \tag{3.30}$$

For $0\leqq x\leqq r$, it follows from (3.29) that

$$\frac{1}{q_n(x)}\geqq e^{-x}-\frac{e^{-r}}{2}\geqq\frac{1}{2e^x},$$

and hence

$$|e^x-q_n(x)|\leqq e^x q_n(x)\eta\leqq e^{2x-r}, \quad \text{for} \quad 0\leqq x\leqq r. \tag{3.31}$$

Since $t\geqq\ln(1+t)$ for any $t\geqq 0$, we have

$$xA=\sum_{k=1}^{n}\frac{x}{a_k}\geqq\sum_{k=1}^{n}\ln\left(1+\frac{x}{a_k}\right)=\ln q_n(x), \quad x\geqq 0,$$

which together with (3.31) implies that

$$xA\geqq\ln q_n(x)\geqq\ln(e^x-e^{2x-r}), \quad \text{for} \quad 0\leqq x<r.$$

Putting $x=r/2$ in the last inequality gives

$$\frac{r}{2}A\geqq\ln q_n(r/2)\geqq\frac{r}{2}+\ln(1-e^{-r/2}), \tag{3.32}$$

or

$$A\geqq 1+\frac{2}{r}\ln(1-e^{-r/2}). \tag{3.33}$$

As the right-hand side of (3.33) tends to 1 as $r\to+\infty$, we can choose r_0 so large that

$$A\geqq 1/2, \quad \text{for} \quad r\geqq r_0. \tag{3.34}$$

To establish an upper bound for A we first note that

$$xA \leqq \prod_{k=1}^{n}\left(1+\frac{x}{a_k}\right)-1=q_n(x)-1, \qquad x\geqq 0.$$

Thus from (3.31) we have

$$xA \leqq e^x+e^{2x-r}-1, \qquad 0\leqq x\leqq r,$$

or

$$A\leqq \frac{e^x-1}{x}+\frac{e^{2x-r}}{x}, \qquad 0<x\leqq r. \tag{3.35}$$

But since

$$\frac{e^x-1}{x}=1+\frac{x}{2!}+\frac{x^2}{3!}+\cdots\leqq 1+\frac{1}{2}(x+x^2+x^3+\cdots), \qquad \text{for} \quad 0<x<1,$$

it follows that

$$\frac{e^x-1}{x}\leqq 1+\frac{x}{2(1-x)}, \qquad 0<x<1,$$

and hence from (3.35) we obtain

$$A\leqq 1+\frac{x}{2(1-x)}+\frac{e^{2x-r}}{x}, \qquad 0<x<1.$$

Choosing $x=e^{-r/2}<1$, the last inequality becomes

$$A\leqq 1+e^{-r/2}\lambda(r), \tag{3.36}$$

where

$$\lambda(r):=\frac{1}{2(1-e^{-r/2})}+\exp\,(2e^{-r/2}).$$

As $\lambda(r)$ decreases to 1.5 as $r\to+\infty$, we can additionally require r_0 to be so large that $\lambda(r)\leqq 1.6$ for all $r\geqq r_0$, whence

$$A\leqq 1+1.6\,e^{-r/2}, \qquad r\geqq r_0. \tag{3.37}$$

Now from (3.37) and (3.32) we obtain

$$\frac{r}{2}A-\ln q_n(r/2)\leqq 0.8\,r\,e^{-r/2}-\ln(1-e^{-r/2}),$$

and again we can choose r_0 large enough so that

$$0.8\,r\,e^{-r/2}-\ln(1-e^{-r/2})\leqq r\,e^{-r/2}, \qquad \text{for all} \quad r\geqq r_0;$$

consequently

$$\frac{r}{2}A-\ln q_n(r/2)=\sum_{k=1}^{n}\left[\frac{r}{2a_k}-\ln\left(1+\frac{r}{2a_k}\right)\right]\leqq r\,e^{-r/2}, \qquad r\geqq r_0. \tag{3.38}$$

Next, observe that each term in the above sum is positive, and thus, no one term can exceed the right-hand side, i.e.,

$$0<\frac{r}{2a_k}-\ln\left(1+\frac{r}{2a_k}\right)\leqq r\,e^{-r/2}, \qquad \text{for} \quad 1\leqq k\leqq n, \qquad r\geqq r_0. \tag{3.39}$$

Making the final assumption that r_0 be so large that

$$\frac{1}{2}-\ln\left(1+\frac{1}{2}\right)\geqq r_0 e^{-r_0/2},$$

it follows from (3.39) that

$$0 < \frac{r}{2\,a_k} \leqq \frac{1}{2}, \qquad \text{for} \quad 1 \leqq k \leqq n, \quad r \geqq r_0. \tag{3.40}$$

