The Tanh Rule for Numerical Integration

Seymour Haber


Stable URL: http://links.jstor.org/sici? Sidney=0036-1429%28197709%2914%3A4%3C668%3ATTFRNI%3E2.0.CO%3B2-2

*SIAM Journal on Numerical Analysis* is currently published by Society for Industrial and Applied Mathematics.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/siam.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.
THE TANH RULE FOR NUMERICAL INTEGRATION*

SEYMOUR HABER†

Abstract. The tanh rule for numerical integration is analyzed in the context of the Hardy space $H^2$. The optimal parameter choice is determined, and it is shown that the norm of the error functional is asymptotic to $C \exp\left(-(\pi/2)/\sqrt{M}\right)$, where $M$ is the number of points used and $C$ is a certain constant. The result is related to recent theorems on approximation of piecewise analytic functions by rational functions.

1. Introduction. It has been known for some time that the trapezoid rule for the interval $(-\infty, \infty)$,

$$
\int_{-\infty}^{\infty} f(x) \, dx \approx h \sum_{n=-\infty}^{\infty} f(nh),
$$

often gives remarkably good results; that is, the expression on the right-hand side of (1) converges very rapidly to the integral as $h$ decreases toward zero, and the infinite sum itself is often rapidly convergent. E. T. Goodwin [1] analyzed this phenomenon for certain integrands, in 1949. In 1968 C. Schwartz [2] suggested using the phenomenon for evaluating integrals over finite intervals, by a change of variables. Choosing any function $\psi$ that maps $(-\infty, \infty)$ onto $(-1, 1)$ monotonically, Schwartz suggested writing

$$
\int_{-1}^{1} g(x) \, dx = \int_{-\infty}^{\infty} g(\psi(t))\psi'(t) \, dt
$$

$$
= h \sum_{n=-N}^{N} \psi'(nh)g(\psi(nh))
$$

and using the last sum, for appropriate $h$ and $N$, to approximate the original integral. One of the specific functions Schwartz suggested was

$$
\psi(t) = \tanh \frac{t}{2},
$$

and one choice he suggested for the step $h$ was

$$
h = \pi \sqrt{\frac{2}{N}}.
$$

The combination of (2) and (3) has become known as the “tanh rule”, and has been studied by H. Takahasi and M. Mori [3] and by F. Stenger [4]. From the classical point of view it is a strange quadrature method, as it is not exact for any polynomials—not even constants—for any positive $h$ and finite $N$. Nevertheless it is remarkably accurate. Stenger, and Takahasi and Mori, showed that if the integrand $g$ is analytic in the unit disk and does not grow too quickly as $z$

---

* Received by the editors April 23, 1976, and in revised form August 31, 1976.
approaches the unit circle, then the quadrature error is

\[ O(e^{-CN}) \]

if \( h \) is specified by (4); \( C \) is some constant related to the growth properties of the integrand.

This result is intriguing, not only because the convergence speed is faster than \( N^{-k} \) for any \( k \), but also because it suggests a possible connection with D. J. Newman's fundamental result [5] on the rational approximation of \( |x| \), where an \( e^{-CN} \) type of convergence was found.

The present paper arose from an attempt to determine whether the convergence of the tanh rule is indeed of \( e^{-CN} \) type—the above quoted papers gave such expressions mainly as upper bounds. (For integrands analytic on the closed interval \((-1, 1)\) it is well known that the Gauss–Legendre formula provides a faster rate of convergence—of \( e^{-CN} \) type.) Other questions were also in mind: is (4) the best prescription for \( h \), in some sense; and is (3) the best choice of \( \psi \). As we shall see, it turns out that a slightly different choice of \( h \) is optimal—in a sense to be defined below—for the tanh rule, and with the optimal \( h \) the rule’s convergence is indeed of \( e^{-CN} \) type. The consideration of alternative \( \psi \)'s—several have been suggested by Takahasi and Mori—is something I hope yet to do.

2. The spaces \( H^2 \) and \( C^2 \). A convenient context for the analysis of the tanh rule is the Hardy space \( H^2 \). This consists of those functions \( f \) which are analytic in the unit disk and whose Taylor coefficients satisfy the condition

\[ \sum_{n=0}^{\infty} |a_n|^2 < \infty \]

(where \( f(z) = \sum_{n=0}^{\infty} a_n z^n \)). Included are some functions, such as \((1 - z^2)^{-a}, \ a < \frac{1}{2}\), which are unbounded as \( z \to \pm 1 \). As is well known (see, e.g. [6, pp. 17 and 21]) functions in \( H^2 \) have nontangential limits almost everywhere on the unit circle and belong to \( L^2 \) on the unit circle. With the inner product

\[ (f, g) = \frac{1}{2\pi} \int_{|z|=1} f(z)\overline{g(z)}|dz|, \]

\( H^2 \) is a Hilbert space, and \([1, z, z^2, \cdots] \) is an orthonormal basis. \( H^2 \) has the reproducing kernel

\[ K(z, w) = \sum_{n=0}^{\infty} z^n\overline{w}^n = 1/(1 - z\overline{w}). \]

An exposition of the basic properties of reproducing kernels in Hilbert spaces may be found in [7, pp. 316–320].

From the existence of the reproducing kernel it follows (Thm. 12.6.1 of [7]) that the "point functionals" \( P_z \), defined by

\[ P_z f = f(z) \]

are bounded linear functionals whenever \( |z| < 1 \); and in fact

\[ |P_z f|^2 \leq K(z, z)\|f\|^2. \]
The integral over \((-1, 1)\) also defines a bounded linear functional (see, e.g. [8, p. 417]). Let us set

\[ I_f = \int_{-1}^{1} f, \]

\[ T_{h,N} f = h \sum_{n=-N}^{N} \frac{1}{2 \cosh^2(nh/2)} f(\tanh \frac{nh}{2}), \]

\[ T_h f = \lim_{N \to \infty} T_{h,N} f, \]

\[ E_{h,N} = I - T_{h,N}, \]

\[ E_h = I - T_h; \]

then \(I, T_{h,N},\) and \(E_{h,N}\) are bounded linear functionals on \(H^2\) (and, as we shall see below, so are \(T_h\) and \(E_h\)).

