A BOUND FOR THE ERROR TERM IN THE BRET-MCMILLAN ALGORITHM

RICHARD P. BRENT AND FREDRIK JOHANSSON

Abstract. The Brent-McMillan algorithm B3 (1980), when implemented with binary splitting, is the fastest known algorithm for high-precision computation of Euler’s constant. However, no rigorous error bound for the algorithm has ever been published. We provide such a bound and justify the empirical observations of Brent and McMillan. We also give bounds on the error in the asymptotic expansions of functions related to the Bessel functions $I_0(x)$ and $K_0(x)$ for positive real $x$.

1. Introduction

Brent and McMillan [3, 5] observed that Euler’s constant

$$
\gamma = \lim_{n \to \infty} (H_n - \ln(n)) \approx 0.5772156649, \quad H_n = \sum_{k=1}^{n} \frac{1}{k},
$$

can be computed rapidly to high accuracy using the formula

$$
\gamma = \frac{S_0(2n) - K_0(2n)}{I_0(2n)} - \ln(n),
$$

where $n > 0$ is a free parameter (understood to be an integer), $K_0(x)$ and $I_0(x)$ denote the usual Bessel functions, and

$$
S_0(x) = \sum_{k=0}^{\infty} \frac{H_k}{(k!)^2} \frac{x^{2k}}{2^k}.
$$

The idea is to choose $n$ optimally so that an asymptotic series can be used to compute $K_0(2n)$, while $S_0(2n)$ and $I_0(2n)$ are computed using Taylor series.

When all series are evaluated using the binary splitting technique (see [4, §4.9]), the first $d$ digits of $\gamma$ can be computed in essentially optimal time $O(d^{1+\varepsilon})$. This approach has been used for all recent record calculations of $\gamma$, including the current world record of 29,844,489,545 digits set by A. Yee and R. Chan in 2009 [9].
Brent and McMillan gave three algorithms (B1, B2 and B3) to compute $\gamma$ via (1.1). The most efficient, B3, approximates $K_0(2n)$ using the asymptotic expansion

$$2xI_0(x)K_0(x) = \sum_{k=0}^{m/2-1} \frac{b_k}{x^{2k}} + T_m(x), \quad b_k = \left[\frac{(2k)!}{(k)!^4 8^k}\right]^3,$$

where one should take $m \approx 4n$. The expansion (1.2) appears as formula 9.7.5 in Abramowitz and Stegun [1], and 10.40.6 in the Digital Library of Mathematical Functions [7]. Unfortunately, neither work gives a proof or reference, and no bound for the error term $T_m(x)$ is provided. Brent and McMillan observed empirically that $T_{4n}(2n) = O(e^{-4n})$, which would give a final error of $O(e^{-8n})$ for $\gamma$, but left this as a conjecture.

Brent [2] recently noted that the error term can be bounded rigorously, starting from the individual asymptotic expansions of $I_0(x)$ and $K_0(x)$. However, he did not present an explicit bound at that time. In this paper, we calculate an explicit error bound, allowing the fastest version of the Brent-McMillan algorithm (B3) to be used for provably correct evaluation of $\gamma$.

To bound the error in the Brent-McMillan algorithm we must bound the errors in evaluating the transcendental functions $I_0(2n)$, $K_0(2n)$ and $S_0(2n)$ occurring in (1.1) (we ignore the error in evaluating $\ln(n)$ since this is well-understood).

The most difficult task is to bound the error associated with $K_0(2n)$. For reasons of efficiency, the algorithm approximates $I_0(2n)$ using the asymptotic expansion (1.2), and then the term $K_0(2n)/I_0(2n)$ in (1.1) is computed from $I_0(2n)/K_0(2n)^2$.

Sections 2–3 contain bounds on the size of various error terms that are needed for the main result. For example, Lemma 2.1 bounds the error in the asymptotic expansion for $I_0(x)$, which is nontrivial as the terms do not have alternating signs.

