TWO FORMULAS FOR NUMERICAL INDEFINITE INTEGRATION

SEYMOUR HABER

ABSTRACT. We derive two formulas for approximating the indefinite integral over a finite interval. The approximation error is $O(c^{-c\sqrt{m}})$ uniformly, where m is the number of integrand evaluations. The integrand is required to be analytic in the interior of the integration interval, but may be singular at the endpoints. Some sample calculations indicate that the actual convergence rate accords with the error bound derived.

1. INTRODUCTION

The formulas we propose are

(A)
$$\int_{-1}^{s} g(u)\, du = \frac{h}{\pi} \sum_{k=-N}^{N} \psi'(kh) \cdot g(\psi(kh)) \cdot \mathrm{Si}\left(\frac{\pi\varphi(s)}{h} - \pi k\right)$$
$$+ \frac{1}{2} I^* + O(N^{1/2} e^{-\sqrt{\pi d \alpha N}})$$

and

(B)
$$\int_{-1}^{s} g(u)\, du = h \sum_{k=-N}^{N} \sum_{j=-N}^{N} \psi'(jh) \cdot g(\psi(jh)) \cdot \sigma_{k-j} \cdot \mathrm{sinc}\left(\frac{\varphi(s)}{h} - k\right)$$
$$+ I^* \cdot \left(\eta(s) - \sum_{k=-N}^{N} \left(-\frac{1}{2} + \eta(\psi(kh)) \right) \mathrm{sinc}\left(\frac{\varphi(s)}{h} - k\right) \right)$$
$$+ O(N^{3/2} e^{-\sqrt{\pi d \alpha N}}).$$

In each formula

$$I^* = h \sum_{k=-N}^{N} \psi'(kh) g(\psi(kh))$$

and

$$h = \sqrt{\frac{\pi d}{\alpha N}};$$

d and α are parameters related to the integrand g, and ψ and φ are the functions $\psi(x) = \tanh(x/2)$, $\varphi(x) = \log((1+x)/(1-x))$. The function Si is the sine integral:

$$\mathrm{Si}(x) = \int_{0}^{x} \frac{\sin t}{t}\, dt;$$

σ_n, for n an integer, is $(1/\pi)\operatorname{Si}(n\pi)$, and "sinc($x$)" is a common engineering notation for the function $\frac{\sin \pi x}{\pi x}$. Finally, η is an auxiliary function whose required properties are described below; an example is $\eta(x) = (1 + x)/2$.

The attractive thing about these complicated formulas is their accuracy. Using only $2N + 1$ values of the integrand g, they approximate the indefinite integral uniformly within $O(e^{-c\sqrt{N}})$ for some $c = c(g)$. This fast convergence rate is characteristic of a large class of numerical approximations developed by Frank Stenger and his school—see [7] for an introduction and survey. Formulas (A) and (B) belong to this class and are related to formula (4.58) of [7] and Theorem 3.3 of [3].

For such a fast convergence rate the integrand g must be very smooth, and indeed we require it to be analytic on $(-1, 1)$. But g is allowed to be singular at ± 1, with some restriction on its growth rate as $x \to 1$ or -1. This contrasts with classical requirements such as "$g \in C^2$" or "$g \in C^4$" to have convergence rates $O(N^{-2})$ and $O(N^{-4})$ for the trapezoid rule and Simpson's rule. The requirement of analyticity seems much more stringent, but in practice is not likely to be restrictive at all. Functions that come up for numerical integration are almost always analytic with isolated singularities, and it is usually possible to arrange that the integration take place on intervals where the singularities are only at endpoints.

Formulas (A) and (B) are based on the standard strategy for numerical quadrature: approximate the integrand by a function that is easy to integrate, and integrate the latter. In this case the approximation is by linear combinations of translates of $\frac{\sin \pi x}{\pi x}$, whose integral is not elementary. Both formulas involve values of the sine integral, formula (B) through the quantities σ_{k-j}. The sine integral is one of the most thoroughly studied of the "higher transcendental functions" and is not hard to evaluate. Good methods are given in [6] and in the Appendix to this paper.

The analysis leading to these formulas takes place most naturally on the interval $(-\infty, \infty)$ rather than $(-1, 1)$. The central part of the analysis is largely taken from a paper by McNamee, Stenger, and Whitney [4] and from later developments by Stenger and his students. I shall give a full derivation of (A) and (B) for the convenience of the reader, as it seems that even the central part has appeared in print only incompletely.

2. DERIVATION OF (A)

We start from the contour integral

$$\frac{1}{2\pi i} \int_{L_{n,\varepsilon}} \frac{f(z)\,dz}{(z - t)\sin \frac{\pi z}{h}},$$

where $L_{n,\varepsilon}$ is the rectangular contour whose vertical sides are at $x = \pm(n + 1/2)h$ (n a positive integer) and whose horizontal sides are at $y = \pm(d - \varepsilon)$. The parameters h and d are positive. We assume

(H$_1$) f is analytic in the strip $|y| < d$.

We suppose that the real number t is $\leq nh$ in absolute value, and evaluate the integral by residues. The singularities inside $L_{n,\varepsilon}$ are at $z = t$ with residue $f(t)/\sin(\pi t/h)$ and at $z = kh$, k any integer between $-n$ and n, with residue

(if $t \neq kh$)

$$(-1)\frac{h}{\pi}\frac{f(kh)}{kh-t}.$$

After some algebra we get the equation

$$f(t) = \sum_{k=-n}^{n} f(kh)\operatorname{sinc}\left(\frac{t}{h}-k\right) + R_{n,\varepsilon}(t),$$

(1)

$$R_{n,\varepsilon}(t) = \frac{1}{2\pi i}\sin\left(\frac{\pi t}{h}\right)\int_{L_{n,\varepsilon}}\frac{f(z)\,dz}{(z-t)\sin(\pi z/h)}.$$

It is immediate that (1) holds also for $t = kh$, $|k| \leq n$, and so for all t in $[-nh, nh]$. We now further assume that

(H$_2$)   for small positive ε each of the integrals $\int_{-\infty}^{\infty}|f(x \pm i(d - \varepsilon))|\,dx$ exists, and these integrals are bounded in ε;

and

(H$_3$)   for each small positive ε the integral $\int_{-d+\varepsilon}^{d-\varepsilon}|f(x + iy)|\,dy$ is a bounded function of x.