We can now define the idea of a “best” step-size \(h\): For any \(N, h_\ast = h_\ast(N)\) is “best” if

\[ \|E_{h_\ast,N}\| = \inf_{h > 0} \|E_{h,N}\|. \]

This is a natural definition because of the sharp inequality

\[ |I - T_{h,N} f| \leq \|E_{h,N}\| \cdot \|f\|. \]

The change of variables (3) means that we shall effectively be dealing with functions that are defined on \((-\infty, \infty)\); and it is convenient to use the same change of variables to define a space of such functions, corresponding to \(H^2\). Let us first denote by “\(S_\ast\)” the infinite strip \(\{z | -a < \text{Im}(z) < a\}\). Now with \(\psi(w) = \tanh(w/2)\) as in (3), we let \(G^2\) be the set of all functions of the form

\[ \psi(w) f(\psi(w)), \quad f \in H^2. \]

Since \(\psi\) maps \(S_{\pi/2}\) onto the unit disk, the functions in \(G^2\) are analytic on \(S_{\pi/2}\) and may be regarded as defined almost everywhere on \(\partial S_{\pi/2}\). Let us use “\(^*\)” to denote the transition from \(H^2\) to \(G^2\), so that we will, for example, set \(\hat{f}(w) = \psi(\psi(w)) f(\psi(w))\), for each \(f \in H^2\). Since

\[ (f, g)_{H^2} = \frac{1}{2\pi} \int_{\{|z| = 1\}} f(z) \overline{g(z)} \, dz \]

(6)

\[ = \frac{1}{2\pi} \int_{\partial S_{\pi/2}} f(\psi(w)) \overline{g(\psi(w))} |\psi'(w)| \, dw \]

\[ = \frac{1}{2\pi} \int_{\partial S_{\pi/2}} \hat{f}(w) \overline{\hat{g}(w)} \frac{dw}{|\psi'(w)|}, \]
we may use the last expression to define an inner product in \( G^2 \). Thus \( G^2 \) is a Hilbert space, and the mapping \( f \rightarrow \hat{f} \) an isomorphism of \( H^2 \) onto \( G^2 \). Therefore the functions

\[
\phi_n(z) = \psi'(z)(\psi(z))^n, \quad n = 0, 1, 2, \ldots,
\]

constitute a basis for \( G^2 \), and \( G^2 \) has the reproducing kernel

\[
\hat{K}(z, w) = \frac{1}{4 \cosh (z/2) \cosh (\bar{w}/2) \cosh (z - \bar{w})/2}.
\]

An alternative, intrinsic, description of \( G^2 \) could be given: it is the Hilbert space consisting of all functions \( \hat{f} \) such that

1. \( \hat{f} \) is analytic on \( S_{\pi/2} \).
2. There is a constant \( C(\hat{f}) \) such that

\[
\int_{-\infty}^{\infty} \left\{ \left| \hat{f}(x + i\left(\frac{\pi}{2} - \epsilon\right)) \right|^2 + \left| \hat{f}(x - i\left(\frac{\pi}{2} - \epsilon\right)) \right|^2 \right\} e^{\epsilon|x|/2} dx < C(\hat{f})
\]

whenever \( 0 < \epsilon < \pi/2 \);

with the last expression in (6) as inner product.

Since

\[
\int_{-1}^{1} f(x) \, dx = \int_{-\infty}^{\infty} \hat{f}(u) \, du
\]

and

\[
T_{h,N} f = h \sum_{n=-N}^{N} \hat{f}(nh)
\]

we shall denote the quantities on the right sides of these two equations by \( \hat{f} \) and \( \hat{T}_{h,N} \hat{f} \) respectively. We similarly define \( \hat{T}_h, \hat{E}_{h,N} \) and \( \hat{E}_h \).

The inequality (5), transferred to \( G^2 \), tells us that

\[
|\hat{f}(nh)| \leq (\hat{K}(nh, nh))^{1/2} |\hat{f}|
\]

for each \( \hat{f} \) in \( G^2 \). Therefore

\[
|\hat{T}_h \hat{f}| \leq h |\hat{f}| \sum_{n=-\infty}^{\infty} (K(nh, nh))^{1/2}
\]

\[
\leq h |\hat{f}| \sum_{n=-\infty}^{\infty} \left( 2 \cosh \frac{nh}{2} \right)^{-1}
\]

\[
\leq h |\hat{f}| \sum_{n=-\infty}^{\infty} e^{-h|n|/2} = h \coth \frac{h}{4} |\hat{f}|.
\]

This implies:

**Lemma 1.** \( T_h, E_h, \hat{T}_h \) and \( \hat{E}_h \) are bounded linear functionals.

It should be noted that all the functionals we are dealing with are real—that is they satisfy the condition

\[
L \hat{f} = \overline{L f}.
\]
It follows from the properties of the reproducing kernel (12.6.7 and 12.6.12 of [7]) that if \( L \) is any of the functionals on \( G^2 \) that we are considering,

\[
\|L\|^2 = L(z)L(w)\hat{K}(z, w).
\]

(The parenthesized subscripts, here and below, indicate the variable with respect to which the operator is acting; that is, \( L(w) \) is acting on \( K \) regarded as a function of \( w \).) Our purpose is to calculate the norm of \( E_{h,N} \); by the correspondence between \( H^2 \) and \( G^2 \), that is the same as the norm of \( \hat{E}_{h,N} \) which we will calculate from (9).

3. The infinite trapezoidal rule. We will first detour, however, for a discussion of the optimality of the coefficients of the "infinite trapezoidal rule" \( \hat{T}_h \). The discussion will throw some light on our analytical situation, and yield by-products which will be needed in the calculation of \( \|E_{h,N}\| \).

Let \( A = \{ \cdots, a_{-2}, a_{-1}, a_0, a_1, \cdots \} \) be a sequence of complex numbers, infinite in both directions. For any \( \hat{f} \in G^2 \), we set

\[
\hat{T}_{h:A}\hat{f} = \sum_{n=-\infty}^{\infty} a_n\hat{f}(nh); \quad \hat{E}_{h:A} = \hat{I} - \hat{T}_{h:A}.
\]

We will restrict attention to those \( A \)'s for which \( \hat{T}_{h:A} \) is a bounded linear functional on \( G^2 \). Is \( \|\hat{E}_{h:A}\| \) minimized by choosing \( a_0 = h \)?

