Asymptotic Bit Cost of Quadrature Formulas Obtained by Variable Transformation

P. Favati
Istituto di Matematica Computazionale del C.N.R
v. S. Maria 46, 56127 Pisa, Italy
favati@iei.pi.cnr.it

(Received August 1996; revised and accepted January 1997)

Abstract—In this paper, the asymptotic bit operation cost of a family of quadrature formulas, especially suitable for evaluation of improper integrals, is studied. More precisely, we consider the family of quadrature formulas obtained by applying \( k \) times the variable transformation \( x = \sinh(y) \) and then the trapezoidal rule to the transformed integral. We prove that, if the integrand function is analytic in the interior part of the integration interval and approaches zero at a rate which is at least the reciprocal of a polynomial, then the computational bit cost is bounded above by a polynomial function of the number of exact digits in the result. Moreover, disregarding logarithmic terms, the double exponential transformation (\( k = 2 \)) leads to the optimal cost among the methods of this family.

Keywords—Improper integral, Double exponential method, Trapezoidal rule.

1. INTRODUCTION AND PRELIMINARY

The asymptotic bit operation cost of a family of quadrature formulas, very suitable for evaluation of improper integrals, is studied. Without loss of generality, the problem taken into consideration is that of approximating

\[
I = \int_{0}^{\infty} f(x) \, dx,
\]

where \( f \) is an even analytic function on \(( -\infty, +\infty )\); all the other cases of improper integrals with integrand function analytic in the interior part of the integration interval can be reduced to the previous one.

More precisely, the family of quadrature methods under investigation (see [1–3]) is obtained by applying \( k \geq 0 \) times the variable transformation \( x = \sinh(y) \) and then the trapezoidal rule to the integral

\[
I = \int_{0}^{\infty} f_k(y) \, dy,
\]

where

\[
f_k(y) = f_{k-1}(\sinh(y)) \cosh(y) \quad \text{and} \quad f_0(y) = f(y).
\]

The computational cost measure used in this analysis is the number of bit operations as a function of the number of exact decimal digits in the result. As should be expected, the cost of these quadrature methods is characterized by

(i) the order of decay of the integrand function \( f \), when its argument goes to infinity,
(ii) the number \( k \) of variable transformations applied, and
(iii) some parameters associated with the computational cost of the algorithm used for evaluating the integrand function at a point.

If the integrand function approaches zero at a rate which is at least the reciprocal of a polynomial and the pointwise evaluation cost of the integrand function does not increase too much, then the bit cost is upper bounded by a polynomial function of the number of exact digits in the result. Moreover, disregarding logarithmic terms, the double exponential transformation ($k = 2$) leads to the optimal cost among the methods of this family.

In the rest of this section, some notations and basic definitions are introduced.

### 1.1. Notations

Given two real functions $f(x), g(x)$ and a point $x_0 \in \mathbb{R} \cup \{+\infty\}$,

$$f(x) = O(g(x)) \quad \text{means that} \quad \limsup_{x \to x_0} \left| \frac{f(x)}{g(x)} \right| = c, \quad c \in \mathbb{R},$$

$$f(x) = \Omega(g(x)) \quad \text{means that} \quad g(x) = O(f(x)),$$

$$f(x) = \Theta(g(x)) \quad \text{means that} \quad f(x) = O(g(x)) \text{ and } f(x) = \Omega(g(x)),$$

$$f(x) = O_{\epsilon}(g(x)) \quad \text{means that} \quad f(x) = O\left(g(x)^{1+\epsilon}\right), \text{ for any } \epsilon > 0.$$

In the following, the point $x_0$ will be implicitly defined by the context.

### 1.2. Bit-Operation Cost of Function Evaluation at a Point

Let $f(x)$ be a continuous function defined on an interval $\Gamma \subset \mathbb{R}$. We suppose that there is an algorithm $\varphi$ which, for any integer $d$ and any point $x \in \Gamma$, computes an approximation $\varphi(x, d)$ of $f(x)$ with accuracy $d$, that is,

$$|f(x) - \varphi(x, d)| < e^{-d}. \quad (2)$$

Let $B_{\varphi}(x, d)$ be the pointwise bit-operations cost of computing $\varphi(x, d)$ at $x \in \Gamma$. We consider the class of functions for which the cost $B_{\varphi}(x, d)$ can be bounded by a product of two separate factors, the first one depending only on the algorithm $\varphi$ and the point $x$, and the second one depending only on $d$.