Let $n \to \infty$ in (1). The integrals over the vertical side of $L_{n,\varepsilon}$ approach zero, and the integrals over the horizontal sides approach

(2) $$I_{\pm} = \int_{-\infty}^{\infty}\frac{f(x \pm i(d - \varepsilon))\,dx}{(x - t \pm i(d - \varepsilon))\sin\frac{\pi}{h}(x \pm i(d - \varepsilon))}.$$

We state this as

Lemma 1. *If f satisfies conditions* (H$_1$), (H$_2$), *and* (H$_3$), *then*

(3) $$f(t) = \sum_{k=-\infty}^{\infty} f(kh)\operatorname{sinc}\left(\frac{t}{h}-k\right) + \frac{1}{2\pi i}\cdot\sin\frac{\pi t}{h}\cdot(I_{-}-I_{+})$$

for all real t.

Conditions (H$_2$) and (H$_3$) are somewhat obscure. In the transition to the interval $(-1, 1)$ they shall be replaced by simpler conditions.

We now integrate (3) to get

$$\int_{-A}^{x} f(t)\,dt = \sum_{k=-\infty}^{\infty} f(kh)\int_{-A}^{x}\operatorname{sinc}\left(\frac{t}{h}-k\right)dt + \frac{1}{2\pi i}(R_{-}-R_{+}),$$

(4)

$$R_{\pm} = \int_{-A}^{x}\sin\frac{\pi t}{h}\int_{-\infty}^{\infty}\frac{f(s \pm i(d - \varepsilon))}{(s - t \pm i(d - \varepsilon))\sin\frac{\pi}{h}(s \pm i(d - \varepsilon))}\,ds\,dt.$$

The interchange of integration and summation is enabled by imposing a new condition (which we shall also need below) on f:

(H$_4$)   there is a constant $\alpha > 0$ such that for x real, $$f(x) = O(e^{-\alpha|x|}) \text{ as } |x| \to \infty.$$

We can exchange the order of integrations in $R_{\pm}$, since

$$\int_{B}^{\infty}\frac{f(s \pm i(d - \varepsilon))\,ds}{(s - t \pm i(d - \varepsilon))\sin\frac{\pi}{h}(s \pm i(d - \varepsilon))} \to 0$$

as $B \to \infty$, uniformly for $t \in [-A, x]$ (using condition (H_2)), and the same holds for the integral from $-\infty$ to $-B$. The resulting inner integral

$$\int_{-A}^{x} \frac{\sin(\pi t/h)\, dt}{s - t \pm i(d - \varepsilon)}$$

is $O(h)$ uniformly in A, x, s, and ε. Since

$$\left| \sin \frac{\pi}{h}(s \pm i(d - \varepsilon)) \right| \geq \sinh \frac{\pi(d - \varepsilon)}{h} \geq \frac{1}{2} e^{\pi(d-\varepsilon)/h},$$

we conclude that

$$|R_\pm| \leq Che^{-\pi(d-\varepsilon)/h} \int_{-\infty}^{\infty} |f(s \pm i(d - \varepsilon))|\, ds.$$

Letting $A \to \infty$ and then $\varepsilon \to 0$ in (4), we have that

$$\int_{-\infty}^{x} f(t)\, dt = h \sum_{k=-\infty}^{\infty} f(kh) \cdot \left(\frac{1}{2} + \frac{1}{\pi} \operatorname{Si}\left(\frac{\pi x}{h} - \pi k \right) \right) + O(he^{-\pi d/h})$$

as $h \to 0$. Now we use (H_4) (and the boundedness of Si) to write

$$\int_{-\infty}^{x} f(t)\, dt = h \sum_{k=-N}^{N} f(kh) \cdot \left(\frac{1}{2} + \frac{1}{\pi} \operatorname{Si}\left(\frac{\pi x}{h} - \pi k \right) \right) + O(he^{-\pi d/h})$$
$$+ O(h^{-1} e^{-\alpha Nh})$$

as $h \to 0$ and $Nh \to \infty$. Setting

$$(5) \qquad\qquad\qquad h = h(N) = \sqrt{\frac{\pi d}{\alpha N}}$$

nearly balances the two error terms and we have:

Theorem 1. *If f satisfies conditions (H_1), (H_2), (H_3), and (H_4), and h is defined by (5), then*

$$\int_{-\infty}^{x} f(t)\, dt = h \sum_{k=-N}^{N} f(kh) \cdot \left(\frac{1}{2} + \frac{1}{\pi} \operatorname{Si}\left(\frac{\pi x}{h} - \pi k \right) \right) + O(N^{\frac{1}{2}} e^{-\sqrt{\pi d \alpha N}})$$

as $N \to \infty$, uniformly for $x \in (-\infty, \infty)$.

(We could have, by modifying (5), eliminated the $N^{\frac{1}{2}}$ factor in the error term in the theorem. Calculations done with the modified h do not, however, show a consistent advantage for practical values of N.)

The transition to $(-1, 1)$ is made via the change of variables

$$w = \psi(z) = \tanh z/2.$$

With $z = x + iy$ and $w = u + iv$, ψ takes the line $-\infty < x < \infty$ monotonically onto $-1 < u < 1$. Each horizontal line $y = \delta$, with $|\delta| < \pi$, becomes a circular arc going from $w = -1$, through $w = i\tan(\delta/2)$, to $w = +1$; the center and radius of the circle it is part of are $-i \cot \delta$ and $\operatorname{cosec} \delta$, respectively. A vertical line segment

$$z = x + iy, \quad |y| < \delta \qquad (\delta < \pi)$$

becomes a circular arc whose center and radius are

(6) $$\frac{1 + e^{-2x}}{1 - e^{-2x}} \operatorname{sgn}(x) \quad \text{and} \quad \frac{2e^{-x}}{1 - e^{-2x}} \operatorname{sgn}(x),$$

respectively.

The strip $|y| < d$ $(d < \pi)$ is conformally mapped to the lenticular region Λ_d whose boundary consists of the two circular arcs that are the images of the lines $y = d$ and $y = -d$. The vertical width of Λ_d is $2\tan(d/2)$; the region $\Lambda_{\pi/2}$ is the unit disk and Λ_π is the whole w-plane, slit from $-\infty$ to -1 and from 1 to ∞ on the u-axis.

We write

$$\int_{-\infty}^{x} f(t)\,dt = \int_{-1}^{s} g(u)\,du,$$

using the change of variables $u = \psi(t)$. Then

$$s = \psi(x), \quad x = \varphi(s) = \log((1+s)/(1-s)), \quad f(t) = g(\psi(t)) \cdot \psi'(t),$$

and the equation in Theorem 1 becomes (A).