The asymptotic expansion (1.2) can be obtained formally by multiplying the asymptotic expansions for $K_0(x)$ and $I_0(x)$ occurring in (1.1) (we ignore the error in evaluating $\ln(n)$ since this is well-understood). To obtain $m$ terms in the asymptotic expansion, we multiply the polynomials $P_m(-1/z)$ and $P_m(1/z)$ occurring in (2.1)–(2.2), then discard half the terms (here $z = 1/x$ is small when $x \approx 2n$ is large, so we discard the terms involving high powers of $z$). To bound the error, we show in Lemma 3.1 that the discarded terms are sufficiently small, and also take into account the error terms $R_m$ and $Q_m$ in the asymptotic expansions for $K_0$ and $I_0$.

The main result, Theorem 4.1, is given in Section 4. Provided the parameter $N$ (the number of terms used to approximate $S_0(2n)$ and $I_0(2n)$) is sufficiently large, the error is bounded by $24e^{-8n}$. Corollary 4.3 shows that it is sufficient to take $N \approx 4.971n$.

2. Bounds for the Individual Bessel Functions

Asymptotic expansions for $I_0(x)$ and $K_0(x)$ are given by Olver [8, pp. 266–269] and can be found in [7, §10.40]. They can be written as

$$K_0(x) = e^{-x} \left( \frac{\pi}{2x} \right)^{1/2} (P_m(-x) + R_m(x)),$$

where
and

\begin{equation}
I_0(x) = \frac{e^x}{(2\pi x)^{1/2}} (P_m(x) + Q_m(x)),
\end{equation}

where \( R_m(x) \) and \( Q_m(x) \) denote error terms,

\begin{equation}
P_m(x) = \sum_{k=0}^{m-1} a_k x^{-k}, \quad \text{and} \quad a_k = \frac{[(2k)!]^2}{(k!)^3 32^k}.
\end{equation}

For \( n \geq 1, \)

\begin{equation}
\sqrt{2\pi n^{n+1/2} e^{-n}} \leq n! \leq e n^{n+1/2} e^{-n},
\end{equation}

so the coefficients \( a_k \) in (2.3) satisfy

\begin{equation}
a_k \leq \frac{e^2}{\pi^{3/2} 2^{1/2} k^{1/2}} \left( \frac{k}{2e} \right)^k \left( \frac{k}{2e} \right)^k
\end{equation}

for \( k \geq 1 \) (the first term is \( a_0 = 1 \)).

For \( x > 0, \) we also have the global bounds

\begin{equation}
0 < K_0(x) < e^{-x} \left( \frac{\pi}{2x} \right)^{1/2}
\end{equation}

and

\begin{equation}
I_0(x) > \frac{e^x}{(2\pi x)^{1/2}}.
\end{equation}

Observe that the bound on \( K_0(x) \) and equation (2.1) imply that

\begin{equation}
|P_m(-x) + R_m(x)| < 1.
\end{equation}

For \( x > 0, \) the series (2.1) for \( K_0(x) \) is alternating, and the remainder satisfies

\begin{equation}
|R_m(x)| \leq \frac{a_m}{x^m} \leq \frac{1}{m^{1/2}} \left( \frac{m}{2e} \right)^m \frac{1}{x^m}.
\end{equation}

The series (2.2) for \( I_0(x) \) is not alternating. The following lemma bounds the error \( Q_m(x). \)

**Lemma 2.1.** Let \( Q_m(x) \) be defined by (2.2). Then for \( m \geq 1 \) and real \( x \geq 2 \) we have

\[ |Q_m(x)| \leq 4 \left( \frac{m}{2e x} \right)^m + e^{-2x}. \]

**Proof.** The identity \( I_0(x) = i(K_0(xe^{\pi i}) - K_0(x))/\pi \) (see [10.34.5]) gives

\[ Q_m(x) = R_m(xe^{\pi i}) - i \left( \frac{2\pi x}{e} \right)^{1/2} K_0(x). \]

According to Olver [8, p. 269],

\[ |R_m(xe^{\pi i})| \leq 2 \chi(m) \exp \left( \frac{1}{8} \pi x^{-1} \right) a_m x^{-m}, \]

where

\[ \chi(m) = \pi^{1/2} \frac{\Gamma(m/2 + 1)}{\Gamma(m/2 + 1/2)} \leq \frac{\pi}{2} m^{1/2} \]

(the bound on \( \chi(m) \) follows as \( \chi(m)/m^{1/2} \) is monotonic decreasing for \( m \geq 1 \)).