**Definition 1.** [4] The algorithm $\varphi$ is said to be $\alpha$-separable if, for any $x \in \Gamma$ and any $\epsilon > 0$, its cost can be upper bounded in the following way:

$$B_{\varphi}(x, d) \leq C_{\varphi}(x) d^{\alpha + \epsilon}, \quad (3)$$

where $C_{\varphi}(x)$ is a positive valued function independent of $d$.

Bounds on the cost having a simple form like (3) are very useful when estimating the average cost of function evaluations, as in our case. In [4,5], it has been shown that the classical algorithms for the evaluation of the most popular elementary functions are $\alpha$-separable, with $\alpha = 1$. In the following sections, we will use the lower bounds on the function $C_{\varphi}(x)$ given in the following theorem.

**Theorem 1.** [4] If $f(x) \in C^1(\Gamma)$ and $\varphi$ is an $\alpha$-separable algorithm for evaluating $f$, then

$$C_{\varphi}(x) \geq 1 + \max \{0, 1 + \log |xf'(x)|\}, \quad (4)$$

$$C_{\varphi}(x) \geq 1 + \max \{0, 1 + \log |f(x)|\}, \quad (5)$$

$$\alpha \geq 1, \quad (6)$$

for any $x$ such that $f(x) \neq 0$ and $xf'(x) \neq 0$. 

**DEFINITION 2.** Given a measure μ on Γ, the average cost of φ on A ⊆ Γ with respect to μ is the integral

$$B_φ(A, d) = \frac{1}{μ(A)} \int_A B_φ(x, d) dμ,$$

where

$$μ(A) = \int_A dμ.$$

**1.3. Order of Decay of Integrand Function**

We introduce the following sequence of functions having different rate of growth, when their argument goes to infinity:

$$Q_0(x) = x;$$

$$Q_{i+1}(x) = e^{Q_i(x)};$$

$$Q_{i-1}(x) = \log Q_i(x); \quad i \in \mathbb{Z}.$$  

This sequence of functions is used to characterize the order of decay of the integrand function.

**DEFINITION 3.** We say that a function f, integrable in (0, ∞), has order of decay i if i is the largest integer for which a function F(x) exists such that

$$F(x) = O\left(Q_i(x)^{-b}\right), \quad b > 0, \quad f(x) = O\left(F'(x)\right),$$

and F'(x) is nonnegative and monotone.

**2. BIT COMPLEXITY OF THE TRAPEZOIDAL RULE APPLIED TO APPROXIMATE INFINITE INTEGRALS**

In this section, we study the bit complexity of the direct application of the trapezoidal rule to a function f satisfying the following assumptions.

**ASSUMPTION H1.** f is an even function defined on (−∞, +∞), which is analytic in the infinite strip |Im(z)| < s.

**ASSUMPTION H2.** |f(x)| is integrable in [0, +∞), with decay i.

**ASSUMPTION H3.** There exists an α-separable algorithm φ, which approximates f, for which the function C_φ(x) is integrable in [0, A] for any A.

**ASSUMPTION H4.** C_φ(x) is an increasing function and there exist r and β > 0 such that C_φ(x) = O_λ(Q_r(x)^β) (that is, we suppose that the difficulty of computing φ increases with a known rate when x goes to infinity).

**2.1. Error Analysis**

The exact integral (1) can be approximated by a trapezoidal rule with mesh size h, introducing the total error

$$E_{tot} = I - I_{computed} = E_{trap} + E_{trim} + E_f + E_{round},$$

where

$$E_{trap} = \int_0^∞ f(x) dx - h \sum_{i=0}^{+∞} f(hi)$$

is the discretization error,

$$E_{trim} = h \sum_{hi > U} f(hi)$$

is the trimming error,

$$E_f = h \sum_{hi < U} f(hi) - h \sum_{hi < U} \phi(hi, d)$$

is the function evaluation error, and, finally,

$$E_{round} = h \sum_{hi < U} \phi(hi, d) - I_{computed}$$

is the roundoff error.
with \( I_{\text{computed}} \) denoting the summation performed by using an arithmetic with relative precision \( \epsilon_1 \).

