JOURNAL OF APPROXIMATION THEORY 32, 85–95 (1981)

Approximation of e^{-x} by Rational Functions with Concentrated Negative Poles

JAN-ERIK ANDERSSON

Department of Mathematics, University of Göteborg, Göteborg, Sweden

Communicated by Richard S. Varga

Received April 30, 1979

1. INTRODUCTION

In a number of papers uniform approximation of e^{-x} on $[0, \infty)$ by rational functions has been investigated. Because of its importance in certain numerical methods, Saff *et al.* [4] studied the degree of approximation by functions of the form $p(x)(x+nq)^{-n}$, $q>0$, p polynomial of degree at most n. The method used in [4] gave the best order of approximation for $q=1$. In this case it was shown that the order is at most $O(2^{-n})$.

Recently, Kaufman and Taylor [3] commented on this result. Among other things they asked if the choice $q=1$ is optimal. In this paper we shall prove a theorem that shows that it is not. The best choice is $q=1/\sqrt{2}$ and then the degree of approximation is essentially of the order $(\sqrt{2}+1)^{-n}$.

2. FOMULATION OF THE RESULT

We give a result that determines the asymptotic behaviour of the degree of approximation of e^{-x} on $[0, \infty)$ by functions of the form $p(x)(x+a_n)^{-n}$, for any sequence of $a_n>0$. For reasons that will be clear later we prefer to rewrite a_n as nq_n.

For a sequence $Q=(q_n)_1^\infty$ of positive real numbers let

$$R_n(Q) = \{\text{functions of type } p(x)(x+nq_n)^{-n},\ p \in \Pi_n\}.$$

Here and in the sequal Π_n denotes the class of polynomials of degree at most n. Furthermore, let

$$\rho_n(Q) = \inf\{\sup_{x \geq 0} |e^{-x} - r(x)| : r \in R_n(Q)\}.$$

0021-9045/81/060085–11\$02.00/0

The following function G will be central:

$$G(z, q) = \log \left| \frac{z-1}{z+1} \right| + q \operatorname{Re} z^2, \qquad q > 0 \text{ and } z \in \mathbb{C}.$$

In stating our theorem we need another

DEFINITION. For $q > 0$ let q^* be the root in $\operatorname{Im} z \geqslant 0$ of the equation

$$q(z^3 - z) + 1 = 0,$$

that has smallest positive real part. Also let

$$G^*(q) = G(q^*, q).$$

The case when $q_n = q$ is fixed for all n is easiest to handle. In fact we shall prove that in this case

$$\lim \rho_n(Q)^{1/n} = \exp G^*(q).$$

The minimum value of the limit is $\sqrt{2} - 1$ and is obtained for $q = 1/\sqrt{2}$.

For the general situation we have the following result.

THEOREM. (i) $\overline{\lim}_{n \to \infty} \rho_n(Q)^{1/n} = \overline{\lim}_{n \to \infty} \exp G^*(q_n)$,

(ii) $\underline{\lim}_{n \to \infty} \rho_n(Q)^{1/n} = \underline{\lim}_{n \to \infty} \exp G^*(q_n)$,

(iii) $\exp G^*(q) \geqslant \sqrt{2} - 1$ *for all* $q > 0$ *and with equality if and only if* $q = 1/\sqrt{2}$.

The proof of the theorem will be given first for $q_n = q \in (0, \infty)$, $n = 1, 2, \ldots$, and then generalized. In this special case we consequently must prove

$$\lim \rho_n(Q)^{1/n} = \exp G^*(q).$$

3. Preliminary Transformations

In the definition of $\rho_n(Q)$ we perform the transformation

$$t = (nq - x)/(nq + x)$$

and find that

$$\rho_n(Q) = \inf\{\|f_n - p\|_{[-1,1]}: p \in \Pi_n\}, \tag{1}$$

where $f_n(t) = \exp[nq(t-1))/(t+1)]$ and $\|\cdot\|$ denotes the sup-norm.

For technical reasons we prefer to approximate on the closed complex unit disc Δ. Andersson and Ganelius [1] discussed a way to reduce problems concerning rational approximation on more general sets, so-called Faber sets, to approximation on Δ. In our special case with polynomial approximation on $[-1, 1]$ this can be done as follows.