Conditions (H_1) and (H_4) on f translate straightforwardly into the conditions

(H_1') \qquad\qquad\qquad\qquad g is analytic on Λ_d

and

(H_4') \quad there is a constant $\alpha > 0$ such that for real u, $g(u) = O((1 - u^2)^{-1+\alpha})$ as $u \to \pm 1$ from inside $(-1, 1)$.

The integral in condition (H_3) becomes

$$\int_{C_{x,\varepsilon}} |g(w)\,dw|,$$

where $C_{x,\varepsilon}$ is the image under ψ of the vertical line segment

$$z = x + iy, \qquad |y| < d - \varepsilon;$$

by (6) it is an arc of a circle of radius $2e^{-|x|} + O(e^{-2|x|})$ and center $\pm 1 + O(e^{-2|x|})$. So, if we require of g that

$$g(w) = O(|1 - w^2|^{-1})$$

as $w \to \pm 1$ from inside Λ_d, then (H_3) would be satisfied. However, we shall require the stronger condition

(H_3') \quad there is a constant $\alpha > 0$ such that $g(w) = O(|1 - w^2|^{-1+\alpha})$ as $w \to \pm 1$ from inside Λ_d.

This makes condition (H_4') unnecessary, and it also simplifies the translation of (H_2).

Let P_a be the image under ψ of the line $y = a$, and let $P_{a,\beta}$, with $\beta > 0$, be that part of P_a that lies outside circles of radius β centered at 1 and at -1. The integrals in (H_2) become

(7) $$\int_{P_{\pm(d-\varepsilon)}} |g(w)\,dw|,$$

which surely exist once (H'_3) is assumed. The condition (H'_3) also implies that the contribution to (7) from the portion of $P_{\pm(d-\varepsilon)}$ that lies within distance β of either $+1$ or -1 is bounded in ε, for small β. Thus we need only assume that

(H'_2) for small $\beta > 0$ the integrals $\int_{P_{\pm(d-\varepsilon),\beta}} |g(w)\,dw|$ are bounded in ε for $0 < \varepsilon < d$.

To determine whether (H'_2) is satisfied, for a given g that satisfies (H'_1) and (H'_3), we need only consider singularities of g at points on the boundary of Λ_d other than $+1$ and -1. Furthermore, if d' is any positive number that is less than d, then (H'_2) will automatically hold with d' in place of d. So we have

Theorem 2. *If g is analytic in Λ_d for some positive $d \leq \pi$ and $g(w) = O(|1 - w^2|^{-1+\alpha})$ for some $\alpha > 0$ as $w \to \pm 1$ from inside Λ_d, then*

$$\int_{-1}^{s} g(u)\,du = \frac{h}{\pi} \sum_{k=-N}^{N} \psi'(kh) \cdot g(\psi(kh)) \cdot \mathrm{Si}\left(\frac{\pi\varphi(s)}{h} - \pi k\right) + \frac{1}{2}I^*$$
$$+ O(N^{1/2}e^{-\sqrt{\pi d'\alpha N}})$$

uniformly in $[-1, 1]$, where d' is any number in $(0, d)$ and $h = h(N)$ is defined by (5) with d' in place of d. Here,

$$(8) \qquad\qquad I^* = h \sum_{k=-N}^{N} \psi'(kh)g(\psi(kh)).$$

We also have

Theorem 2a. *If $0 < d \leq \pi$ and $\alpha > 0$, and g satisfies conditions (H'_1), (H'_2), and (H'_3), then (A) holds uniformly in $[-1, 1]$, with $h = h(N)$ defined by (5) and I^* defined by (8).*

3. Derivation of (B)

The point of formula (B) is to reduce the need for values of the sine integral; in (B) we use only the values $\mathrm{Si}(n\pi)$, n an integer. These can be calculated more simply (see the appendix) than general values of Si.

We again work primarily in the context of integrals over the real line. In outline, we first set

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

and approximate F by an interpolation formula that uses evaluations of F at integral multiples of a parameter h. Then those values of F are themselves approximated using Theorem 1.

We get our interpolation formula from Lemma 1 by bounding the remainder there and then truncating the infinite sum. In the integrals in (2) the numerator is in $L^1(-\infty, \infty)$ by (H_2), and the integral of its absolute value is bounded. In addition, the first factor in the denominator is bounded away from zero uniformly in x and t. The second factor there is greater than $ce^{\pi(d-\varepsilon)/h}$ uniformly in x, where c is independent of ε. So

$$|I_\pm| = O(e^{-\pi d/h}) \quad \text{as } h \to 0.$$

Imposing condition (H_4) on f, we obtain

$$f(t) = \sum_{k=-N}^{N} f(kh) \operatorname{sinc}\left(\frac{t}{h} - k\right) + O(e^{-\pi d/h}) + O(h^{-1}e^{-\alpha Nh})$$

as $h \to 0$, $N \to \infty$. Using (5) to specify h, we then have

Lemma 2. *If f satisfies conditions* (H_1), (H_2), (H_3), *and* (H_4), *then*

$$f(t) = \sum_{k=-N}^{N} f(kh) \operatorname{sinc}\left(\frac{t}{h} - k\right) + O(N^{1/2}e^{-\sqrt{\pi d\alpha N}})$$

uniformly in $(-\infty, \infty)$, *with h defined by* (5).

Now we define

$$F(z) = F(x + iy) = \int_{-\infty}^{x} f(t)\,dt + \int_{0}^{y} f(x + is)\,ds$$

for z in the strip $|y| < d$. We cannot however apply Lemma 2 directly to F, since F does not necessarily satisfy (H_4). Indeed, since f satisfies (H_4),

$$F(x) = I + O(e^{-\alpha|x|}) \quad \text{as } x \to +\infty,$$

where

$$I = \int_{-\infty}^{\infty} f(t)\,dt$$

may not be zero. We therefore introduce

$$F^*(z) = F(z) - I \cdot \xi(z),$$

where ξ is an auxiliary function satisfying (H_1) and the new condition

(H_5)   for the same $\alpha > 0$ as in (H_4), $\xi(x) = O(e^{-\alpha|x|})$ *as* $x \to -\infty$
and $\xi(x) = 1 + O(e^{-\alpha|x|})$ *as* $x \to +\infty$.

(An example is $\xi(z) = 1/(1 + e^{-\alpha z})$.) The function F^* satisfies (H_1) and (H_4).