To see whether this is so, we first note that \( \hat{T}_{h:A} \) can be written as a linear combination of point functionals:

\[
\hat{T}_{h:A} = \sum_{n=-\infty}^{\infty} a_nP_{nh}.
\]

If \( S_0 \) is the closed subspace of \( G^{2*} \) spanned by the \( P_{nh} \), then \( \hat{T}_{h:A} \) is any element of \( S_0 \), and we wish to determine whether \( \hat{T}_h \) is that element which is closest to \( \hat{I} \). That is the case if and only if \( \hat{I} - \hat{T}_h \) is orthogonal to every one of the \( P_{nh} \).

Let \( F_{nh} \) be the representor of \( P_{nh} \)—the element of \( G^2 \) for which

\[
P_{nh}\hat{f} = (f, F_{nh})
\]

for every \( \hat{f} \). Then

\[
F_{nh} = P_{nh}(w)\hat{K}(z, w) = \frac{1}{4 \cosh \frac{z}{2} \cosh \frac{nh}{2} \cosh \frac{z - nh}{2}}
\]

([7, 12.6.6 and 12.6.10]). If \( J \) is the representor of \( \hat{I} - \hat{T}_h \), then (using "\( (\cdot, \cdot) \)" for the inner product both in \( G^2 \) and in \( G^{2*} \)),

\[
(P_{nh}, \hat{I} - \hat{T}_h) = (F_{nh}, J) = \hat{I}F_{nh} - \hat{T}_hF_{nh}.
\]

Therefore the coefficients of \( \hat{T}_h \) are optimal if and only if \( \hat{T}_h \) integrates \( F_{nh} \) exactly, for all integers \( n \). Now

\[
\hat{I}F_{nh} = \frac{1}{2 \cosh \frac{nh}{2}} \int_{-\infty}^{\infty} dx \frac{dx}{2 \cosh \frac{x}{2} \cosh \frac{x - nh}{2}}
\]

\[
= \frac{1}{2 \cosh \frac{nh}{2}} \int_{-\infty}^{\infty} dx \frac{nh}{2 + \cosh \frac{x - nh}{2}}.
\]
Using the identity

\[
\frac{\sinh a}{\cosh a + \cosh b} = \frac{1}{1 + e^{b-a}} - \frac{1}{1 + e^{b+a}},
\]

we see that the last integral is equal to \(nh/\sinh (nh/2)\), and therefore

\[
\hat{F}_{nh} = \frac{nh}{\sinh (nh)}
\]

for \(n \neq 0\); and \(\hat{F}_0 = 1\), since both sides of (11) are continuous functions of \(n\).

Using the same manipulations, we see that for \(n \neq 0\)

\[
\hat{F}_{nh} = \frac{h}{\sinh (nh)} \sum_{r=-\infty}^{\infty} \left( \frac{1}{1 + e^{(r-n)h}} - \frac{1}{1 + e^{rh}} \right).
\]

If we set

\[
G(z) = \sum_{r=-\infty}^{\infty} \left( \frac{1}{1 + e^{(r-z)h}} - \frac{1}{1 + e^{rh}} \right),
\]

then \(G(0) = 0\) and

\[
G(z + 1) = \lim_{M,N \to \infty} \sum_{r=-M}^{N} \left( \frac{1}{1 + e^{(r-z-1)h}} - \frac{1}{1 + e^{rh}} \right)
\]

\[
= \lim_{M,N \to \infty} \left\{ \sum_{r=-M}^{N} \left( \frac{1}{1 + e^{(r-z)h}} - \frac{1}{1 + e^{rh}} \right) + \frac{1}{1 + e^{(-M-z-1)h}} - \frac{1}{1 + e^{(N-z)h}} \right\}
\]

\[
= G(z) + 1.
\]

Thus

\[
G(n) = n, \quad \text{and} \quad \hat{F}_{nh} = \hat{F}_{nh}, \quad n = \pm 1, \pm 2, \ldots.
\]

For \(n = 0\), we use the Poisson sum formula (see e.g., [9, p. 37]; \(F_0\) clearly satisfies the conditions for the validity of the formula) to obtain:

\[
\hat{F}_0 - \hat{F}_0 = -\sum_{r=1}^{\infty} \int_{-\infty}^{\infty} \frac{\cos (2\pi rx/h)}{1 + \cosh x} \, dx
\]

\[
= -2 \sum_{r=1}^{\infty} \frac{2\pi^2 r/h}{\sinh (2\pi^2 r/h)}
\]

([10, p. 137, formula 14c]). The last expression is, for small \(h\), asymptotic to

\[
-\frac{8\pi^2}{h} e^{-2\pi^2/h}.
\]

This is not zero, though it is quite small for \(h\) less than, say, 1. (The values of \(h\) used in practical application of the tanh rule generally range from \(\frac{1}{4}\) to \(\frac{1}{2}\).)

So the coefficients of the infinite trapezoid rule are in fact not optimal in our sense. This tends to call our analytical framework into question as much as it casts
doubt on the coefficients. One feels intuitively that those coefficients are right. But in fact the quantity (16) is very small relative to the norm of \( \hat{T}_h \), which is of the order of magnitude of \( \exp(-\pi^2/2h) \), as we shall see below. Thus while \( \hat{T}_h - \hat{T}_h \) is not orthogonal to \( P_0 \), the angle between them differs from \( \pi/2 \) by only \( O(\exp(-\pi^2/2h)) \), and the change in the coefficients of \( \hat{T}_h \) needed to make them optimal in our sense is also that small. Such a change would lower the error norm by a quantity which is \( O(||E_h||^2) \), and that is of no significance. The “artificiality” in our analysis which causes this small annoyance is just the limitation of the analyticity of the functions considered to the strip \( S_{\pi/2} \), but this strip is naturally related to the unit disk via the tanh transformation.