Let $\tilde{f}$ be continuous on Δ and analytic in the interior of Δ. Then we define

$$E_n(\tilde{f}) = \inf\{\|\tilde{f} - p\|_\Delta : p \in \Pi_n\}.$$

Now let $\psi(w) = (w + w^{-1})/2$, mapping the exterior of Δ onto the exterior of $[-1, 1]$. For a continuous function f on $[-1, 1]$ we define

$$\tilde{f}(w) = \frac{1}{2\pi i} \int_C \frac{f \circ \psi(u)}{u - w} \, du, \qquad w \in \mathbb{C} \text{ and } |w| < 1.$$

Here C is the positively oriented unit circle in $\mathbb{C}$. The crucial properties that we shall need are the following. The mapping $\tilde{f} \to f$

(1) is bounded (i.e., there is a constant K such that $\|f\|_{[-1,1]} \leqslant K \|\tilde{f}\|_\Delta$),

(2) maps w^n, $n \geqslant 0$, on a polynomial of degree n.

For details and further references we refer the reader to [1]. Let us just mention that property (1) is a consequence of

$$\int_C |d_w \arg[\psi(rw) - t]| = 2\pi$$

for $r > 1$ and $t \in [-1, 1]$.

From these observations it follows that

$$\rho_n(Q) \leqslant K E_n(\tilde{f}_n),$$

where f_n is as in (1). To simplify our notation we denote $F_n(u) = f_n \circ \psi(u)$. Thus we have

$$F_n(u) = \exp\left[nq\left(\frac{u-1}{u+1}\right)^2\right]$$

and

$$2\pi i \tilde{f}_n(w) = \int_C F_n(u)(u - w)^{-1} \, du.$$

Let P_n be the Taylor polynomial of order $n-1$ at the origin for $\tilde{f}_n$. Then for $|w|<1$

$$2\pi i(\tilde{f}(w)-P_n(w)) = w^n \int_C F_n(u)\, u^{-n}(u-w)^{-1}\, du$$

and we find that

$$E_n(\tilde{f}_n) \leqslant \sup_{|w|\leqslant 1} \left| \int_C F_n(u)\, u^{-n}(u-w)^{-1}\, du \right|. \tag{2}$$

The function $F_n(u)$ is analytic in the whole extended plane except for $u=-1$. Hence the contour C in (2) can be replaced by other curves γ as long as they behave adequately at $u=-1$. We want to choose a curve so that $\max|F_n(u)u^{-n}|$ is as small as possible on it.

To simplify our discussion we make a final mapping,

$$z=(u-1)/(u+1), \qquad \text{i.e.,} \quad u=u(z)=(1+z)/(1-z). \tag{3}$$

Then

$$|F_n(u)u^{-n}| = \exp[nG(z,q)]$$

and for later use we define for each $q>0$ a function H by

$$H(z)=\log((1-z)/(1+z))+qz^2.$$

Consequently $G(z,q)=\operatorname{Re} H(z)$.

4. An Extremal Curve

Definition. Let Σ denote the class of all closed curves σ in the extended right half-plane that are symmetric with respect to the real axis and passes through the point of infinity.

For $\sigma \in \Sigma$ we let $g(\sigma)=\max\{G(z,q): z \text{ on } \sigma\}$ and

$$\hat{g}=\inf\{g(\sigma): \sigma\in\Sigma\}. \tag{4}$$

Because of the behaviour of $G(z,q)$ as $z\to\infty$ we may assume that for R sufficiently large the set $\sigma\cap\{|z|\geqslant R\}$ consists of the two rays $z=t(1\pm i\sqrt{3})$, $t\geqslant R/2$. Hence we restrict our attention to the set $\{z: |z|\leqslant R, \operatorname{Re} z\geqslant 0\}$. On this set G has only one singular point, namely, $z=1$. Since $G(z,q)\to-\infty$ as $z\to 1$, it is possible to find curves $\sigma\in\Sigma$ such that $g(\sigma)=\hat{g}$.