But for $0 < t \leqq 1/2$, we have

$$t - \ln(1+t) = t$$
$$-\left(t - \frac{t^2}{2} + \frac{t^3}{3} - \cdots\right) \geqq \frac{t^2}{2} - \frac{t^3}{3} + \cdots \geqq \frac{t^2}{2} - \frac{t^3}{3} \geqq \frac{t^2}{3}\left(\frac{3}{2} - t\right) \geqq \frac{t^2}{3};$$

whence from (3.40) and (3.38) we deduce

$$\sum_{k=1}^{n} \frac{1}{3}\left(\frac{r}{2\,a_k}\right)^2 \leqq r\,e^{-r/2},$$

or

$$\sum_{k=1}^{n} \frac{1}{2a_k} \leqq \frac{12}{r}\,e^{-r/2}, \qquad r \geqq r_0. \tag{3.41}$$

Combining (3.34) and (3.41) with Schwarz's inequality gives

$$\frac{1}{2} \leqq A = \sum_{k=1}^{n} \frac{1}{a_k} \leqq \sqrt{n}\left(\sum_{k=1}^{n} \frac{1}{2\,a_k}\right)^{1/2} \leqq \left(\frac{12\,n}{r}\right)^{1/2} e^{-r/4},$$

and hence for $\gamma := 1/(2\sqrt{12})$ we have

$$\gamma\left(\frac{r}{n}\right)^{1/2} \leqq e^{-r/4},$$

or

$$\frac{\gamma^4 r^2}{n^2} \leqq e^{-r}, \qquad r \geqq r_0. \tag{3.42}$$

Since the function

$$h(t) := \frac{\gamma^4 t^2}{n^2} - e^{-t}$$

is strictly increasing for $t \geqq 0$, and since, as is easily verified,

$$h\{\ln(n^2/\gamma^4 \ln^2 n)\} > 0$$

for all n sufficiently large, say $n \geqq n_0$, it follows from (3.42) that

$$r < \ln(n^2/\gamma^4 \ln^2 n), \qquad n \geqq n_0.$$

Hence,

$$\eta = \frac{e^{-r}}{2} \geqq \frac{\gamma^4 \ln^2 n}{2\,n^2} = C\left(\frac{\ln n}{n}\right)^2, \qquad n \geqq n_0, \qquad r \geqq r_0,$$

which completes the proof. ∎

4. Numerical Results and Applications

As in (3.1), we set

$$\mu_n(n) := \inf_{p \in \pi_{n-1}} \left\| e^{-x} - \frac{p(x)}{\left(1 + \frac{x}{n}\right)^n} \right\|_{[0,+\infty)}, \qquad n \geqq 1. \tag{4.1}$$

By means of a modified Remez algorithm, the numbers $\{\mu_n(n)\}_{n=1}^{20}$ were computed, and these, along with $\mu_n(n)^{1/n}$, are given in Table I. Now, from (2.10) of Theorem 2.4, we know that there is a $\tilde{q} \geqq 2$ such that

$$3 - 2\sqrt{2} \leqq \overline{\lim_{n\to\infty}}\, \mu_n(n)^{1/n} := \frac{1}{\tilde{q}} \leqq \frac{1}{2}. \tag{4.2}$$

and the computations of Table I appear to indicate that $5.828 \doteq (3-2\sqrt{2})^{-1} > \tilde{q} > 2$. The exact value of $\tilde{q}$ satisfying (4.2) is, as yet, unknown.

Finally, if

$$\mu_n(n) = \left\| e^{-x} - \frac{\tilde{p}_{n-1}(x)}{\left(1+\frac{x}{n}\right)^n} \right\|_{[0,+\infty)}, \quad \text{where } \tilde{p}_{n-1} \in \pi_{n-1} \text{ for all } n \geqq 1, \tag{4.3}$$

then the sequence of rational functions

$$\left\{ \tilde{R}_n(z) := \frac{\tilde{p}_{n-1}(z)}{\left(1+\frac{z}{n}\right)^n} \right\}_{n=1}^{\infty} \tag{4.4}$$

has attractive features in applications. Consider the numerical solution of the linear system of ordinary differential equations

$$\frac{d\underline{u}(t)}{dt} = -A\,\underline{u}(t) + \underline{k}, \quad t > 0, \qquad \underline{u}(0) = \underline{u}_0, \tag{4.5}$$

where $\underline{u}(t) = [u_1(t), \ldots, u_m(t)]^T$ is a column vector with m components, and where A is a fixed real $m x m$ positive definite symmetric matrix. Such systems arise (cf. [1]) from semi-discretization of heat-conduction (parabolic) partial differential equations. The solution of (4.5) is of course given explicitly by

$$\underline{u}(t) = A^{-1}\underline{k} + \exp(-tA)\{\underline{u}_0 - A^{-1}\underline{k}\}, \quad \text{for all } t \geqq 0, \tag{4.6}$$

where $\exp(-tA) := \sum_{k=0}^{\infty} (-tA)^k/k!$.