In order to make F^* satisfy (H_2) and (H_3) we require that f and ξ' satisfy, in place of (H_3), the stronger condition

(H_{3a})   for any small positive ε the function $V_\varepsilon(x) = \int_{-d+\varepsilon}^{d-\varepsilon} |f(x + iy)|\,dy$
is bounded and is in $L^1(-\infty, \infty)$. Furthermore, $\int_{-\infty}^{\infty} V_\varepsilon(x)dx$
is bounded in ε.

That F^* now satisfies (H_2) and (H_3) can be seen by writing, e.g.,

$$\int_{-\infty}^{\infty} |F^*(x + i(d - \varepsilon))|\,dx$$

$$= \int_{-\infty}^{\infty} \left| F^*(x) + \int_{0}^{d-\varepsilon} (f(x + iy) - I\xi'(x + iy))\,dy \right| dx$$

$$\leq \int_{-\infty}^{\infty} |F^*(x)|\,dx + \int_{-\infty}^{\infty} \int_{0}^{d-\varepsilon} |f(x + iy)|\,dy\,dx$$

$$+ |I| \cdot \int_{-\infty}^{\infty} \int_{0}^{d-\varepsilon} |\xi'(x + iy)|\,dy\,dx.$$

Applying Lemma 2 to F^*, we obtain:

Lemma 3. *Assume that f satisfies* (H_1), (H_2), (H_{3a}), *and* (H_4), *that ξ satisfies* (H_1) *and* (H_5), *and ξ' satisfies* (H_{3a}), *and that h is defined by* (5) *and*

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \ \text{and} \ I = \int_{-\infty}^{\infty} f(t)\,dt\,.$$

Then

$$F(x) = \sum_{k=-N}^{N} F(kh)\,\mathrm{sinc}\left(\frac{x}{h} - k\right)$$

$$+ I \cdot \left(\xi(x) - \sum_{k=-N}^{N} \xi(kh)\,\mathrm{sinc}\left(\frac{x}{h} - k\right)\right)$$

$$+ O(N^{1/2}e^{-\sqrt{\pi d\alpha N}})\,.$$

Now we use Theorem 1 to replace each $F(kh)$ in the last equation by

$$h \sum_{j=-N}^{N} f(jh)\left(\frac{1}{2} + \frac{1}{\pi}\,\mathrm{Si}(\pi(j-k))\right)$$

and I by

$$I^* = h \sum_{j=-N}^{N} f(jh)\,.$$

Each of the $2N + 1$ replacements of an $F(kh)$ introduces an error that is no greater than $CN^{1/2}e^{-\sqrt{\pi d\alpha N}}$ in absolute value; one constant C will do for all k's. Replacing I by I^* introduces another such error, and the quantity multiplying I is itself $O(N)$ because ξ is bounded. Using the notation

$$\sigma_n = \frac{1}{\pi}\,\mathrm{Si}(n\pi)\,,$$

we have, after some algebra:

Theorem 3. *Assume that f satisfies* (H_1), (H_2), (H_{3a}), *and* (H_4), *that ξ satisfies* (H_1) *and* (H_5) *and ξ' satisfies* (H_{3a}), *and that h is defined by* (5). *Then*

$$\int_{-\infty}^{x} f(t)\,dt = h \sum_{k=-N}^{N} \sum_{j=-N}^{N} f(jh) \cdot \sigma_{k-j} \cdot \mathrm{sinc}\left(\frac{x}{h} - k\right)$$

$$+ I^* \cdot \left(\xi(x) - \sum_{k=-N}^{N}\left(\xi(kh) - \frac{1}{2}\right)\mathrm{sinc}\left(\frac{x}{h} - k\right)\right)$$

$$+ O(N^{3/2}e^{-\sqrt{\pi d\alpha N}})$$

uniformly for $-\infty < x < \infty$. *Here,* $I^* = h\sum_{j=-N}^{N} f(jh)$. *If it is known that*

$$\int_{-\infty}^{\infty} f(t)\,dt = 0\,,$$

then I^* may be replaced by 0.

To get formula (B), we again use $w = \psi(z) = \tanh(z/2)$ and set $f(t) = g(\psi(t)) \cdot \psi'(t)$; and we also set

$$\eta(w) = \xi(\varphi(w))\,.$$

Conditions (H'_2) and (H'_3) on g imply conditions (H_2), (H_{3a}), and (H_4) on f, and condition (H'_1) on g is equivalent to condition (H_1) on f. Similarly, (H'_1) for η is equivalent to condition (H_1) for ξ. Condition (H_5) on ξ is guaranteed by requiring that

(H'_5) η is continuous on $[-1, 1]$ and $\eta(-1) = 0$, $\eta(1) = 1$

and that η satisfy a Hölder condition of order α on $[-1, 1]$. In order to have ξ' satisfy (H_{3a}), we impose condition (H'_3) on η'; condition (H'_3) for η' and (H'_5) for η together imply the Hölder condition for η. We conclude:

Theorem 4a. *If $0 < d \le \pi$ and $\alpha > 0$ and g satisfies conditions (H'_1), (H'_2), and (H'_3), and η satisfies (H'_1) and (H'_5), and η' satisfies (H'_3), then* (B) *holds uniformly in $[-1, 1]$, with $h = h(N)$ defined by* (5) *and I^* defined by* (8).

The conditions on η are not onerous. The function $\eta(w) = (1 + w)/2$ satisfies the conditions for all d and any $\alpha \le 1$; the function $\eta(w) = (2 + 3w - w^3)/4$ has $\eta'(-1) = \eta'(1) = 0$ and so satisfies the conditions for all d and any $\alpha \le 2$.

As with Theorem 2, we can avoid dealing with condition (H'_2) by accepting a smaller value of d in the error term. We state this as

Theorem 4. *If $0 < d \le \pi$ and $\alpha > 0$ and*

1) *g is analytic in Λ_d and $g(w) = O(|1 - w^2|^{-1+\alpha})$ as $w \to \pm 1$ from inside Λ_d;*
2) *η is analytic in Λ_d and $\eta(-1) = 0$ and $\eta(1) = 1$; and $\eta'(w) = O(|1 - w^2|^{-1+\alpha})$ as $w \to \pm 1$ from inside Λ_d,*

then

$$\int_{-1}^{s} g(u)\,du = h \sum_{k=-N}^{N} \sum_{j=-N}^{N} \psi'(jh) \cdot g(\psi(jh)) \cdot \sigma_{k-j} \cdot \mathrm{sinc}\left(\frac{\varphi(s)}{h} - k\right)$$

$$+ I^* \cdot \left(\eta(s) - \sum_{k=-N}^{N} \left(\eta(\psi(kh)) - \frac{1}{2}\right) \mathrm{sinc}\left(\frac{\varphi(s)}{h} - k\right)\right)$$

$$+ O(N^{3/2} e^{-\sqrt{\pi d' \alpha N}})$$

uniformly in $[-1, 1]$, where $d' \in (0, d)$ and $h = h(N)$ is defined by (5) *with d' in place of d. The constant I^* is defined by* (8); *if*

$$\int_{-1}^{1} g(u)\,du = 0,$$

then I^* may be replaced by zero.