Returning to $G(z, q) = \text{Re}\, H(z)$ we find

$$G'_x = \text{Re}\, H' = 2\, \text{Re}((z^2 - 1)^{-1} + qz) \qquad (z = x + iy),$$

and consequently $G'_x < 0$ for x close to 0. Therefore we can take our optimal σ so that it never meets the imaginary axis. Hence we can choose our optimal σ so that $\hat{g}$ is obtained only at points where grad $G = 0$. By the Cauchy–Riemann equations these points are determined by $H'(z) = 0$, i.e.,

$$1 + q(z^3 - z) = 0.$$

Since $q > 0$ this equation has always one negative root which we shall neglect in the sequel. The other two roots are in $\text{Re}\, z \geqslant 0$ and are

—complex conjugated if $q < 3\sqrt{3}/2$,

—both equal to $z = 1/\sqrt{3}$ if $q = 3\sqrt{3}/2$,

—real, not equal, between 0 and 1 if $q > 3\sqrt{3}/2$.

These points all yield saddle points for $G(z, q)$.

Later we shall use a steepest descent method to get the lower bound. Therefore we must choose the curves σ with some precaution.

In the first case we can take σ so that it follows the directions of steepest descent for $G(z, q)$ at least close to the critical points q^* and $\overline{q^*}$.

In the third case that function $G(z, q)$ has a local minimum along the real axis at the critical point with smallest real part. At this point q^*, the directions of steepest descent are orthogonal to the real axis. The second critical point gives a local maximum for $G(z, q)$ along the real axis, and the directions of steepest descent are along the real axis. Since σ is supposed to be symmetric with respect to the real axis, the optimal curve will never pass this second critical point.

In the remaining case, $q = 3\sqrt{3}/2$ and $q^* = 1/\sqrt{3}$, we see that

$$G(z, q) = G^*(q) + 6^{-1} H^{(3)}(q^*)\, \text{Re}(z - q^*)^3 (1 + o(1)),$$

where $H^{(3)}(q^*) < 0$. Hence from this point there are three different directions of steepest descent, namely, along the positively oriented real axis and with angles $\pm 2\pi/3$ with respect to this direction. The optimal curve can be taken to follow the last two directions. Owing to the symmetry, the first direction will just take us to $z = 1$ and back.

Hence in all cases we find an optimal curve σ such that $\hat{g} = g(\sigma) = G(q^*, q) = G^*(q)$ and close to q^* (and $\overline{q^*}$) σ follows the directions of steepest descent for G. The regularity of G guarantees that σ can be taken at least piecewise smooth.

Part (iii) of our theorem now follows easily. The optimal curve σ must pass the ray $z = t(1 + i)$, $t \geqslant 0$. On this ray $G(z, q) = \log|(z - 1)/(z + 1)|$.

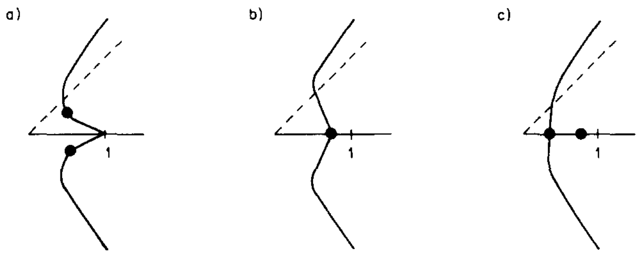

FIG. 1. The curve σ in the cases (a) $q < 3\sqrt{3}/2$, (b) $q = 3\sqrt{3}/2$ and (c) $q > 3\sqrt{3}/2$. The saddle points and the ray $\arg z = \pi/4$ are also indicated.

Hence for $q > 0$,

$$G^*(q) = g(\sigma) \geqslant \min\{G(z, q) : z = t(1 + i), t \geqslant 0\} = \log(\sqrt{2} - 1).$$

The minimum value is obtained for $z = (1 + i)/\sqrt{2}$. Therefore the only chance to get $G^*(q) = \log(\sqrt{2} - 1)$ is that $q^* = (1 + i)/\sqrt{2}$. This happens if and only if $q = 1/\sqrt{2}$ and thus (iii) in the theorem is proved.

5. The Upper Estimate when $q_n = q$

We now assume that $q_n = q > 0$ for all n. The optimal curve σ from Section 4 is mapped back to the u-plane by (3). Thus we get a piecewise smooth curve γ on which

$$|F_n(u)\, u^{-n}| \leqslant \exp[nG^*(q)]. \tag{5}$$

The behaviour of σ as $z \to \infty$ implies that γ approaches -1 along the circles $\arg[(u - 1)/(u + 1)] = \pm\pi/3$.