The function \( G \) deserves some further comment. If we differentiate (13) term by term we find that

\[
G'(z) = \frac{h}{2} \sum_{r=-\infty}^{\infty} \frac{1}{1 + \cosh (r - z)/h}.
\]

It is immediate that \( G' \) is periodic with periods 1 and \( 2\pi i/h \). Its poles are just the poles of its summands; and in its period parallelogram there is just one, a pole of order 2 at \( z = \pi i/h \). Thus \( G' \) is an elliptic function of the second order, and in fact

\[
-hG'(z) = \mathcal{P} \left( \frac{z - \pi i}{h} \right) \left( \frac{\pi i}{2h} \right) + C
\]

where \( \mathcal{P} \) is the Weierstrass \( \mathcal{P} \)-function and \( C \) depends only on \( h \). It follows from elementary properties of elliptic functions that the only values of \( z \) for which \( G(z) = z \) are the integers \( n \) and the half-integers.

4. The error norm. Returning to the calculation of \( ||E_{h,N}|| \), we will express the main quantities involved as contour integrals, using the following standard result.

**Lemma 2.** If \( \alpha \) and \( h \) are positive real numbers and \( f \) is a function satisfying the following conditions:

1. \( f \) is analytic on the closure of the strip \( S_{\alpha} \),
2. \( \lim_{\alpha \to \pm \infty} \int_{-\alpha}^{\alpha} |f(z)| \, dz = 0 \),
3. \( \int_{-\infty}^{\infty} f(x) \, dx \) and \( \sum_{n=-\infty}^{\infty} f(nh) \) exist,

then

\[
\int_{-\infty}^{\infty} f(x) \, dx = \frac{1}{2\pi i} \int_{S_{\alpha}} f(z) \Psi(z) \, dz
\]

and, if \( \alpha' \leq \min \{\alpha, \alpha/h\} \),

\[
h \sum_{n=-\infty}^{\infty} f(nh) = \frac{1}{2\pi i} \int_{S_{\alpha}} f(z) \cdot \pi \cot \frac{\pi z}{h} \, dz.
\]

Here \( \Psi(z) \) is defined, for all \( z \) off the real axis, by:

\[
\Psi(z) = \begin{cases} 
-\pi i, & \text{Im}(z) > 0, \\
\pi i, & \text{Im}(z) < 0.
\end{cases}
\]
Set $\hat{\sigma}_{h,N} = \hat{T}_h - \hat{T}_{h,N}$. By (9),

$$\|\hat{E}_{h,N}\|^2 = \|\hat{E}_h + \hat{\sigma}_{h,N}\|^2$$

$$= \hat{E}_{h,(z)} \hat{E}_{h,(w)} \hat{K}(z, w) + \hat{\sigma}_{h,N,(z)} \hat{\sigma}_{h,N,(w)} \hat{K}(z, w)$$

$$+ \hat{\sigma}_{h,N,(z)} \hat{E}_{h,(w)} \hat{K}(z, w) + \hat{\sigma}_{h,N,(z)} \hat{\sigma}_{h,N,(w)} \hat{K}(z, w).$$

The second and third of the last 4 terms are easily seen to be equal; and since

$$\hat{\sigma}_{h,N,(w)} \hat{K}(z, w) = h \sum_{|n| > N} \hat{K}(z, nh)$$

$$= h \sum_{|n| > N} F_{nh}(z),$$

the second term is just

$$h \sum_{|n| > N} \hat{E}_h F_{nh}$$

which is zero, by (14).

To evaluate the first term on the right in (17), we note that we may replace $\hat{K}(z, w)$ in it by $\hat{K}(z, \bar{w})$ without changing the value of the term; and that $\hat{K}(z, \bar{w})$, regarded as a function of $w$, satisfies the conditions of Lemma 2 if we take $\alpha' = \alpha < \min \{\pi/2, \pi/(2h)\}$. So

$$\hat{E}_h \hat{K}(z, w) = \hat{E}_h \hat{K}(z, \bar{w})$$

$$= \frac{1}{2\pi i} \int_{S_\alpha} \frac{\Phi(w)}{4 \cosh \frac{z}{2} \cosh \frac{w}{2} \cosh \frac{z - w}{2}} \, dw,$$ 

where

$$\Phi(z) = \Phi(z; h) = \Psi(z) - \pi \cot \frac{\pi z}{h}$$

$$= \begin{cases} 
\frac{-2\pi i}{1 - \exp(-2\pi iz/h)}, & \text{Im}(z) > 0, \\
\frac{2\pi i}{1 - \exp(2\pi iz/h)}, & \text{Im}(z) < 0.
\end{cases}$$

We initially take $\alpha$ sufficiently small so that $S_\alpha$ contains no poles of

$$\hat{K}(z, \bar{w}) = \frac{1}{4 \cosh \frac{z}{2} \cosh \frac{w}{2} \cosh \frac{z - w}{2}}$$

the latter being regarded as a function of $w$. If we specify that $z$ is not any one of the numbers $(2n + 1)\pi i$, $n$ an integer, then those poles are:

- a first order pole at each point $(2n + 1)\pi i$, $n$ an integer, with residue $(\sinh z)^{-1}$
- and

- a first order pole at each point $z + (2n + 1)\pi i$, $n$ an integer, with residue $-(\sinh z)^{-1}$. 


If, for the integral in (18), we let $\alpha$ grow to infinity through values which keep $\partial S_{\alpha}$ away from these poles, we pick up residues at the poles; while the integral itself goes to zero because $|\Phi(x + iy)| \sim 2\pi \exp (-2\pi |y|/h)$. We thus have

$$
\hat{E}_{h,w}(z, w) = \frac{-1}{\sinh z} \sum_{n=-\infty}^{\infty} \Phi((2n + 1)\pi i) + \frac{1}{\sinh z} \sum_{n=-\infty}^{\infty} \Phi(z + (2n + 1)\pi i).
$$

But $\Phi(-z) = -\Phi(z)$, and so the first sum is zero. The second sum is analytic in $S_{mr}$ and periodic with period $h$ since $\Phi$ has period $h$. It follows that

$$
H(z) = \hat{E}_{h,w}(z, w)
$$

satisfies the conditions for applicability of the Poisson sum formula, and so

$$
\|\hat{E}_h\|^2 = \hat{E}_h H = -2 \sum_{r=1}^{\infty} \int_{-\infty}^{\infty} H(x) \cos (2\pi rx/h) \, dx.
$$