Since \( x \geq 2, \) applying (2.5) gives

\[ |R_m(xe^{m})| \leq \pi e^{\pi/16} \left( \frac{m}{2e x} \right)^m \frac{1}{x^m} < 4 \left( \frac{m}{2e x} \right)^m. \]
Combined with the global bound $\frac{1}{2}2^n$ for $K_0(x)$, we obtain
\begin{equation}
|Q_m(x)| \leq |R_m(xe^{\pi i})| + \frac{1}{\pi} \left(2\pi x\right)^{1/2} e^{x} K_0(x) \leq 4 \left(\frac{m}{2e^x}\right)^m e^{-2x}.
\end{equation}
}\hfill\square

**Corollary 2.2.** For $x \geq 2$, we have $0 < I_0(x) K_0(x) < 1/x$.

**Proof.** The first inequality is obvious, since both $I_0(x)$ and $K_0(x)$ are positive. Also, using (2.2) and (2.10) with $m = 1$ gives
\[ I_0(x) \leq \frac{e^x}{(2\pi x)^{1/2}} (1 + e^{-1} + e^{-4}), \]
so from (2.6) we have
\[ I_0(x) K_0(x) \leq \frac{1 + e^{-1} + e^{-4}}{2x} < \frac{1}{x}. \]\hfill\square

**Lemma 2.3.** If $R_m(x)$ and $Q_m(x)$ are defined by (2.1) and (2.2) respectively, then
\begin{equation}
|R_{4n}(2n)| \leq \frac{e^{-4n}}{2n^{1/2}} \text{ and } |Q_{4n}(2n)| \leq 5e^{-4n}.
\end{equation}

**Proof.** Taking $x = 2n$ and $m = 4n$, the inequality (2.9) gives the first inequality, and Lemma 2.1 gives the second inequality. \hfill\square

We also need the following lemma.

**Lemma 2.4.** If $P_m(x)$ is defined by (2.3), then
\begin{equation}
|P_{4n}(2n)| < 2 \text{ and } |P_{4n}(-2n)| < 1.
\end{equation}

**Proof.** Using (2.3) and (2.5), we have
\[ P_{4n}(2n) = 1 + \sum_{k=1}^{4n-1} \frac{a_k}{(2n)^k} \leq 1 + \sum_{k=1}^{4n-1} k^{-1/2} \left(\frac{k}{4en}\right)^k \leq 1 + \sum_{k=1}^{4n-1} e^{-k} < \frac{e}{e-1} < 2. \]
The right inequality in (2.12) can be proved in a similar manner, taking the sign alternations into account. \hfill\square

### 3. Bounds for the product

We wish to bound the error term $T_m(x)$ in (1.2) when evaluated at $x = 2n$, $m = 4n$. The result is given by the following lemma.

**Lemma 3.1.** If $T_m(x)$ is defined by (1.2), then $T_{4n}(2n) < 7e^{-4n}$. 

Proof. In terms of the expansions for $I_0(x)$ and $K_0(x)$, we have

$$2xI_0(x)K_0(x) = (P_m(-x) + R_m(x))(P_m(x) + Q_m(x))$$

(3.1) $$= P_m(x)P_m(-x) + [P_m(-x) + R_m(x)]Q_m(x) + P_m(x)R_m(x).$$

It follows from (2.8), (2.11) and (2.12) that the expression $[\cdots]$ in (3.1), evaluated at $x = 2n$, $m = 4n$, is bounded in absolute value by

$$5e^{-4n} + e^{-4n}/n^{1/2} \leq 6e^{-4n}.$$ 

(3.2)