This means that in (2)

$$|u - w|^{-1} \leqslant K\,|u + 1|^{-1}, \qquad u \in \gamma, \quad |w| \leqslant 1,$$

for some constant K.

In the construction of σ we can avoid the point $z = 1$. Together with the regularity properties this means that γ has finite length. On γ we also have

$$|F_n(u)| \leqslant \exp[-Kn\,|u + 1|^{-2}] \qquad (K = \text{constant depending on } q).$$

We now observe that the function $x^{-1}\exp[-Knx^{-2}]$ is increasing for $0 < x \leqslant (2Kn)^{1/2}$. For $n \geqslant -[2G^*(q)]^{-1}$ and $|u + 1| \leqslant (-K/G^*(q))^{1/2}$ we consequently get

$$|F_n(u)| \cdot |u + 1|^{-1} \leqslant (-G^*(q)/K)^{1/2} \exp[nG^*(q)].$$

On the other hand, using (5) for $|u+1| \geqslant (-K/G^*(q))^{1/2}$ we obtain

$$|F_n(u)\, u^{-n}| \cdot |u+1|^{-1} \leqslant (-G^*(q)/K)^{1/2} \exp[nG^*(q)].$$

In (2) this implies that

$$E_n(\tilde{f}_n) \leqslant K_q \exp[nG^*(q)].$$

Since $\rho_n(Q) \leqslant KE_n(\tilde{f}_n)$ we consequently get

$$\overline{\lim}\, \rho_n(Q)^{1/n} \leqslant \exp G^*(q).$$

6. The Lower Estimate when $q_n = q$

Returning to (1) in Section 3, by the Hahn–Banach theorem we obtain

$$\rho_n(Q) = \sup \left| \int f_n(t)\, d\mu(t) \right|,$$

where the sup is taken over all measures $d\mu$ on $[-1, 1]$ such that

(i) $\int t^k\, d\mu = 0,\ k = 0,\dots, n,$

(ii) $\|d\mu\| \leqslant 1.$

For $m = 2, 3$ we let $d\mu_m$ be given by

$$2\pi i \int h\, d\mu_m = \int_C h \circ \psi(u) \cdot u^{-n-m}\, du,$$

for all functions h continuous on $[-1, 1]$. As before $\psi(u) = (u + u^{-1})/2$ and C: the complex unit circle. Conditions (i) and (ii) are satisfied for $d\mu_m$. Hence for $m = 2, 3$

$$2\pi \rho_n(Q) \geqslant \left| \int_C F_n(u) \cdot u^{-n-m}\, du \right| = \left| \int_\gamma F_n(u)\, u^{-n-m}\, du \right|, \tag{6}$$

where γ is the optimal curve from Section 5, oriented to match the orientation of C.

With the same transformations as those in the previous sections we have

$$F_n(u) \cdot u^{-n} = \exp[nH(z)].$$

In the z-plane we have chosen the optimal curve σ so that close to a critical point σ is an orthogonal trajectory of the level curves of $\operatorname{Re} H(z)$. This means that $\operatorname{Im} H(z)$ is constant on σ close to a critical point. Let α be this constant at the critical point q^* in the upper half-plane.

Pick $\varepsilon > 0$ small. Divide γ into the three parts

$$\gamma_1 = \{u \in \gamma: G(z, q) \leqslant G^*(q) - \varepsilon\},$$
$$\gamma_2 = \text{the part of } \gamma \backslash \gamma_1, \text{ in } \operatorname{Im} u \geqslant 0,$$
$$\gamma_3 = \text{the part of } \gamma \backslash \gamma_1, \text{ in } \operatorname{Im} u \leqslant 0.$$

Let

$$I_1 = \int_{\gamma_1} F_n(u)\, u^{-n-m}\, du;$$

then it holds that

$$|I_1| \leqslant K \exp n[G^*(q) - \varepsilon]. \tag{7}$$

Furthermore we notice that because of the symmetry

$$\overline{\int_{\gamma_2} F_n(u)\, u^{-n-m} du} = \int_{\gamma_2} F_n(\bar{u})(\bar{u})^{-n-m}\, d\bar{u} = -\int_{\gamma_3} F_n(u)\, u^{-n-m}\, du.$$

