Improvement of selection formulas of mesh size and truncation numbers for the double-exponential formula

Tomoaki Okayama$^1$ and Chisei Kurogi$^2$

$^1$ Graduate School of Information Sciences, Hiroshima City University, 3-4-1 Ozuka-higashi, Asaminami-ku, Hiroshima 731-3194, Japan
$^2$ Energia Communications, Inc., NHK Hiroshima Broadcasting Center Building, 2-11-10 Otemachi, Naka-ku, Hiroshima 730-0051, Japan

E-mail okayama@hiroshima-cu.ac.jp

Received December 12, 2019, Accepted February 10, 2020

Abstract

The double-exponential formula is known as a very efficient quadrature formula. An important point in eliciting its high performance is properly selecting mesh size and truncation numbers depending on a given positive integer. However, the standard selection formulas are not optimal, and there is room for improvement. In this paper, we propose improved selection formulas that reduce the error of the double-exponential formula. We also present a computable error bound of the modified double-exponential formula.

Keywords double-exponential formula, error bound, verified numerical integration

Research Activity Group Quality of Computations

1. Introduction and summary

The double-exponential formula was proposed by Takahashi–Mori [1] to efficiently approximate definite integrals $\int_{a}^{b} f(t) \, dt$. The formula consists of the following two steps. The first is to apply a variable transformation $t = \phi(x)$, which translates the given integral to an integral over the infinite interval $(-\infty, \infty)$ as

$$\int_{a}^{b} f(t) \, dt = \int_{-\infty}^{\infty} f(\phi(x)) \phi'(x) \, dx,$$

where the transformed integrand decays double-exponentially, i.e., $f(\phi(x))\phi'(x) = O(\exp(-\gamma \exp(|x|)))$ with some $\gamma > 0$. Such a transformation is called the double-exponential transformation. The second step is to apply the trapezoidal formula, which is suitably truncated as

$$\int_{-\infty}^{\infty} f(\phi(x))\phi'(x) \, dx \approx h \sum_{k=-\infty}^{\infty} f(\phi(kh))\phi'(kh) \quad (1)$$

and

$$h \sum_{k=-M}^{N} f(\phi(kh))\phi'(kh). \quad (2)$$

Here, the mesh size $h$ and truncation numbers $M$ and $N$ are selected depending on a given positive integer $n$. The convergence rate of the double-exponential formula was analyzed as $O(n \log n)$ with some $c > 0$ in several cases [2], which is quite fast convergence. In addition, by examining the error analysis precisely, a computable error bound was given [3, 4]. The error bound is useful for verified numerical integration [5].

This study improves the selection formulas of $h$, $M$ and $N$. In the previous study, they were selected to enable the discretization error (error of the approximation (1)) and the truncation error (error of the approximation (2)) to become approximately equal. However, there is still a non-negligible gap between the two errors. In this study, we improve the selection formulas of $h$, $M$ and $N$ to reduce the gap and consequently derive a sharper error bound than the existing one. Furthermore, we remove one artificial condition imposed on $n$.

The remainder of this paper is organized as follows. The existing and new error bounds are stated in Section 2. Numerical results are shown in Section 3. Proofs for the presented theorem are provided in Section 4.

2. Existing and new theorems

First, we introduce a function space that supposes integrands translated by the double-exponential transformation. Here, $\mathcal{D}_d$ denotes a strip complex domain defined by $\mathcal{D}_d = \{z \in \mathbb{C} : \text{Im} \, z < d\}$ for $d > 0$.

Definition 1. Let $L$, $R$, $\alpha$ and $\beta$ be positive constants, and $d$ be a constant where $0 < d < \pi/2$. Then, $L_{\alpha,R,\beta}(\mathcal{D}_d)$ denotes a family of functions $F$ that are analytic on $\mathcal{D}_d$, and satisfy for all $z \in \mathcal{D}_d$ and $x \in \mathbb{R}$ that

$$|F(z)| \leq \frac{L \cosh \alpha}{1 + e^{-\pi \sinh \alpha} |z|^\beta},$$

$$|F(x)| \leq \frac{R \cosh x}{(1 + e^{-\pi \sinh x})^\alpha (1 + e^{\pi \sinh x})^\beta}. \quad (3)$$

For convenience, we further define $x_\gamma$ for $\gamma > 0$ by

$$x_\gamma = \begin{cases} \frac{\arcsinh \left(1 + \sqrt{1 + (2\gamma)^2} \right)}{2\pi \gamma}, & (0 < \gamma < \frac{1}{2\pi}) \\ \arcsinh(1), & (\frac{1}{2\pi} \leq \gamma). \end{cases} \quad (4)$$

Then, the existing error bound is described as follows.