4. COMPUTATIONAL CONSIDERATIONS

The double sum in (B) apparently involves $2N + 1$ values of sine; but since the k occurring there is an integer,

$$\mathrm{sinc}\left(\frac{\varphi(s)}{h} - k\right) = \frac{\sin(\frac{\pi\varphi(s)}{h} - \pi k)}{\frac{\pi\varphi(s)}{h} - \pi k} = \frac{(-1)^k}{\varphi(s) - kh} \cdot \frac{h}{\pi} \cdot \sin\frac{\pi\varphi(s)}{h},$$

and the double sum can be written as

$$\frac{h^2}{\pi} \sin \frac{\pi\varphi(s)}{h} \sum_{k=-N}^{N} \sum_{j=-N}^{N} \frac{(-1)^k \sigma_{k-j}}{\varphi(s) - kh} \psi'(jh)g(\psi(jh)).$$

This involves only one sine evaluation.

Because of the double sum, (B) involves a good deal more arithmetic than (A), for a single value of s. But if the integral is to be computed for many values of s, we may rewrite the double sum as

$$\frac{h^2}{\pi} \sin \frac{\pi\varphi(s)}{h} \sum_{k=-N}^{N} \frac{S_k}{\varphi(s) - kh},$$

where

$$S_k = \sum_{j=-N}^{N} (-1)^k \sigma_{k-j} \psi'(jh)g(\psi(jh)),$$

and note that the S_k are independent of s. So the S_k may be computed first, and then each new value of s requires the calculation of only a single sum. As a result, formula (B) is faster than formula (A) when many values of the indefinite integral are wanted, because each new value of s requires only a logarithm, a sine, and two very simple sums in (B), while in (A) we must calculate $2N+1$ new values of the sine integral. In one of the examples presented below, formula (A) required roughly $CMN^{3/2}$ seconds to calculate the integral for M values of s (C being a machine-dependent constant), while (B) required $C(MN^{1/2} + N^2/2)$ seconds.

As has been remarked elsewhere ([5, p. 148], [2, p. 682]), some of the abscissas $\psi(jh)$ are very close to ± 1 when N is moderately large. Sometimes they are too close for the computer to distinguish them from ± 1, and the integrand g may be infinite at ± 1. This difficulty might be avoided by going over to higher-precision arithmetic or by direct study of the function $\psi'(x)g(\psi(x))$ and its accurate calculation for large values of x.

5. COMPUTATIONAL EXAMPLES

We shall look at the results of applying formulas (A) and (B) to four integrands, chosen to show the effects of different singularities. The integrands are

$$g_1(u) = \frac{1}{\pi\sqrt{1-u^2}}, \qquad g_2(u) = \frac{1}{4\log 2} \log\left(\frac{1+u}{1-u}\right),$$

$$g_3(u) \equiv \frac{1}{2}, \qquad g_4(u) = \frac{2}{\pi}\sqrt{1-u^2}.$$

The constants were chosen to "normalize" the functions in the sense that

$$\int_{-1}^{1} |g(u)| \, du = 1.$$

The first and fourth functions have $x^{-1/2}$- and $x^{1/2}$-type singularities at $u = \pm 1$; the second function goes to infinity weakly, and the third has no singularities. It should be remarked that formulas (A) and (B) are not expected

to integrate g_3 exactly; in contrast to the classical numerical integration formulas, no considerations of "degree of precision"—exactness for polynomial integrands—have entered into the present theory.

A negative feature of our formulas is that they require adjustment to the integrand that is treated. The parameter h, used in the computation, is defined in terms of d and α, which are numbers describing some of the behavior of the integrand. For our four examples the appropriate values of α are clearly $1/2$ for g_1, 1 for g_3, and $3/2$ for g_4; it is not so clear what α to use for g_2. Any α that is less than 1 will do, but 1 itself will not do.

None of our four functions has singularities anywhere but at ± 1, so at first thought it seems that $d = \pi$ is appropriate to all of them. But in fact none of them satisfies condition (H_2') with $d = \pi$. That is because each contour $P_{\pm(\pi-\varepsilon),\beta}$ is not bounded in length as ε approaches zero. Its length is asymptotic to $2\pi/\varepsilon$, so that it is immediate that g_3 does not satisfy (H_2'); and it is not hard to see that is also the case for the other g's.

Of course, all four g's satisfy (H_2') for any d less than π. This makes the problem of applying Theorems 2a and 4a the same as that of applying Theorems 2 and 4: all values of d less than a given d_0 may validly be used, but not d_0 itself if the error estimate in the theorem is to hold. What we shall do in practice is use d_0 itself.

Similarly, we shall use $\alpha = 1$ for the integrand g_2.

We present the results of applying each formula with several values of N, in order to see its convergence behavior.

Since the main factor in the error bound is $e^{-\sqrt{\pi d \alpha N}}$, we have used as successive values of N the successive squares $1, 4, 9, 16, \ldots$. If our error bounds represent the actual behavior of (A) and (B), this should make the errors decrease by a factor of about $e^{\sqrt{\pi d \alpha}}$ as we go from each N to the next. The results of using formula (A) with integrands g_1 through g_4 are in Tables 1a through 4a, respectively; similarly with formula (B) and Tables 1b through 4b.