Consequently with

$$I_2 = \int_{\gamma_2} + \int_{\gamma_3} F_n(u)\, u^{-n-m}\, du$$

we have

$$I_2 = 2i \operatorname{Im} \int_{\gamma_2} F_n(u)\, u^{-n-m}\, du.$$

For ε sufficiently small we know that

$$F_n(u) u^{-n} = |F_n(u) u^{-n}| \exp(in\alpha).$$

Let $\beta(u) = \arg u$ and $du = |du| \cdot \exp(i\delta(u))$ on γ. Then

$$\operatorname{Im} \int_{\gamma_2} F(u)\, u^{-n-m}\, du = \int_{\gamma_2} \sin[n\alpha - m\beta(u) + \delta(u)] \cdot |F_n(u) u^{-n-m}|\, |du|.$$

There exists a constant $k > 0$ such that at least for one of the two possible values of m it holds that

$$|\sin[n\alpha - m\beta(u) + \delta(u)]| \geqslant k > 0$$

for u close to $u(q^*)$. This follows since $\beta(u) \neq 0 \pmod{\pi}$ if u is not real, which we can assume if $u(q^*)$ is not real. On the other hand, if $u(q^*)$ is real then $\alpha = 0 \pmod{\pi}$, $\beta(u) \to 0 \pmod{\pi}$ and $\delta(u) \to \pi/2$ (if $q > 3\sqrt{3}/2$) or $2\pi/3$ (if $q = 3\sqrt{3}/2$) as $u \to u(q^*)$ on γ_2.

Since $\log |F_n(u) u^{-n}|$ is differentiable at $u(q^*)$, by standard saddle point methods (see e.g., de Bruijn [2]) we get

$$\max_{m=2,3} |I_2|^{1/n} \to \exp G^*(q), \qquad n \to \infty. \tag{8}$$

Returning to (6) we have

$$(2\pi \rho_n(Q))^{1/n} \geqslant |I_1 + I_2|^{1/n} = |I_2|^{1/n} \, |1 + I_2^{-1} \cdot I_1|^{1/n}.$$

But (7) and (8) yield $I_2^{-1} \cdot I_1 \to 0$ and consequently

$$\underline{\lim}\, \rho_n(Q)^{1/n} \geqslant \underline{\lim}\, |I_2|^{1/n} = \exp G^*(q).$$

This concludes the proof of the theorem in case $q_n = q$ for all n.

7. The General Case

The simplest generalization of the proof in the previous sections is to the case when $q_n \to q \in (0, \infty)$. Continuity arguments in the previous proof immediately gives

$$\lim \rho_n(Q)^{1/n} = \exp G^*(q),$$

at least for $q \neq 3\sqrt{3}/2$. In case $q = 3\sqrt{3}/2$ the lower estimate is not as obvious.

The case $q_n \to \infty$ can be handled in the same manner. Since then $q_n^* \sim 1/q_n \to 0$ and $G^*(q_n) \to 0$ the upper estimate is trivial. The lower estimate follows as in Section 6, since the optimal curves σ_n will approach the imaginary axis as $n \to \infty$.

The cases $q_n \to 0$ and $q_n \to 3\sqrt{3}/2$ can be handled by a different method. The approach is the same for the two cases so we just concentrate on $q_n \to 0$. We observe that $q_n^* \sim q_n^{-1/3} \cdot (1 + i\sqrt{3})/2$. This gives $G^*(q_n) \to 0$ so the upper estimate is trivial. It remains to prove that

$$\underline{\lim}\, \rho_n(Q)^{1/n} = 1.$$

To get a contradiction we assume that

$$\underline{\lim}\, \rho_n(Q)^{1/n} \leqslant \exp(-2a), \qquad \text{for some } a > 0.$$

Then for a subsequence N' of the natural numbers we can find $r \in R_n(Q)$ such that

$$|e^{-x} - r(x)| \leqslant e^{-an}, \qquad \text{for all } x \geqslant 0.$$

For every $0 < \varepsilon < a$ we get

$$|e^{-(x-\varepsilon n)} - r(x)\, e^{\varepsilon n}| \leqslant e^{-(a-\varepsilon)n}, \qquad x \geqslant 0.$$