- 13 –
Theorem 2 (Okayama [4, Theorem 5.3]) Assume that $F \in \mathbf{L}_{L,R,\alpha,\beta}(\mathcal{D})$. Let $\mu = \min\{\alpha, \beta\}$, let $\nu = \max\{\alpha, \beta\}$, let $h$ be defined as
$$h = \frac{\log(4d\mu/n)}{n},$$
and let $M$ and $N$ be defined as
$$\begin{cases} M = n, & N = n - \frac{\log(\beta/\alpha)}{h} \quad (\text{if } \mu = \alpha), \\ N = n, & M = n - \frac{\log(\alpha/\beta)}{h} \quad (\text{if } \mu = \beta). \end{cases}$$
Furthermore, let $n$ be taken sufficiently large to enable $n \geq (\nu \nu_c)/(4d)$, $Mh \geq x_\alpha$ and $Nh \geq x_\beta$ to hold. Then, it holds that
$$\left| \int_{-\infty}^{\infty} F(x) \, dx - h \sum_{k=-M}^{N} F(kh) \right| \leq \frac{2C}{\pi \mu} e^{-2\pi dn/\log(4d/\mu)} e^{\pi \nu/2} R.$$

The concept of the selection formulas (5) and (6) is described as follows. The discretization error (say $E_D$) is analyzed as $O(e^{-\pi \nu d/h})$ (see also Lemma 4). Furthermore, by setting $M$ and $N$ as (6), the truncation error (say $E_T$) is analyzed as $O(e^{-\pi \nu d/\log(4d/\mu)})$. Finally, by setting $h$ as (5), we have
$$E_D = O(e^{-2\pi dn/\log(4d/\mu)}),$$
$$E_T = O(e^{-2\pi dn}).$$

Although these two orders are approximately equal, the former is lower than the latter. Therefore, the convergence rate in total is analyzed as $O(e^{-2\pi dn/\log(4d/\mu)})$.

By improving the selection formulas of $h$, $M$ and $N$, we show the following theorem. Here, $p(x)$ and $q(x)$ are defined by
$$p(x) = \frac{x}{\arcsinh(x) \arcsinh x},$$
$$q(x) = \frac{x}{\arcsinh x}.$$

Theorem 3 Assume that $F \in \mathbf{L}_{L,R,\alpha,\beta}(\mathcal{D})$. Let $\mu = \min\{\alpha, \beta\}$, let $h$ be defined as
$$h = \frac{\arcsinh(q(2dn/\mu))}{n},$$
and let $M$ and $N$ be defined as
$$\begin{cases} M = n, & N = \frac{\arcsinh\left(\frac{\mu}{2} q(2dn/\mu)\right)}{h} \quad (\text{if } \mu = \alpha), \\ N = n, & M = \frac{\arcsinh\left(\frac{\beta}{2} q(2dn/\mu)\right)}{h} \quad (\text{if } \mu = \beta). \end{cases}$$
Furthermore, let $n$ be taken sufficiently large to enable $Mh \geq x_\alpha$ and $Nh \geq x_\beta$ to hold. Then, it holds that
$$\left| \int_{-\infty}^{\infty} F(x) \, dx - h \sum_{k=-M}^{N} F(kh) \right| \leq \frac{2C}{\pi \mu} e^{-2\pi dn/\arcsinh(2dn/\mu)},$$
where $C$ is a constant defined by $c = 2d/\mu$ as
$$C = 2L \left(1 - e^{-\pi \mu c/2}\right) \cos^{\alpha+\beta/2} \left(\frac{\pi}{2} \sin d\right) d + \frac{\pi}{2} R.$$
that the error bound from Theorem 3 (dashed line) is sharper than that from Theorem 2 (dotted line). We also show observed errors of the existing and new formulas for reference, but we cannot conclude which is better.

4. Proofs

4.1 Sketch of the proof

The general strategy to analyze the error is to split it into two terms as

\[
\left| \int_{-\infty}^{\infty} F(x) \, dx - h \sum_{k=-\infty}^{N} F(kh) \right| \leq \left| \int_{-\infty}^{\infty} F(x) \, dx - h \sum_{k=-\infty}^{M-1} F(kh) \right| + \left| h \sum_{k=-\infty}^{M-1} F(kh) + h \sum_{k=N+1}^{\infty} F(kh) \right|.
\]

The first term and second terms are called the discretization and truncation errors, respectively. The discretization error is bound as follows.

Lemma 4 (Okayama [4, part of Theorem 5.3])
Assume that \( F \in L_{L,R,\alpha,\beta}(\mathcal{D}_d) \). Then, setting \( \mu = \min\{\alpha, \beta\} \), we have

\[
\left| \int_{-\infty}^{\infty} F(x) \, dx - h \sum_{k=-\infty}^{\infty} F(kh) \right| \leq \frac{4L}{\pi \mu} (1 - e^{-2\pi d/h}) \cos \alpha + \beta (\frac{\pi}{2} \sin d) e^{-2\pi d/h}.
\]

We bound the truncation error as follows, which is proved in Section 4.2.