In order to have \( |I - I_{\text{computed}}| = O(e^{-D}) \), we require that each error is of order less than or equal to \( e^{-D} \). Since \( |E_{\text{trap}}| = \Theta(e^{-2\pi s/h}) \), see [6, p. 211], we can choose

\[
h = \frac{2\pi s}{D}, \tag{7}
\]

and denoting by \( N \) the number of function evaluations, obtain

\[
N = \Theta \left( \frac{U}{h} \right) = \Theta(DU).
\]

Moreover, since

\[
|E_{\text{round}}| \leq N \epsilon_1 h \sum_{hi < U} |\varphi(hi, d)| \approx N \epsilon_1 \int_0^U |f(x)| \, dx
\]

and

\[
|E_f| = h \left| \sum_{hi < U} (f(hi) - \varphi(hi, d)) \right| \leq h \sum_{hi < U} |f(hi) - \varphi(hi, d)| \leq U e^{-d},
\]

it is sufficient to choose \( -\log \epsilon_1 = \Theta(D + \log U) \) and \( d = \Theta(D + \log U) \) to have \( E_{\text{round}} \) and \( E_f \) of order not greater than \( e^{-D} \). By considering the order of decay of the integrand function, we detect a value for \( U \) which guarantees \( |E_{\text{trim}}| = O(e^{-D}) \). From Definition 3, if \( f \) has decay \( i \), then a constant \( c \) exists such that

\[
|E_{\text{trim}}| = h \left| \sum_{hi > U} f(hi) \right| \leq h \sum_{hi > U} |f(hi)| \leq c h \sum_{hi > U} F'(hi) \leq c \int_U^\infty F'(x) \, dx = -cF(U)
\]

with \( F(x) = O(Q_i(x)^{-b}) \), for some \( b > 0 \). Hence,

\[
U = \Theta(Q_{1-i}(\lambda D)), \quad \lambda > \frac{1}{b} \tag{8}
\]

ensures \( |E_{\text{trim}}| = O(e^{-D}) \). We have disregarded the inherent error, assuming that the points \( hi \) are exactly represented for \( i = 0, 1, \ldots, N/h \). If \( h \) has the form \( h = 2^{-n} \) with \( n \geq \log D - \log(2\pi s) \), in order to satisfy (7), the points \( hi \) for \( i = 0, 1, \ldots, N/h \) can be exactly represented by using \( \log(N + n) \) digits. This cost does not affect the total computational bit cost.

2.2. Bit Complexity

In the previous section, we gave estimates of \( h, N, \epsilon_1, d, \) and \( U \) as functions of \( D \) and of the decay parameter \( i \), so that the error was of order less than or equal to \( e^{-D} \). Now we are ready to state the following theorem.

**Theorem 2.** Let \( f \) be a function satisfying Assumptions H1–H3, and let \( H_f(U) = \int_0^U C_\varphi(x) \, dx \). Let \( BC(D) \) be the bit-operation cost of approximating \( I = \int_0^\infty f(x) \, dx \) by the trapezoidal rule with an absolute error \( |E_{\text{tot}}| = O(e^{-D}) \). Then the relation \( BC(D) = O_{\epsilon}(D^{1+\alpha}H_f(Q_{1-i}(\lambda D))) \) holds.

**Proof.** The bit-operation cost \( BC(D) \) consists of two parts: a combinatory cost due to the sum and an information cost due to the function evaluations. Since for the trapezoidal rule the density of function evaluation is constant, by using Definition 2 with the Lagrange measure, we have

\[
\overline{B_\varphi([0, U], d)} = \frac{1}{U} \int_0^U B_\varphi(x, d) \, dx \leq \frac{1}{U} H_f(U)d^{n+\epsilon}
\]
Asymptotic Bit Cost

for any $\epsilon > 0$. By taking into account the $O_\epsilon$ notation, the inequality $H_f(U) \geq U$, and the previous section, we conclude that

$$BC(D) = N \left[ -\log \epsilon + B_\varphi([0, U], d) \right] = O \left( DU \left[ D + \log U + \frac{H_f(U)}{U} [D + \log U]^{\alpha + \epsilon} \right] \right)$$

$$= O \left( D[D + Q_{-i}(\lambda D)]^{\alpha + \epsilon} H_f(Q_{1-i}(\lambda D)) \right) = O_\epsilon \left( D^{1+\alpha} H_f(Q_{1-i}(\lambda D)) \right).$$

In the following, we consider the behaviour of $BC(D)$ with respect to different growths of $C_\varphi(x)$; by taking Assumption H4 into account, that is, $C_\varphi(x) = O_{\varphi}(Q_r(x)^\beta)$, $\beta > 0$, we obtain the following relations:

$$H_f(Q_{1-i}(\lambda D)) = \int_0^{Q_{1-i}(\lambda D)} C_\varphi(x) \, dx = O_\epsilon \left( \int_0^{Q_{1-i}(\lambda D)} Q_r(x)^\beta \, dx \right)$$

$$= \left\{ \begin{array}{ll}
O_\epsilon(Q_{1-i}(\lambda D)), & r < 0, \\
O_\epsilon(Q_{1-i}(\lambda D)^{\beta+1}), & r = 0, \\
O_\epsilon(Q_{r-i+1}(\lambda D)^{\beta}), & r > 0.
\end{array} \right.$$

By selecting the cases in which the upper bound on the bit cost is polynomial in $D$, we state the following result.