In each table "maxerr" denotes the maximum of the absolute value of the error of the approximation for those values of s for which the integral was calculated. Those were the 370 values

$$S = -0.999, -0.998, -0.997, \ldots, -0.9, -0.89, -0.88,$$
$$-0.87, \ldots, +0.91, +0.911, +0.912, \ldots, +0.999.$$

TABLE 1a		
N	maxerr	ratio
1	5.92E − 2	
4	5.80E − 3	10.2
9	6.67E − 4	8.7
16	7.58E − 5	8.8
25	8.45E − 6	9.0
36	9.34E − 7	9.0
49	1.03E − 7	9.1
64	1.13E − 8	9.1

TABLE 1b		
N	maxerr	ratio
1	1.08E − 1	
4	1.70E − 2	6.4
9	2.09E − 3	8.1
16	2.36E − 4	8.8
25	2.63E − 5	9.0
36	2.90E − 6	9.1
49	3.18E − 7	9.1
64	3.48E − 8	9.1

TABLE 2a		
N	maxerr	ratio
1	1.67	
4	1.06E − 2	15.8
9	6.01E − 4	17.6
16	3.35E − 5	17.9
25	1.77E − 6	19.0
36	9.10E − 8	19.4
49	4.55E − 9	20.0
64	2.24E − 10	20.3
81	1.08E − 11	20.6
100	5.20E − 13	20.9

TABLE 2b		
N	maxerr	ratio
1	1.67	
4	1.06E − 2	15.8
9	6.01E − 4	17.6
16	3.35E − 5	17.9
25	1.77E − 6	19.0
36	9.10E − 8	19.4
49	4.55E − 9	20.0
64	2.24E − 10	20.3
81	1.08E − 11	20.6
100	5.20E − 13	20.9

TABLE 3a		
N	maxerr	ratio
1	7.82E − 2	
4	2.08E − 3	37.6
9	1.03E − 4	20.2
16	4.67E − 6	22.0
25	2.11E − 7	22.3
36	9.33E − 9	22.6
49	4.11E − 10	22.7
64	1.80E − 11	22.8
81	7.86E − 13	22.9

TABLE 3b		
N	maxerr	ratio
1	7.93E − 2	
4	2.80E − 3	28.3
9	1.42E − 4	19.7
16	6.38E − 6	22.3
25	2.91E − 7	21.9
36	1.28E − 8	22.8
49	5.68E − 10	22.5
64	2.48E − 11	22.9
81	1.09E − 12	22.8

TABLE 4a		
N	maxerr	ratio
1	8.32E − 2	
4	1.42E − 3	58.7
9	4.35E − 5	32.6
16	1.16E − 6	37.7
25	2.95E − 8	39.2
36	7.30E − 10	40.4
49	1.76E − 11	41.5
64	4.20E − 13	41.9

TABLE 4b		
N	maxerr	ratio
1	9.84E − 2	
4	4.74E − 3	20.8
9	2.07E − 4	23.0
16	7.41E − 6	27.9
25	2.39E − 7	31.0
36	7.16E − 9	33.3
49	2.05E − 10	35.0
64	5.64E − 12	36.3

For each value of N, "ratio" denotes the ratio of maxerr for the previous value of N to maxerr for the current value. The calculations were done in double precision on a clone of the IBM PC, using an 8087 coprocessor; thus, about 16-significant-figure arithmetic was used. For each integrand g the results

are presented only for those values of N for which the approximation error was large compared to the likely roundoff error. In the case of integrand g_1 calculation with $N \geq 81$ was prevented by the "blowup" discussed at the end of the last section.

It is clear that formula (A) gave somewhat more accurate results than (B), except for g_2. The tabulated results for g_2 are identical for the two formulas, but that is not because (A) and (B) gave identical errors for all values of s. Instead, each formula had its largest error, for g_2, at $s = 0$; and at $s = 0$ formulas (A) and (B) do give the same approximation when the integrand is an odd function.

For g_1 we took $d = \pi$ and $\alpha = 1/2$, so that the quantity $e^{\sqrt{\pi d \alpha}}$ is 9.2. The values of "ratio" in Tables 1a and 1b are quite close to 9.2 for the higher values of N. This indicates that the convergence behaviors of formulas (A) and (B) are, for this integrand, just what one might expect from the error bounds in our theorems, and that this behavior is seen clearly for moderate values of N.

For g_2, g_3, and g_4 the corresponding values of $e^{\sqrt{\pi d \alpha}}$ are 23.1, 23.1, and 46.9, respectively. (These represent quite rapid convergence!) The results for g_3 fit very nicely, those for g_2 and g_4 somewhat less so. For g_2 this might be expected, because of the infinities of the integrand; for g_4 we can only say that the expected behavior is not yet taking place for the values of N that we have used.

Nevertheless, the convergence rate that we see for g_4 is very fast—considerably faster than for g_3, which is the smoothest possible function. This is in sharp contrast to what classical formulas—say the Gaussian—would do; the singularities of g_4 at ± 1 would very much limit their convergence speed. That points up the fact that it is not, indeed, the nonanalyticity of the integrand at the endpoints that decides the convergence rate of (A) and (B), but only its *growth* as the endpoints are approached. A function such as g_4, which has "half-zeros" at the endpoints, has its integral approximated more accurately than that of a function that is analytic at the endpoints but not zero there.

(For completeness in describing the calculations, we note that for integrands g_1, g_2 and g_3 the function η used in (B) was $(1 + w)/2$, while for g_4 it was $(2 + 3w - w^3)/4$.)

It is interesting to see the actual form of the error function—the true indefinite integral, minus the approximation—for these formulas. Figures 1 and 2 show this function for the case of formula (A) with $N = 25$, for integrands g_2 and g_4 respectively. In all cases the horizontal axis range is from -1 to $+1$; in Figure 1 the vertical range is from -10^{-5} to $+10^{-5}$ and in Figure 2 it is from -10^{-7} to $+10^{-7}$. The curves obtained for the other g's are similar in shape to that in Figure 2.

Figure 3 is another graph of the error function from Figure 2, but now the horizontal axis is scaled according to $\varphi(s)$ instead of s; here, $\varphi(s)$ runs from -20 to 20. This stretches out the ends of the interval $-1 < s < 1$ so that the behavior of the error for s near ± 1 is shown more clearly.