Next, we rewrite

$$P_m(x)P_m(-x) = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} (-1)^ia_ia_jx^{-(i+j)}$$
as $L + U$, where

$$L = \sum_{k=0}^{m-1} \left( \sum_{j=0}^{k} (-1)^ja_ja_{k-j} \right) x^{-k}$$

(3.3) and

$$U = \sum_{k=m}^{2m-2} \left( \sum_{j=k-(m-1)}^{m-1} (-1)^ja_ja_{k-j} \right) x^{-k}.$$ 

(3.4)

The "lower" sum $L$ is precisely $\sum_{k=0}^{m/2-1} b_kx^{-2k}$. Replacing $k$ by $2k$ in (3.3) (as the odd terms vanish by symmetry), we have to prove

$$\sum_{j=0}^{2k} \frac{(-1)^j[(2j)!]^2[(4k-2j)!]^2}{(j!)^2[(2k-j)!]^32^{2k}} = \frac{(2k)!^3}{(k!)^48^{2k}}.$$ 

(3.5)

This can be done algorithmically using the creative telescoping approach of Wilf and Zeilberger. For example, the implementation in the Mathematica package HolonomicFunctions by Koutschan can be used. The command

```mathematica
a = ((2j)!)^2 / ((j!)^3 32^j);
CreativeTelescoping[(-1)^j a (a /. j -> 2k-j), {S[j]-1}, S[k]]
```

outputs the recurrence equation

$$(8 + 8k)b_{k+1} - (1 + 6k + 12k^2 + 8k^3) b_k = 0$$

matching the right-hand side of (3.5), together with a telescoping certificate. Since the summand in (3.5) vanishes for $j < 0$ and $j > 2k$, no boundary conditions enter into the telescoping relation, and checking the initial value ($k = 0$) suffices to prove the identity.

It remains to bound the "upper" sum $U$ given by (3.4). The coefficients of $U = \sum_{k=m}^{2m-2} c_kx^{-k}$ can be written as

$$c_k = \sum_{j=1}^{2m-k-1} (-1)^{j+k+m}a_{k-m+j}a_{m-j}.$$ 

1Curiously, the built-in Sum function in Mathematica 9.0.1 computes a closed form for the sum

(3.5), but returns an answer that is wrong by a factor 2 if the factor $[(4k-2j)!]^2$ in the summand is input as $[(2(2k-j))!]^2$. 
By symmetry, this sum is zero when $k$ is odd, so we only need to consider the case of $k$ even. We first note that, if $1 \leq i < j$, then $a_ia_j \geq a_{i+1}a_{j-1}$. This can be seen by observing that the ratio satisfies

$$\frac{a_ia_j}{a_{i+1}a_{j-1}} = \frac{(i+1)(2j-1)^2}{j(2i+1)^2} \geq 1.$$ 

Thus, after adding the duplicated terms, $c_k$ can be written as an alternating sum in which the terms decrease in magnitude, e.g. for $m = 10$ we have

$$c_{10} = -2a_1a_3 + 2a_2a_8 - 2a_3a_7 + 2a_4a_6 - a_5a_5,$$
$$c_{12} = -2a_3a_9 + 2a_4a_8 - 2a_5a_7 + a_6a_6,$$
$$c_{14} = -2a_5a_9 + 2a_6a_8 - a_7a_7,$$
$$c_{16} = -2a_7a_9 + a_8a_8,$$
$$c_{18} = -a_9a_9.$$ 

Hence $|c_k|$ is bounded by $2a_{1+k-m}a_{m-1}$, giving

$$\left| \sum_{k=m}^{2m-2} \frac{c_k}{x^k} \right| \leq \sum_{k=m}^{2m-2} t_k, \quad t_k = \frac{2a_{1+k-m}a_{m-1}}{x^k}.$$ 