Set

$$
H_n(x) = \Phi(x + (2n - 1)\pi i) + \Phi(x - (2n - 1)\pi i),
$$

so that

$$
H(x) = \sum_{n=1}^{\infty} \frac{H_n(x)}{\sinh x}.
$$

For $n > 1$, we integrate twice by parts to obtain

$$
\int_{-\infty}^{\infty} \frac{H_n(x)}{\sinh x} \cos (2\pi rx/h) \, dx = -\left(\frac{h}{2\pi r}\right)^2 \int_{-\infty}^{\infty} \frac{H_n(x)}{\sinh x} \cos (2\pi rx/h) \, dx.
$$

Now

$$
\frac{H_n(x)}{\sinh x} = \exp \left(- (4n - 2)\pi^2 / h \right) \cdot \frac{\exp ((4n - 2)\pi^2 / h) H_n(x)}{\sin (2\pi x/h)} \cdot \frac{\sin (2\pi x/h)}{\sinh x},
$$

and the middle factor on the right is equal to

$$
\frac{-4\pi}{1 - 2 \exp \left(- (4n - 2)\pi^2 / h \right) \cos (2\pi x/h) + \exp \left(- (8n - 4)\pi^2 / h \right)}.
$$

It is easy to show that the $i$th derivative of this middle factor is less than $Ah^{-i}$ for $i = 0, 1, 2$; for each $i$, $A$ is some absolute constant. Similarly, the $i$th derivative of the third factor is less than $Ah^{-i} e^{-|x|}$, $i = 0, 1, 2$. It follows that

$$
\frac{H_n(x)}{\sinh x}'' < Ah^{-2} e^{-|x|}
$$

for some constant $A$; and so, because of (21),

$$
\left| \int_{-\infty}^{\infty} \frac{H_n(x)}{\sinh x} \cos (2\pi rx/h) \, dx \right| \leq \frac{A}{r^2} \exp \left(- (4n - 2)\pi^2 / h \right)
$$
for some \( A \). Therefore

\[
\left| \sum_{r=1}^{\infty} \int_{-\infty}^{\infty} \sum_{n=2}^{\infty} \frac{H_n(x)}{\sinh x} \cos \left( \frac{2\pi rx}{h} \right) \, dx \right| = O(e^{-6\pi^2/h}).
\]

We now evaluate

\[
\int_{-\infty}^{\infty} \frac{\Phi(x + \pi i) - \Phi(x - \pi i)}{\sinh x} \cos \left( \frac{2\pi rx}{h} \right) \, dx
\]

\[
= -4\pi e^{-\pi^2/h} \int_{-\infty}^{\infty} \frac{\sin \left( \frac{2\pi x}{h} \right) \cos \left( \frac{2\pi rx}{h} \right)}{\sinh x} L(x) \, dx,
\]

where

\[
L(x) = \frac{1}{1 - 2 \exp \left( -2\pi^2/h \right) \cos \left( \frac{2\pi x}{h} \right) + \exp \left( -4\pi^2/h \right)}.
\]

By an argument similar to the one just used, we can show that our quantity is equal to

\[
-4\pi e^{-\pi^2/h} \int_{-\infty}^{\infty} \frac{\sin \left( \frac{2\pi x}{h} \right) \cos \left( \frac{2\pi rx}{h} \right)}{\sinh x} \, dx + O(r^{-2} e^{-4\pi^2/h}).
\]

the "\( O \)" being uniform in \( r \). The numerator of the integrand in (23) can be written as a sum of two sines, and each resulting integral may be evaluated from the formula

\[
\int_{-\infty}^{\infty} \frac{\sin ax}{\sinh x} \, dx = \pi \tanh \left( a\pi/2 \right)
\]

([10, p. 136]). We then obtain, from (23), (22), and (20), the equation

\[
\hat{E}_h H = 4\pi^2 e^{-2\pi^2/h} \sum_{r=1}^{\infty} \left( \tanh \frac{\pi^2(1-r)}{h} + \tanh \frac{\pi^2(1+r)}{h} \right) + O(e^{-4\pi^2/h}).
\]

For \( r > 1 \), the \( r \)th summand in the last sum is equal to

\[
\tanh \frac{\pi^2(r+1)}{h} - \tanh \frac{\pi^2(r-1)}{h},
\]

and that is \( O(\exp \left( -2(r-1)\pi^2/h \right)) \). The 1st summand is just

\[
\tanh \left( 2\pi^2/h \right) = 1 + O(e^{-4\pi^2/h}).
\]

So altogether

\[
\|\hat{E}_h\|^2 = \hat{E}_{h,(z)} \hat{E}_{h,(w)} \hat{K} = 4\pi^2 e^{-2\pi^2/h} + O(e^{-4\pi^2/h}).
\]

Taking square roots, we have

THEOREM 1. For the space \( C^2 \), the error norm of the infinite trapezoid rule with step \( h \) is given by

\[
\|\hat{E}_h\| = 2\pi e^{-\pi^2/h} + O(e^{-3\pi^2/h})
\]

(as \( h \to 0+ \)).
We now must evaluate the last term in (17). That is equal to
\[ \frac{h^2}{4} \sum_{|r|,|n|>N} \left( \cosh \frac{rh}{2} \cosh \frac{nh}{2} \cosh \left( \frac{n-r|h|}{2} \right) \right)^{-1}, \]
which can be written as
\[ \frac{h^2}{2} \sum_{r,n>N} \left( \cosh \frac{rh}{2} \cosh \frac{nh}{2} \cosh \left( \frac{n-r|h|}{2} \right) \right)^{-1} + \frac{h^2}{2} \sum_{r,-n>N} \left( \cosh \frac{rh}{2} \cosh \frac{nh}{2} \cosh \left( \frac{n-r|h|}{2} \right) \right)^{-1}. \]

(25)

The second of the last two sums is equal to
\[ \sum_{r,n>N} \left( \cosh \frac{rh}{2} \cosh \frac{nh}{2} \cosh \left( \frac{n+r|h|}{2} \right) \right)^{-1}, \]
which is less than
\[ 8 \sum_{r,n>N} (e^{rh/2} e^{nh/2} e^{(n+r|h|/2)}^{-1}. \]