Figure 4 shows the error curve for formula (B) and the integrand g_4; again, N is 25, the vertical range is from -10^{-6} to $+10^{-6}$. The error curve for B and g_2 is not shown as it is very similar to that in Figure 2.
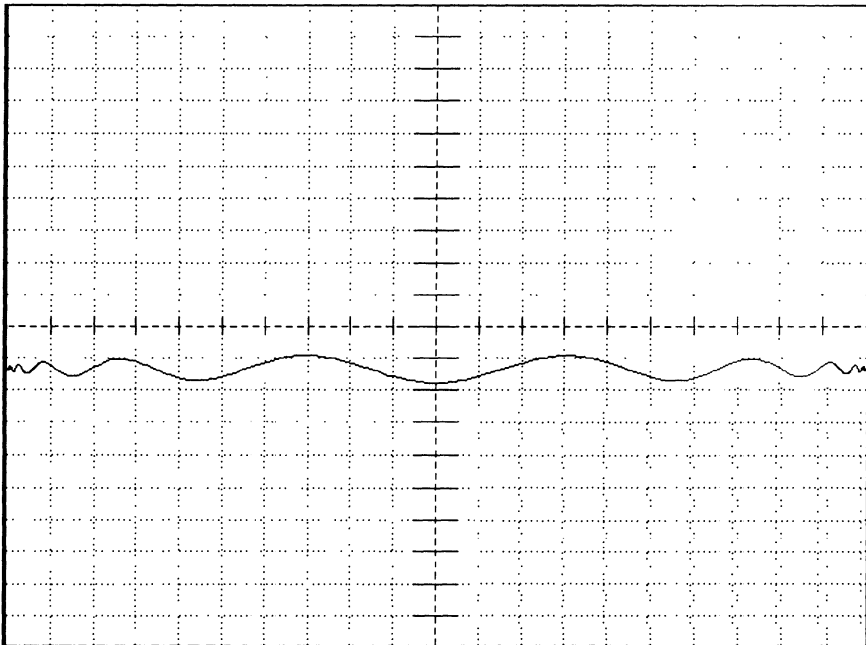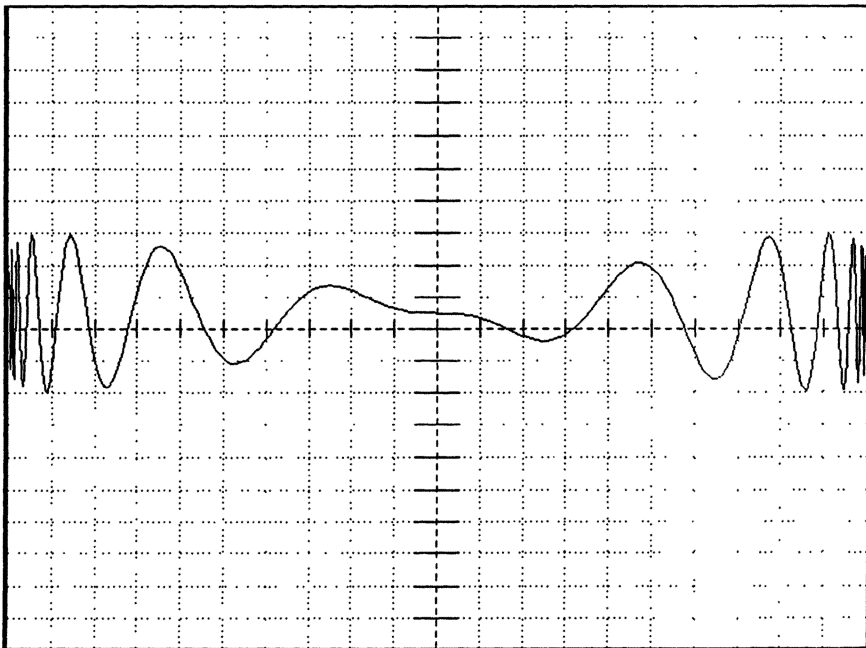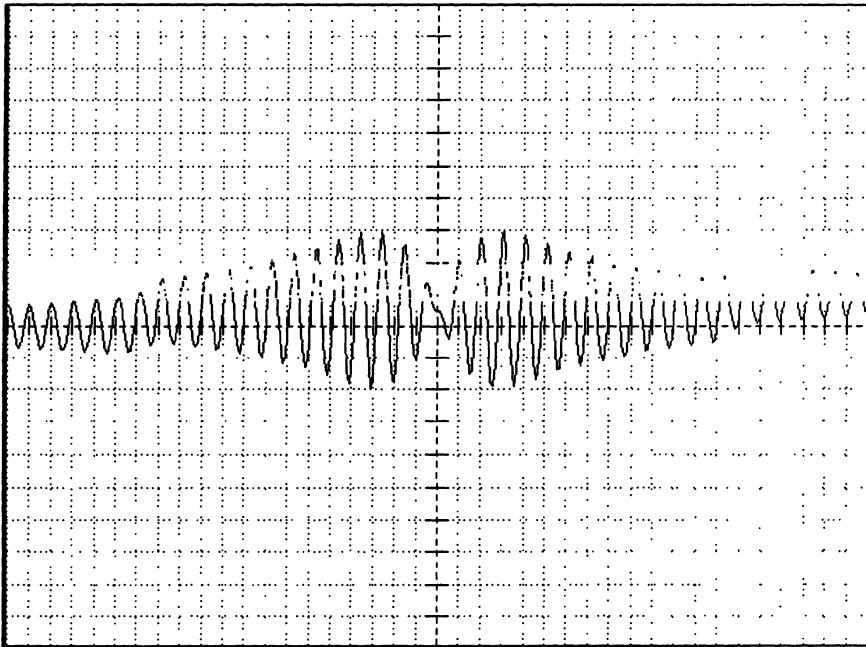
SEYMOUR HABER