**Theorem 3.** Let $f$ be a function satisfying Assumptions H1, H2, H3, H4. If $i \geq 1 + \max(0, r)$, then the cost $BC(D)$ associated to the trapezoidal rule is polynomial; more precisely,

$$BC(D) = \left\{ \begin{array}{ll}
O_\epsilon(D^{\alpha+1}), & i > 1 + \max(0, r), \\
O_\epsilon(D^{\alpha+2}), & i = 1, \ r < 0, \\
O_\epsilon(D^{\alpha+\beta+1}), & i = r + 1 > 1, \\
O_\epsilon(D^{\alpha+\beta+2}), & i = r + 1 = 1.
\end{array} \right. \quad (9)$$

### 3. BIT COMPLEXITY OF TRAPEZOIDAL RULE APPLIED AFTER SINH-TRANSFORMATIONS

In this section, we derive upper bounds on the bit cost $BC(D)$ when the trapezoidal rule is applied after a number $k$ of sinh-transformations. We first consider the function $f_1(x) = f(\sinh(x)) \cosh(x)$. If $f(x)$ satisfies Assumptions H1–H4, we can derive some properties for $f_1$. It is easy to prove that $f_1$ is integrable in $[0, \infty)$, with decay $i + 1$. Moreover, there exist a 1-separable algorithm $\phi$ for the chopped down approximation of $\sinh(x)$ with $C_\phi(x) = (1 + \max\{0, \log |xe^{\epsilon}|\})^{i+\epsilon} = O_\epsilon(x)$ and a 1-separable algorithm $\psi$ for the approximation of $\cosh(x)$ with $C_\psi(x) = O_\epsilon(x)$ (see [4]). In order to have

$$|f(\sinh(x)) \cosh(x) - \varphi(\phi(x, d_1), d_2) \psi(x, d_3) (1 + \epsilon)| = O \left( e^{-D} \right),$$

we choose $d_1 = d + \log(|f'(\eta)| \cosh(x))$, where $|f'(\eta)| = \max_{\sinh(x)=1/2} |f'(t)|$, $d_2 = d + x$, $d_3 = d$, and $-\log \epsilon = d$. In this way, we have an algorithm, say $\varphi_1$, which approximates $f_1$ with total cost

$$B_{\varphi_1}(x, d) = (d + x)^{\alpha+\epsilon} C_\varphi(\phi(x, d_1)) + C_\phi(x) [d + \log(|f'(\eta)| \cosh(x))]^{1+\epsilon} + C_\psi(x) d^{1+\epsilon} + d^{1+\epsilon}.$$

By the inequality (4) of Theorem 1 and Assumption H4, using the $O_\epsilon$ notation, we get

$$B_{\varphi_1}(x, d) = O_\epsilon \left( [C_\phi(x)]^{\alpha} + x (1 + \max\{0, \log(|f'(\eta)| \cosh(x))\}) \alpha \right)$$

$$= O_\epsilon \left( [C_\phi(x)]^{\alpha} \right).$$
Hence, we can conclude that the algorithm \( \varphi_1 \) is \( \alpha \)-separable with function

\[
C_{\varphi_1}(x) = O_\varepsilon \left( C_\varphi \left( e^x \right) x^\alpha \right) = O_\varepsilon \left( Q_\rho \left( e^x \right) x^\beta \right).
\]

In this case, we have \( U = \Theta(Q_{-\lambda}(\lambda D)) \), \( \lambda > 1/b \), and

\[
H_{f_1} \left( Q_{-\lambda}(\lambda D) \right) = \int_0^{Q_{-\lambda}(\lambda D)} C_{\varphi_1}(x) \, dx = \begin{cases} 
O_\varepsilon \left( Q_{-\lambda}(\lambda D)^{\alpha+1} \right), & r < -1, \\
O_\varepsilon \left( Q_{-\lambda}(\lambda D)^{\alpha+\beta+1} \right), & r = -1, \\
O_\varepsilon \left( Q_{-\lambda+1}(\lambda D)^\beta \right), & r \geq 0.
\end{cases}
\]