Lemma 5 Assume that \( F \in L_{L,R,\alpha,\beta}(\mathcal{D}_d) \). Let \( \mu = \min\{\alpha, \beta\} \), and let \( M \) and \( N \) be defined as

\[
M = \left\lceil \frac{1}{h} \operatorname{arcsinh} \left( \frac{2dn/\alpha}{\operatorname{arcsinh}(2dn/\mu)} \right) \right\rceil,
\]

\[
N = \left\lceil \frac{1}{h} \operatorname{arcsinh} \left( \frac{2dn/\beta}{\operatorname{arcsinh}(2dn/\mu)} \right) \right\rceil.
\]

Furthermore, let \( Mh \geq x_{\alpha} \) and \( Nh \geq x_{\beta} \) hold. Then, it holds that

\[
h \sum_{k=-\infty}^{-M-1} |F(kh)| + h \sum_{k=N+1}^{\infty} |F(kh)| \leq \frac{2R}{\pi \mu} e^{-2\pi dn/\operatorname{arcsinh}(2dn/\mu)}.
\]

Combining Lemma 4 with Lemma 5, we prove Theorem 3 in Section 4.3.

4.2 Proof of Lemma 5

The following fact is used to prove Lemma 5.

Proposition 6 (Okayama et al. [3, Proposition 4.17])
Let us define \( x_{\gamma} \) by (4). Then, the function \( G_{\gamma}(x) = \cosh(x) e^{\pi \gamma \sinh x} \) is monotonically increasing for \( x \leq -x_{\gamma} \), and the function \( \tilde{G}_{\gamma}(x) = \cosh(x) e^{-\pi \gamma \sinh x} \) is monotonically decreasing for \( x \geq x_{\gamma} \).

Lemma 5 is proved as follows.

Proof Clearly, it holds that

\[
h \sum_{k=-\infty}^{-M-1} |F(kh)| + h \sum_{k=N+1}^{\infty} |F(kh)| \leq h \sum_{k=-\infty}^{\infty} |F(kh)|.
\]

From the inequality (3), regarding the first sum, we have

\[
h \sum_{k=-\infty}^{-M-1} |F(kh)| \leq \frac{R \cosh(kh)}{(1 + e^{-\pi \sinh(kh)})(1 + e^{\pi \sinh(kh)})} \sum_{k=-\infty}^{-M-1} \cosh(kh) e^{\pi \alpha \sinh(kh)}
\]

\[
= Rh \sum_{k=-\infty}^{-M-1} \cosh(kh) e^{\pi \alpha \sinh(kh)}
\]

\[
\leq Rh \sum_{k=-\infty}^{-M-1} \cosh(kh) e^{\pi \alpha \sinh(kh)}
\]

Furthermore, from Proposition 6 and \( Mh \geq x_{\alpha} \), setting
\[ G_{\gamma} = \cosh(x) e^{\gamma_2 \sinh x}, \]

we have
\[
Rh \sum_{k=-\infty}^{-M-1} G_{\alpha}(kh) \leq R \int_{-\infty}^{-Mh} G_{\alpha}(x) \, dx
\]
\[
= \frac{R}{\pi \alpha} e^{-\pi \alpha \sinh(Mh)}
\]
\[
\leq \frac{R}{\pi \mu} e^{-\pi \alpha \sinh(Mh)}.
\]

Similarly, regarding the second sum, we have
\[
h \sum_{k=N+1}^{\infty} \left| F(kh) \right| \leq \frac{R}{\pi \mu} e^{-\pi \beta \sinh(Nh)}.
\]

Finally, using (13), we obtain the desired result.

(QED)

4.3 Proof of Theorem 3

We prepare two propositions for the proof of Theorem 3.

Proposition 7 It holds for \( t \geq 1 \) that
\[
t + 1 \geq \arcsinh \left( \frac{\sinh t}{t} \right) + \frac{t}{\arcsinh(\sinh(t)/t)}.
\]

Proof As
\[
1 \leq \arcsinh \left( \frac{\sinh t}{t} \right) \leq \arcsinh \left( \frac{\sinh t}{1} \right) = t,
\]
we have
\[
\left[ \arcsinh \left( \frac{\sinh t}{t} \right) - 1 \right] \left[ \frac{t}{\arcsinh(\sinh(t)/t)} - 1 \right] \geq 0,
\]
from which we obtain the claim.

(QED)

Proposition 8 Let \( p(x) \) and \( q(x) \) be defined by (7) and (8), respectively. Then, \( p(x) \) and \( r(x) = p(x) - q(x) \) are monotonically increasing functions for \( x \geq 0 \).