FIGURE 1

FIGURE 2

FIGURE 3
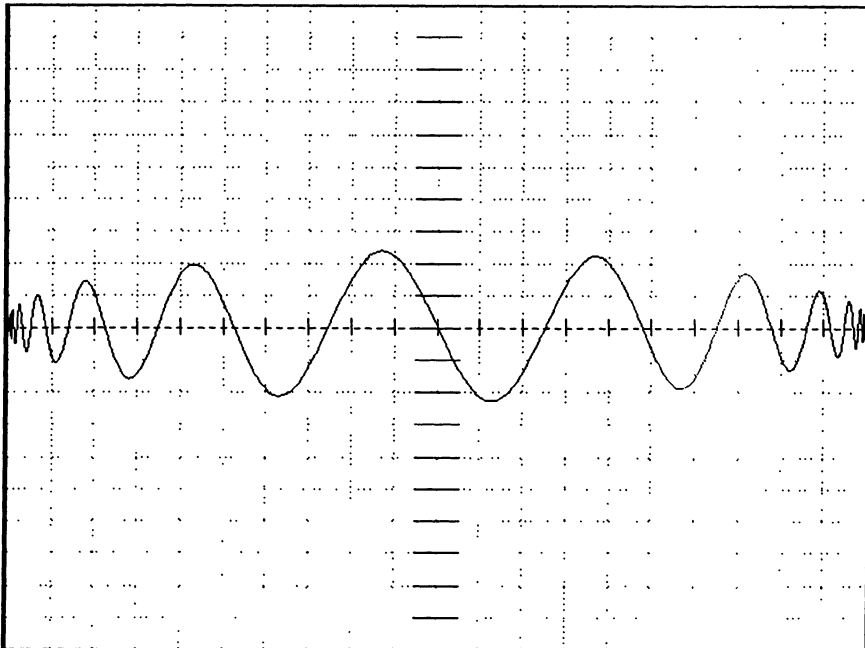
FIGURE 4

APPENDIX

A1. **The calculation of** σ_n . The sine integral is defined as

$$\mathrm{Si}(x) = \int_o^x \frac{\sin t}{t} \, dt \, ,$$

and

$$\sigma_n = \frac{1}{\pi} \, \mathrm{Si}(n\pi).$$

Since $\mathrm{Si}(-x) = -\mathrm{Si}(x)$, we have $\sigma_{-n} = -\sigma_n$ and we need calculate σ_n only for positive n. For large n we propose to use the asymptotic expansion

$$(A1) \qquad \sigma_n \sim \frac{1}{2} + \frac{(-1)^{n+1}}{n\pi^2} \left(1 - \frac{2!}{(n\pi)^2} + \frac{4!}{(n\pi)^4} \cdots \right) ,$$

which is an immediate consequence of the asymptotic expansion for $\mathrm{Si}(x)$ given in [1]. It may be derived directly from the formula

$$\sigma_n = \frac{1}{2} - \frac{1}{\pi} \int_{n\pi}^\infty \frac{\sin t}{t} \, dt$$

by successive integrations by parts, obtaining

$$\sigma_n = \frac{1}{2} + (-1)^{n+1} \left(\frac{1}{n\pi^2} - \frac{2!}{n^3\pi^4} + \cdots + \frac{(-1)^k (2k)!}{n^{2k+1}\pi^{2k+2}} \right.$$
$$\left. + \frac{(-1)^k (2k+1)!}{\pi} \int_{n\pi}^\infty \frac{\cos t}{t^{2k+2}} \, dt \right) .$$

The sign of the last integral is $(-1)^n$, so the remainder term is of opposite sign to the last retained term of the asymptotic series. Thus the error of truncating (A1) is no greater, in absolute value, than the first term not retained. Let us write

$$a_k = \frac{(-1)^{n+k+1} (2k)!}{n^{2k+1}\pi^{2k+2}}.$$

As long as $2k(2k-1) < n^2\pi^2$, we have $|a_k| < |a_{k-1}|$, so that, if we truncate (A1) at the $(K-1)$st term, where

$$K = K(n) = \left[\frac{n\pi}{2} \sqrt{1 + 1/(4n^2\pi^2)} + \frac{1}{4} \right],$$

we will be getting σ_n as accurately as the asymptotic expansion can give it. The error (apart from roundoff) will be no greater than a_K .

TABLE A1

n	$K(n)$	$|a_{K(n)}|$
4	6	$7.82\text{E}-7$
5	8	$3.09\text{E}-8$
6	9	$1.20\text{E}-9$
7	11	$4.81\text{E}-11$
10	15	$3.27\text{E}-15$
12	19	$5.55\text{E}-18$
15	23	$3.99\text{E}-22$
20	31	$5.19\text{E}-29$

Table A1 gives $|a_{K(n)}|$ for several values of n. Using such data, we determine a value n_0 for which the asymptotic expansion will allow the calculation of σ_n to the desired accuracy for $n \geq n_0$. The algorithm for calculating σ_n is completed by adding a table of $\sigma_1, \sigma_2, \ldots, \sigma_{n_0-1}$. Those values may be obtained from Table 2.1 of [7]; or, if more accuracy is needed, they may be calculated by the method of [6] or by numerical integration.

In the calculations described in this paper n_0 was taken to be 12, and 18 terms were used in the series (A1).

A2. Calculation of $\text{Si}(x)$. Since the sine integral is an odd function, we can confine ourselves to positive values of x. For $0 < x < \pi$ the series

$$\text{Si}(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1} x^{2k-1}}{(2k-1)\cdot(2k-1)!}$$

(which is obtained by integrating the series for $(\sin t)/t$) may conveniently be used. Seven terms suffice to give an answer correct to 17 decimal places.

For $x \geq \pi$ we used the following procedure: Set $n = [x/\pi]$ and $\Theta = x - n\pi$. Since

$$(\text{A2}) \qquad \text{Si}(x) = \text{Si}(n\pi) + \int_{n\pi}^{n\pi+\Theta} \frac{\sin t}{t}\, dt = \pi\sigma_n + \frac{(-1)^n}{n\pi} \int_0^{\Theta} \frac{\sin s}{1 + \frac{s}{n\pi}}\, ds,$$

we evaluated the last integral numerically, and evaluated σ_n by the method described above in this appendix. For the numerical integration, we used the 9-point Gauss-Legendre formula. Numerical experiments showed that it will give the integral to 16-decimal-place accuracy for n up to 50, and apparently to 15 decimal places for all n.

BIBLIOGRAPHY

1. M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions*, U.S. Government Printing Office, Washington, DC, 1964, Fomulas 5.2.8, 5.2.34, and 5.2.35.

2. S. Haber, *The tanh rule for numerical integration*, SIAM J. Numer. Anal. **14** (1977), 668–685.

3. R.B. Kearfott, *A Sinc approximation for the indefinite integral*, Math. Comp. **41** (1983), 559–572.

4. J. McNamee, F. Stenger, and E. L. Whitney, *Whittaker's cardinal function in retrospect*, Math. Comp. **25** (1971), 141–154.

5. K. Sikorski and F. Stenger, *Optimal quadratures in H_p spaces*, ACM Trans. Math. Software **10** (1984), 140–151.

6. I. A. Stegun and R. Zucker, *Automatic computing methods for special functions, Part III. The Sine, Cosine, Exponential integrals and related functions*, J. Res. Nat. Bur. Standards B, Mathematical Sciences **80B** (1976), 291–311.

7. F. Stenger, *Numerical methods based on Whittaker cardinal, or Sinc functions*, SIAM Rev. **23** (1981), 165–224.

DEPARTMENT OF MATHEMATICS, TEMPLE UNIVERSITY, PHILADELPHIA, PENNSYLVANIA 19122
E-mail address: shaber@templevm.bitnet