(26)

In the situation that interests us, the quantities \( h \) and \( N \) are interrelated, and we shall henceforth suppose that as \( N \) increases, \( h \) decreases towards zero, in such a manner that the product \( Nh \) approaches infinity. The quantity (26) is then
\[ O(e^{-2Nh}). \]

The first sum in (25) can be split into
\[ \sum_{n>N} \cosh^{-2} \left( \frac{nh}{2} \right) + 2 \sum_{n>r>N} \left( \cosh \frac{rh}{2} \cosh \frac{nh}{2} \cosh \left( \frac{n-r|h|}{2} \right) \right)^{-1}. \]

The first of these two sums is just
\[ 4 \sum_{n>N} e^{-nh} (1 + e^{-nh})^{-2} = \{4 + O(e^{-Nh})\} \sum_{n>N} e^{-nh} = \frac{4e^{-Nh}}{e^h - 1} + O(h^{-1} e^{-2Nh}). \]

Similarly, the second is
\[ \{8 + O(e^{-Nh})\} \sum_{n>r>N} e^{-nh/2} e^{-rh/2} e^{(n+r|h|/2) + e^{|r-n|h|/2}.} \]

This last sum is
\[ \sum_{r=N+1}^{\infty} \sum_{n=r+1}^{\infty} (e^{rh} + e^{nh})^{-1} = \sum_{r=N+1}^{\infty} \sum_{m=0}^{\infty} (e^{rh} + e^{(r+1+m|h|)}^{-1} = \sum_{r=N+1}^{\infty} e^{-rh} \sum_{m=0}^{\infty} (1 + e^{(m+1)|h|})^{-1} = \frac{e^{-Nh}}{e^h - 1} \sum_{n=1}^{\infty} \frac{1}{1 + e^{nh}}. \]
If we define the function $\sigma$ by

$$\sigma(z) = \sum_{n=1}^{\infty} \frac{1}{1 + e^{nz}},$$

we can summarize these calculations in

**Lemma 3.** If $h \to 0$ and $N \to \infty$ simultaneously in such a manner that $Nh \to \infty$, then

$$\hat{\sigma}_{h,N}(z) \hat{\sigma}_{h,N}(w) \hat{K}(z, w) = \frac{2h^2 (1 + 4\sigma(h))}{e^h - 1} e^{-Nh} + O(h \sigma(h) e^{-2Nh}).$$

Combining this lemma and Theorem 1, we have:

$$\|\hat{E}_{h,N}\|^2 = \|E_{h,N}\|^2 = 4\pi^2 e^{-2\pi^2/h} + \frac{2h^2 (1 + 4\sigma(h))}{e^h - 1} e^{-Nh}$$

$$+ O(e^{-4\pi^2/h}) + O(h \sigma(h) e^{-2Nh}).$$

(29)

In order to find the best relation between $h$ and $N$, we must investigate the function $\sigma$.

From (28) it is clear that $\sigma$ is defined and analytic for $z$ in the right half-plane. For such $z$ we write

$$\sigma(z) = \sum_{n=1}^{\infty} \sum_{r=1}^{\infty} (-1)^{r-1} e^{-nrz}$$

(30)

$$= \sum_{r=1}^{\infty} (-1)^{r-1} \frac{e^{-rz}}{e^{rz} - 1} = \sum_{r=1}^{\infty} (-1)^{r-1} \frac{e^{-rz}}{1 - e^{-rz}}.$$

If in the last sum we set $e^{-z} = w$, we obtain a Lambert series in $w$, which is easily seen to have the unit circle as natural boundary. (The proof given in [11, pp. 160–161] for Lambert’s series applies to the present series with almost no change.) Thus the imaginary axis is a natural boundary for $\sigma$. We cannot, therefore, estimate $\sigma(h)$, for $h$ small and positive, from any Taylor series, but must look for an asymptotic expression.

Returning to (28), we write

$$\sigma(h) = \frac{1}{h} \sum_{n=1}^{\infty} \frac{h}{1 + e^{nh}}$$

and

$$\sum_{n=1}^{\infty} \frac{h}{1 + e^{nh}} = \sum_{n=1}^{\infty} \int_{nh - h/2}^{nh + h/2} \frac{dx}{1 + e^x} - \sum_{n=1}^{\infty} \left( \int_{nh - h/2}^{nh + h/2} \frac{dx}{1 + e^x} \right) \frac{h}{1 + e^{nh}}.$$

(31)

Each summand in the last sum is equal to the error in integrating $1/(1 + e^x)$ by the midpoint rule, over a certain interval. Using the usual expression for this error (see, e.g. [12, p. 17]), we see that the $n$th summand is

$$\frac{h^3}{24} \frac{e^{2\epsilon} - e^{\epsilon}}{(e^{\epsilon} + 1)^3}.$$
where \( \xi = \xi(n) \) is some point in \((nh - h/2, nh + h/2)\). Therefore

\[
\left| \sum_{n=1}^{\infty} \left( \int_{nh-h/2}^{nh+h/2} \frac{dx}{1+e^x} - \frac{h}{1+e^{nh}} \right) \right| \leq \frac{h^3}{24} \sum_{n=1}^{\infty} e^{-\xi(n)} \\
\leq \frac{h^3}{24} \sum_{n=1}^{\infty} e^{-(nh-h/2)} \\
\leq \frac{e^{h/2}}{24} \frac{h^3}{e^h - 1} = O(h^2).
\]

Furthermore,

\[
\sum_{n=1}^{\infty} \int_{nh-h/2}^{nh+h/2} \frac{dx}{1+e^x} = \int_{h/2}^{\infty} \frac{dx}{1+e^x} = \log(1+e^{-h/2}) \\
= \log 2 - \frac{h}{4} + O(h^2).
\]