More in general, let \( f_k(x) = f_{k-1}(\sinh(x)) \cosh(x) \), \( k > 1 \), then \( f_k \) has decay \( i+k \) and an \( \alpha \)-separable algorithm \( \varphi_k \) exists with

\[
C_{\varphi_k}(x) = O_\varepsilon \left( C_{\varphi_{k-1}} \left( e^x \right) x^\alpha \right) = O_\varepsilon \left( Q_{k-1}(x)^\alpha Q_{r+k}(x)^\beta \right).
\]

The following relations hold: \( U = \Theta(Q_{-\lambda-\kappa}(\lambda D)) \), \( \lambda > 1/b \), and

\[
H_{f_k} \left( Q_{-\lambda-\kappa}(\lambda D) \right) = \int_0^{Q_{-\lambda-\kappa}(\lambda D)} C_{\varphi_k}(x) \, dx = \begin{cases} 
O_\varepsilon \left( Q_{-\lambda-\kappa}(\lambda D)^\alpha \right), & r < -1, \\
O_\varepsilon \left( Q_{-\lambda-\kappa}(\lambda D)^{\alpha+\beta} \right), & r = -1, \\
O_\varepsilon \left( Q_{-\lambda-\kappa+1}(\lambda D)^\beta \right), & r \geq 0.
\end{cases}
\]

We can summarize these results in the following.

**Theorem 4.** Let \( f \) be a function satisfying Assumptions H1–H4. If \( i > \max(0, r+1) \), then the cost \( BC(D) \) of applying the trapezoidal rule after \( k \) sinh-transformations is polynomial. More precisely,

\[
BC(D) = \begin{cases} 
O_\varepsilon \left( D^{\alpha+1} \right), & i > \max(0, r+1), \\
O_\varepsilon \left( D^{\alpha+\beta+1} \right), & i = r+1 \geq 1, \\
O_\varepsilon \left( D^{2\alpha+p} \right), & i = 0, \ r < -1, \\
O_\varepsilon \left( D^{2\alpha+\beta+p} \right), & i = 0, \ r = -1,
\end{cases}
\]

where

\[
p = \begin{cases} 
2, & \text{if } k = 1, \\
1, & \text{if } k \geq 2.
\end{cases}
\]

Comparing (9) and (10), we can draw some conclusions.

- Whatever the number of sinh-transformations, if \( i \geq \max(0, r+1) \), the bit complexity is polynomial. Otherwise, the upper bound on the bit complexity is exponential. This means that, even if the algorithm for the function evaluation is not asymptotically expensive \( (r < 0) \), the integrand functions with decay less than polynomial might not be tractable with this family of methods.
- With respect to the \( O_\varepsilon \) notation (which in our case hides logarithmic terms), \( k = 2 \) leads to the optimal bit complexity bound for any \( i, r, \alpha, \) and \( \beta \); that is, in terms of asymptotic cost, the double exponential transformation is the optimal one (it is the most suitable method when there is no information about the integrand function).
- In particular, if \( i = 0 \) (respectively, \( i = 1 \) and \( i > 1 \)), the value of \( k \) which gives the optimal bound on \( BC \) is \( k = 2 \) (respectively, \( k = 1 \) and \( k = 0 \)).

The problem of integrating an analytic function on a general interval can be reduced to the problem analyzed above by applying a preliminary variable transformation. For example, consider a function \( f \) defined on \((-1,1)\) with the following properties:

1. \( f \) is an even analytic function in \((-1,1)\);
2. \( |f(x)| \) is integrable in \([0,1]\);
3. there exists an algorithm \( \varphi \) \( \alpha \)-separable with \( B_\varphi(x,d) \leq C_\varphi(x)d^{\alpha+\varepsilon} \).
To integrate $f$ on $[0,1)$ amounts to integrating the even function $g(x) = (\sqrt{1+x^2} - \frac{x}{\sqrt{1+x^2}})^{-3} f(x)$ on $[0,\infty)$. It is possible to derive an algorithm $\tau$ for computing the function $g$ which is $\alpha$-separable. Moreover, if $C_\varphi(1 - 1/x^2) = O_\epsilon(Q_\tau(x)^\beta)$, for suitable constants $r$ and $\beta > 0$, then the relation $C_\tau(x) = O_\epsilon(Q_\tau(x)^\beta \log x) = O_\epsilon(Q_\tau(x)^\beta')$ holds. Finally, if the function $x^{-3}f(1 - 1/x^2)$ has order of decay $i$ when $x$ goes to $+\infty$, then $g(x)$ has order of decay $i$ as well. Hence, we can apply the previous analysis to the even function $g(x)$ integrated in $[0, +\infty)$ with parameters $\alpha, r', \beta', i$.

REFERENCES