Now, since $\tilde{R}_n(z)$ of (4.4) is an approximation of e^{-z}, then we consider approximating $\exp(-\Delta tA)$ by $\tilde{R}_n(\Delta tA) := \left(I + \frac{\Delta tA}{n}\right)^{-n} \tilde{p}_{n-1}(\Delta tA)$. Thus, for fixed n and fixed Δt, we compute

$$\underline{w}^{(r)} := A^{-1}\underline{k} + \tilde{R}_n(\Delta tA) \cdot \{\underline{w}^{(r-1)} - A^{-1}\underline{k}\}, \quad r \geqq 1, \tag{4.7}$$

as an approximation of $\underline{u}(r\Delta t)$, where $\underline{w}^{(0)} := \underline{u}_0$, or equivalently

$$\left(I + \frac{\Delta tA}{n}\right)^n \underline{w}^{(r)} = \left(I + \frac{\Delta tA}{n}\right)^n A^{-1}\underline{k} + \tilde{p}_{n-1}(\Delta tA)\{\underline{w}^{(r-1)} - A^{-1}\underline{k}\}. \tag{4.8}$$

But, because of the particular factored form of the matrix on the left above, this means that $w^{(r)}$ can be obtained from the repeated inversion of

$$\left(I + \frac{\Delta tA}{n}\right)\underline{g}_{l+1} = \underline{g}_l, \quad 0 \leqq l \leqq n-1, \tag{4.9}$$

n times, i.e., with $\underline{g}_0$ set equal to the right-hand side of (4.8), then $\underline{g}_n = \underline{w}^{(r)}$.

Acknowledgement. We are indebted to Mr. David Tolle of the University of South Florida for his computations, on an IBM 360/65, of the entries in Table I. We also are grateful for helpful computational comments from Dr. Wayne Fullerton of the Los Alamos Scientific Laboratory.

References

1. Cody, W. J., Meinardus, G., Varga, R. S.: Chebyshev rational approximations to e^{-x} on $[0, +\infty)$ and applications to heat-conduction problems. J. Approximation Theory **2**, 50–65 (1969)
2. Ehle, B.L.: A-stable methods and Padé approximation to the exponential. SIAM J. Math. Anal., **4**, 671–680 (1973)
3. Newman, D. J.: Rational approximation to e^{-x}. J. Approximation Theory **10**, 301–303 (1974)
4. Saff, E.B., Varga, R. S.: Angular overconvergence for rational functions converging geometrically on $[0, +\infty)$, to appear in the Proceedings of the Approximation Theory Conference. Calgary, Alberta, Canada, August 11–13, 1975
5. Saff, E.B., Varga, R. S., and Ni, W.-C.: Geometric convergence of rational approximations to e^{-z} in infinite sectors. Numer. Math. (to appear).
6. Schönhage, A.: Zur rationalen Approximierbarkeit von e^{-x} über $[0, +\infty)$. J. Approximation Theory **7**, 395–398 (1973)
7. Szegö, G.: Orthogonal Polynomials, Colloq. Publication Vol. 23, Amer. Math. Soc., Providence, R.I., 3rd ed., 1967
8. Underhill, C., Wragg, A.: Convergence properties of Padé approximants to exp (z) and their derivatives. J. Inst. Math. Appl. **11**, 361–367 (1973)
9. Walsh, J.L.: Interpolation and Approximation by Rational Functions in the Complex Domain, Colloq. Publication Vol. 20, Amer. Math. Soc., Providence, R.I., 5th ed., 1969.

E.B. Saff
Dept. of Mathematics
University of South Florida
Tampa, FL 33620
U.S.A.

A. Schönhage
Mathematisches Institut
Universität Tübingen
D-7400 Tübingen 1
Bundesrepublik Deutschland

R. S. Varga
Dept. of Mathematics
Kent State University
Kent, OH 44242
U.S.A.