From these last two estimates and (31) we see that

\[
(32) \quad \sigma(h) = \frac{\log 2}{h} - \frac{1}{4} + O(h),
\]

as \( h \to 0^+ \). We may sum up (29) and (32) in

**Theorem 2.** If \( h \to 0 \) and \( N \to \infty \) simultaneously in such a manner that \( Nh \to \infty \), then

\[
\|E_{h,N}\| = \|E_{h,N}\|^2 = 4\pi^2 e^{-2\pi^2/h} + \frac{2h^2}{e^h - 1} (1 + 4\sigma(h)) e^{-Nh}
\]

\[
+ O(e^{-4\pi^2/h}) + O(e^{-2Nh})
\]

\[
= 4\pi^2 e^{-2\pi^2/h} + 8 \log 2 e^{-Nh}
\]

\[
+ O(e^{-4\pi^2/h}) + O(N e^{-Nh}).
\]

The parameter \( N \) may be taken as a rough measure of the computational effort involved in using the quadrature formula \( T_{h,N} \); \( N \) values of \( \cosh \) and \( \tanh \), and \( 2N + 1 \) values of the integrand, must be computed. It is natural to want to choose \( h = h(N) \) so as to make the numerical quadrature as accurate as possible for the given level of effort. There is no way to do this optimally for each particular integrand; a reasonable approach is to regard that \( h \) as best, which minimizes \( \|E_{h,N}\| \). From (34) it is then clear that for the best \( h \) we must have

\[
2\pi^2/h = Nh + O(1)
\]

as \( N \) grows to infinity (and \( h \) decreases toward zero). It follows that

\[
Nh^2 = 2\pi^2 + O(h)
\]

and so

\[
\frac{h}{\sqrt{2\pi^2/N}} = \frac{h^2}{2\pi^2/N + O(h/N)} = 2\pi^2/N + O(N^{-3/2})
\]
and

\[ h = \pi \sqrt{2/N} + O(1/N). \]  

This justifies the value \( h = \pi \sqrt{2/N} \) suggested by Schwartz.

If we denote that value by \( \"h_0\" \) we have, by Theorem 2,

\[ \|E_{h_0,N}\| \sim (4\pi^2 + 2h_0^2(1 + 4\sigma(h_0)))^{1/2} e^{h_0 - 1} e^{-(\pi/\sqrt{2})\sqrt{N}} \]  

\[ \sim (4\pi^2 + 8 \log 2)^{1/2} e^{-(\pi/\sqrt{2})\sqrt{N}}. \]  

It is useful to know how well these asymptotic forms represent \( \|E_{h_0,N}\| \). The norm itself can be computed directly, for any \( h \) and \( N \), from the formula

\[ \|E_{h,N}\|^2 = \frac{\pi^2}{2} - 2T_{h,N}\left(\frac{1}{x} \log \frac{1+x}{1-x}\right) + T_{h,N,x}T_{h,N,y}\left(\frac{1}{1-xy}\right), \]  

which follows from (9). One has only to carry out the indicated numerical integrations. This was done on the CDC6600 computer at the Hebrew University in Jerusalem; the calculations were done in double-precision, or roughly 29 significant figure, arithmetic) and the results are presented in Table 1. The second column of that table contains the directly computed values of \( \exp(\pi\sqrt{N}/2) \cdot \|E_{h_0,N}\| \) while the third contains the values of the coefficient in (36); the coefficient in (37) is, to 4 figures, 6.710. The approximation by (36) is remarkably close; that by (37) is also fairly good.

<table>
<thead>
<tr>
<th>( N )</th>
<th>computed coefficient</th>
<th>from (36)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>6.421 7508</td>
<td>6.419 6241</td>
</tr>
<tr>
<td>9</td>
<td>6.485 2753</td>
<td>6.485 2504</td>
</tr>
<tr>
<td>16</td>
<td>6.528 1010</td>
<td>6.528 1009</td>
</tr>
<tr>
<td>25</td>
<td>6.557 6153</td>
<td>6.557 6153</td>
</tr>
<tr>
<td>36</td>
<td>6.579 0284</td>
<td>6.579 0283</td>
</tr>
<tr>
<td>49</td>
<td>6.595 2250</td>
<td>6.595 2248</td>
</tr>
</tbody>
</table>

It is clear from (34) and (35) that no other value of \( h \) would increase the rate of convergence beyond \( \exp(-\pi\sqrt{N}/2) \); but we may be able to obtain smaller coefficients than those in (36) and (37). Setting

\[ h = \pi \sqrt{2/N} + \alpha/N, \]

we find that

\[ \exp(-2\pi^2/h) \sim e^\alpha \exp(-\pi \sqrt{2N}), \]

\[ \exp(-Nh) \sim e^{-\alpha} \exp(-\pi \sqrt{2N}), \]
and

\[ \frac{2h^2(1 + 4\sigma(h))}{e^h - 1} \sim 8 \log 2. \]

It follows, after some calculation, that the best value of \( \alpha \) is

\[ \log \frac{\sqrt{2 \log 2}}{\pi}. \]

This number is \(-.9814 \cdots\), which is close to \(-1\). If we set

\[ h_1 = h_1(N) = \pi \sqrt{\frac{2}{N}} + \frac{1}{N} \log \frac{\sqrt{2 \log 2}}{\pi} \]

and, for comparison

\[ h_2 = h_2(N) = \pi \sqrt{\frac{2}{N}} - \frac{1}{N}, \]

we find that

\begin{align*}
(38) & \quad ||E_{h_1,N}|| \sim \sqrt{\pi^2/2 \log 2} e^{-(\pi/\sqrt{2})\sqrt{N}}, \\
(39) & \quad ||E_{h_2,N}|| \sim \sqrt{\pi^2/8 + 8e \log 2} e^{-(\pi/\sqrt{2})\sqrt{N}}.
\end{align*}

The constants in (38) and (39) are 5.43981 \cdots and 5.44028 \cdots respectively; to four figures, they are both 5.440. The best possible \( h \) would differ from \( h_1 \) by only \( o(N^{-1}) \), which would yield no better constant. We therefore have

**Theorem 3.** With \( h \) chosen optimally—to minimize \( ||E_{h,N}|| \)—for each \( N \), we have

\[ ||E_{h,N}|| \sim \sqrt{8\pi \sqrt{2 \log 2}} e^{-(\pi/\sqrt{2})\sqrt{N}}, \]

and with \( h_2(N) = \pi \sqrt{2/N - 1/N} \),

\[ ||E_{h_2,N}|| \sim 5.440 \cdots e^{-(\pi/\sqrt{2})\sqrt{N}}. \]

**5. Some numbers.** In order to help develop a feeling for the actual operation of the tanh rule, we have listed in Table 2 some numerical values of the abscissas and coefficients used. For each value of \( N \) only the positive abscissas, and the corresponding coefficients, are listed; as is clear from (2) and (3), the rule is symmetric about the origin (and the origin itself is always one of the abscissas, with coefficient \( h/2 \)). Table 2 is based on setting \( h = h_2(N) \).