Proof We can write \( p(x) \) as \( p(x) = q(q(x)) \arcsinh x \). Using \( q(x) \geq 0 \) and \( q'(x) \geq 0 \) for \( x \geq 0 \), we have
\[
p'(x) = q'(q(x))q'(x) \arcsinh x + \frac{q(q(x))}{\sqrt{1 + x^2}} \geq 0.
\]

Next, we show \( r'(x) \geq 0 \). Set \( g(t) = r'(\sinh t) \). We show \( g(t) \geq 0 \) for \( t \geq 0 \) below, from which we obtain the conclusion as \( r'(x) = g'(\arcsinh x) \sqrt{1 + x^2} \geq 0 \). We start by estimating \( g(t) \) as
\[
g'(t) = \frac{\cosh t}{\arcsinh(\sinh(t)/t)} \cdot \left\{ 1 - \frac{t - \tanh t}{t^2} \left[ \arcsinh \left( \frac{\sinh t}{t} \right) \right] \right. \\
+ \left. \left[ \frac{\sinh t}{t} \right] \frac{t}{\sqrt{1 + (\sinh(t)/t)^2}} \cdot \frac{t}{\arcsinh(\sinh(t)/t)} \right\}
\]
\[
\geq \frac{\cosh t}{\arcsinh(\sinh(t)/t)} (1 - h(t)),
\]
where
\[
h(t) = \frac{t - \tanh t}{t^2} \left[ \arcsinh \left( \frac{\sinh t}{t} \right) + \frac{t}{\arcsinh(\sinh(t)/t)} \right].
\]

In the case where \( 0 \leq t \leq 1 \), using \( t \leq \sinh t \leq t \sinh 1 \) and \( \tanh t \geq t - t^3/3 \), we have
\[
1 - h(t) \geq 1 - \frac{t^3}{3} \left[ \frac{t}{\arcsinh(\sinh(1/t))} \right] > 0.
\]

In the case where \( t \geq 1 \), from Proposition 7, we have
\[
1 - h(t) \geq 1 - \frac{(t - \tanh t)}{t} (t + 1) = \frac{(t + 1) \tanh(t) - t}{t^2} \geq 0,
\]
as \( (t + 1) \tanh(t) - t \geq 0 \) holds for all \( t \in \mathbb{R} \). This completes the proof.

(QED)

Using the above two propositions, we prove Theorem 3 as follows.

Proof Note that (13) is equivalent to (10) if \( h \) is set as (9). Set \( \kappa = \cos^{\alpha+\beta}(\pi/2) \sin d \). From Lemmas 4 and 5, substituting (9) to \( h \), we have
\[
\int_{-\infty}^{\infty} F(x) \, dx - h \sum_{k=-M}^{N} F(kh) \leq \frac{4L}{\pi \mu (1 - e^{-2\pi d/h})} \left( \frac{2R}{\pi \mu} e^{-2\pi d/h} \arcsinh(2dn/\mu) \right) + R = \frac{2}{\pi \mu} \left( \frac{2L e^{-\pi \mu [p(2dn/\mu) - q(2dn/\mu)]}}{(1 - e^{-\pi \mu [p(2dn/\mu)]})} + R \right) e^{-\pi \mu q(2dn/\mu)},
\]
where \( p(x) \) and \( q(x) \) are defined by (7) and (8), respectively. Furthermore, from Proposition 8, it holds that
\[
\frac{2L e^{-\pi \mu [p(2dn/\mu) - q(2dn/\mu)]}}{(1 - e^{-\pi \mu [p(2dn/\mu)]})} \leq \frac{2L e^{-\pi \mu [p(c) - q(c)]}}{(1 - e^{-\pi \mu [p(c)]})}.
\]

which provides the desired result.

(QED)

Acknowledgments

This work was partially supported by the JSPS Grant-in-Aid for Young Scientists (B) Number JP17K14147.

References

[1] H. Takahasi and M. Mori, Double exponential formulas for numerical integration, Publ. RIMS, Kyoto Univ., 9 (1974), 721–741.

[2] K. Tanaka, M. Sugihara, K. Murota and M. Mori, Function classes for double exponential integration formulas, Numer. Math., 111 (2009), 631–655.

[3] T. Okayama, T. Matsuoka and M. Sugihara, Error estimates with explicit constants for Sinc approximation, Sinc quadrature and Sinc indefinite integration, Numer. Math., 124 (2013), 361–394.

[4] T. Okayama, Error estimates with explicit constants for Sinc quadrature and Sinc indefinite integration over infinite intervals, Reliab. Comput., 19 (2013), 45–65.

[5] N. Yanamaka, T. Okayama, S. Oishi and T. Ogita, A fast verified automatic integration algorithm using double exponential formula, NOLTA, IEICE, 1 (2010), 119–132.