Substituting $t = x - \varepsilon n$ we find

$$|e^{-t} - r^*(t)| \leqslant e^{-(a-\varepsilon)n}, \qquad t \geqslant -\varepsilon n,$$

where $r^*(t) = r(t + n\varepsilon) \exp n\varepsilon$. Letting $Q_\varepsilon = (\varepsilon + q_n)_1^\infty$, we therefore get

$$\underline{\lim}\, \rho_n(Q_\varepsilon)^{1/n} \leqslant \exp(\varepsilon - a).$$

But since $\varepsilon + q_n \to \varepsilon$ we already know that

$$\underline{\lim}\, \rho_n(Q_\varepsilon)^{1/n} = \exp G^*(\varepsilon).$$

Hence for all ε: $0 < \varepsilon < a$

$$G^*(\varepsilon) \leqslant \varepsilon - a,$$

which is impossible since $G^*(\varepsilon) \to 0$ as $\varepsilon \to 0$. Thus

$$\underline{\lim}\, \rho_n(Q)^{1/n} = 1.$$

The general situation when q_n does not tend to a limit now obviously follows by looking at suitable subsequences. So the theorem is proved.

8. A Good Approximant

Finally we further study the optimal situation $q = 1/\sqrt{2}$. In the approximation of $\tilde{f}_n$ in Section 5 we used the Taylor polynomial $\sum_0^n a_k w^k$ of order n. Since we then used the mapping $\tilde{f} \to f$ to get an approximation of f_n it is natural to analyze this mapping somewhat further. For $|w| < 1$ we have

$$\tilde{f}(w) = \frac{1}{2\pi i} \int_C \frac{f \circ \psi(u)}{u - w}\, du = \sum_0^\infty \frac{w^k}{2\pi i} \int_C f \circ \psi(u)\, u^{-k-1}\, du.$$

By definition $\psi(u) = (u + u^{-1})/2$ so we substitute $u = \exp(iv)$ and find

$$\int_C f \circ \psi(u)\, u^{-k-1}\, du = 2i \int_0^\pi f(\cos v) \cos kv$$

$$= 2i \int_{-1}^{1} f(x)\, T_k(x)(1 - x^2)^{-1/2}\, dx,$$

where $T_k(x) = \cos(n \arccos x)$, i.e., the kth Chebyshev polynomial. Hence

$$\tilde{f}(w) = b_0 + 2^{-1} \sum_1^\infty b_k w^k,$$

where b_k is the kth Fourier–Chebyshev coefficient of f. From this fact we can deduce that $2\tilde{T}_k(w) = w^k$, $k > 0$, and $\tilde{T}_0 = 1$. Thus the Taylor series of $\tilde{f}$ corresponds to the Fourier–Chebyshev series of f.

The function $r_n(x)$ that we can use to approximate e^{-x} is then

$$r_n(x) = \sum_0^n b_k T_k \left(\frac{n - x\sqrt{2}}{n + x\sqrt{2}} \right),$$

where b_k is the kth Fourier–Chebyshev coefficient of

$$f_n(t) = \exp \frac{n(t-1)}{(t+1)\sqrt{2}}.$$

From Section 5 we can also get a better estimate than that in our theorem, namely,

$$|e^{-x} - r_n(x)| \leqslant K(\sqrt{2} - 1)^n, \qquad x \geqslant 0,$$

for some constant K.

References

1. J.-E. Andersson and T. H. Ganelius, The degree of approximation by rational functions with fixed poles, *Math. Z.* **153** (1977), 161–166.
2. N. G. de Bruijn, "Asymptotic Methods in Analysis," North-Holland, Amsterdam, 1958.
3. E. H. Kaufman, Jr., and G. D. Taylor, Uniform approximation with rational functions having negative poles, *J. Approx. Theory* **23** (1978), 364–378.
4. E. B. Saff, A. Schönhage, and R. S. Varga, Geometric convergence to e^{-z} by rational functions with real poles, *Numer. Math.* **25** (1976), 307–322.