One sees immediately that most of the abscissas cluster very close to \( \pm 1 \). This is also evident from the definition of the rule, since for \( |nh| \) moderately large (and it can be about as large as \( \pi \sqrt{2N} \)),

\[ \tanh \frac{nh}{2} = \pm 1 + O(e^{-|nh|}). \]
**THE TANH RULE**

**Table 2**

<table>
<thead>
<tr>
<th>Abscissas</th>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N = 1)</td>
<td></td>
</tr>
<tr>
<td>.938036</td>
<td>.206724</td>
</tr>
<tr>
<td>(N = 2)</td>
<td></td>
</tr>
<tr>
<td>.866982</td>
<td>.328010</td>
</tr>
<tr>
<td>.989899</td>
<td>.026548</td>
</tr>
<tr>
<td>(N = 3)</td>
<td></td>
</tr>
<tr>
<td>.806132</td>
<td>.390728</td>
</tr>
<tr>
<td>.977219</td>
<td>.050262</td>
</tr>
<tr>
<td>.997530</td>
<td>.005506</td>
</tr>
<tr>
<td>(N = 4)</td>
<td></td>
</tr>
<tr>
<td>.755532</td>
<td>.423044</td>
</tr>
<tr>
<td>.961953</td>
<td>.073580</td>
</tr>
<tr>
<td>.994614</td>
<td>.010590</td>
</tr>
<tr>
<td>.999248</td>
<td>.001482</td>
</tr>
<tr>
<td>(N = 5)</td>
<td></td>
</tr>
<tr>
<td>.713098</td>
<td>.439127</td>
</tr>
<tr>
<td>.945434</td>
<td>.094844</td>
</tr>
<tr>
<td>.990649</td>
<td>.016631</td>
</tr>
<tr>
<td>.998428</td>
<td>.002807</td>
</tr>
<tr>
<td>.999737</td>
<td>.000471</td>
</tr>
</tbody>
</table>

This can lead to computational pitfalls: integrands containing factors such as \(1 - x\) may be impossible to evaluate directly, on some machines, because of "underflow"; and many integrands will need special evaluation at abscissas very near \(\pm 1\) to avoid excessive loss of significance.

Table 3 gives the values of \(E_{h,N}\) for low \(N\) and for various values of \(h\). The advantage of \(h_2\) over \(h_0\) is seen to be greater for low \(N\) than it is asymptotically as \(N\) grows large. It is a bit surprising that \(h_1\) is not quite as good as \(h_2\) for low \(N\); asymptotically it is somewhat better. The optimal \(h\) was determined numerically for each \(N\) and is given to 3 decimal places in the last column of the table. It is clear from the next to last column that the optimal \(h\) does not afford enough advantage over \(h_2\) to justify the effort of finding and keeping it.

**Table 3**

| \(N\) | \(|E_{h_0,N}|\) | \(|E_{h_2,N}|\) | \(|E_{h_1,N}|\) | \(|E_{h_{opt},N}|\) | \(h_{opt}\) |
|-------|----------------|----------------|----------------|----------------|---------|
| 1     | .705           | .398           | .402           | .331           | 2.893   |
| 2     | .276           | .172           | .173           | .153           | 2.381   |
| 3     | .137           | .0885          | .0890          | .0814          | 2.073   |
| 4     | .0755          | .0502          | .0504          | .0470          | 1.861   |
| 5     | .0448          | .0303          | .0304          | .0287          | 1.705   |
| 10    | .00578         | .00409         | .00411         | .00398         | 1.274   |
| 20    | .000317        | .000233        | .000233        | .000233        | .932    |

6. **Conclusions.** We can conclude that \(h_2(N)\) should be used in practical application of the tanh rule. As compared to the use of \(h_0\), this lowers the error
bound by about twenty to thirty percent without any increase in the computational effort.

Theorem 3 establishes that the convergence of the tanh rule is indeed of the $e^{-c\sqrt{N}}$ type, with optimal choice of the parameters. This remarkable convergence rate is faster than that of any of the classical (trapezoid, Simpson's, etc.) rules based on internal subdivision; and slower than the $e^{-cN}$ rate that Gaussian rules give in integrating functions that are analytic throughout a neighborhood of the integration interval. To the best of my knowledge, this convergence rate made its first appearance in approximation theory in D. J. Newman's theorem on rational approximation of $|x|$. Since then, Newman's theorem has been applied to give other results involving such a rate of convergence. In each case the context involves analyticity and isolated singularities, as is the case in this paper.

Turan and Szűsz [13] showed that Newman's theorem extends to a certain class of piecewise analytic functions. Now all functions of one variable that enter into numerical calculations are piecewise analytic (though not all in the Turan-Szűsz sense), and numerical analysis has generally ignored this. Classically numerical analysis has been based on the idea of polynomial approximation (including piecewise polynomials), and its theorems give convergence rates of the $N^{-k}$ type; or of $e^{-cN}$ type for some classes of analytic functions. Newman's result and its extensions give a tantalizing hint that a new approach based on rational approximation may make $e^{-c\sqrt{N}}$-type convergence available for essentially all functions.

As I was completing this paper, I became aware of the paper [14]. In it H. L. Loeb and H. Werner prove, by a construction which proceeds directly from Newman's construction of his rational approximation to $|x|$, that there are quadrature formulas giving at least $e^{-c\sqrt{N}}$-type convergence for $H^2$. An examination of their proof shows that the tanh rule is in fact essentially the quadrature rule that they derive. This greatly encourages the hope for the new approach mentioned in the previous paragraph.

Acknowledgment. I wish to express my thanks to Martin Milgram for suggesting the use of the identity (10), which considerably simplified the discussion in § 3 of this paper; and to George Weiss for many helpful conversations and for teaching me to use the Poisson sum formula.

REFERENCES