Evaluating at $x = 2n, m = 4n$ as usual, the term ratio

$$\frac{t_{k+1}}{t_k} = \frac{(3 + 2k - 8n)^2}{16n(2 + k - 4n)}$$

is bounded by $1$ when $4n \leq k \leq 8n - 2$. Therefore, using (2.5),

$$\sum_{k=m}^{2m-2} t_k \leq (m-1)t_m \leq e^{-4n} \frac{(4n - 1)^{4n-1/2}}{2^{8n-1}n^{4n}} < e^{-4n}.$$ 

Adding (3.2) and (3.6), we find that $|T_{4n}(2n)| < 7e^{-4n}$. \hfill $\Box$

4. A COMPLETE ERROR BOUND

We are now equipped to justify Algorithm B3. The algorithm computes an approximation $\tilde{\gamma}$ to $\gamma$. Theorem 4.1 bounds the error $|\tilde{\gamma} - \gamma|$ in the algorithm, excluding rounding errors and any error in the evaluation of $\ln n$. The finite sums $S$ and $I$ approximate $S_0(2n)$ and $I_0(2n)$ respectively, while $T$ approximates $I_0(2n)K_0(2n)$.

**Theorem 4.1.** Given an integer $n \geq 1$, let $N \geq 4n$ be an integer such that

$$\frac{2n^{2N}H_N}{(N!)^2} < \varepsilon_0,$$

where

$$\varepsilon_0 = \frac{e^{-6n}}{(4\pi n)^{1/2}(1 + H_N)}.$$ 

Let

$$S = \sum_{k=0}^{N-1} \frac{H_{2k}}{(k!)^2}, \quad I = \sum_{k=0}^{N-1} \frac{n^{2k}}{(k!)^2}, \quad T = \frac{1}{4n} \sum_{k=0}^{2n-1} \frac{[(2k)!]^3}{(k!)^4 8^{2k} (2n)^{2k}},$$

and

$$\tilde{\gamma} = \frac{S}{I} - \frac{T}{T^2} - \ln n.$$ 

Then

$$|\tilde{\gamma} - \gamma| < 24e^{-8n}.$$
Proof. Let
\[ \varepsilon_1 = S_0(2n) - S = \sum_{k=N}^{\infty} \frac{H_k n^{2k}}{(k!)^2}, \]
\[ \varepsilon_2 = I_0(2n) - I = \sum_{k=N}^{\infty} \frac{n^{2k}}{(k!)^2}. \]
Inspection of the term ratios for \( k \geq N \) shows that \( \varepsilon_1 \) and \( \varepsilon_2 \) are bounded by the left side of (4.1). Using (2.7) to bound \( 1/I_0(2n) \), it follows that
\[
\left| \frac{S + \varepsilon_1}{I + \varepsilon_2} - \frac{S}{I} \right| = \left| \frac{\varepsilon_1 I - \varepsilon_2 S}{(I + \varepsilon_2)I} \right| \\
\leq \frac{\varepsilon_0 (I + S)}{(I + \varepsilon_2)I} \\
= \varepsilon_0 \left( \frac{1}{I_0(2n)} \right) \left( 1 + \frac{S}{I} \right) \\
< \frac{e^{-6n}}{(4\pi n)^{1/2}(1 + H_N)} \left( \frac{(4\pi n)^{1/2}}{e^{2n}} \right) (1 + H_N) \\
= e^{-8n}.
\]
We have \( T + \varepsilon_3 = I_0(2n)K_0(2n) \) where, from Lemma [3.1], \( |\varepsilon_3| < 7e^{-4n}/(4n) \). Thus, from Corollary 2.2
\[
T \leq \frac{1}{2n} + \frac{7e^{-4n}}{4n} < \frac{1}{n}.
\]
Therefore, using (2.7) again,
\[
\left| \frac{T + \varepsilon_3}{(I + \varepsilon_2)^2} - \frac{T}{I^2} \right| = \left| \frac{\varepsilon_3 I^2 - T \varepsilon_2 (2I + \varepsilon_2)}{(I + \varepsilon_2)^2I^2} \right| \\
\leq \frac{|\varepsilon_3|}{(I + \varepsilon_2)^2} + T \varepsilon_2 \frac{(2I + \varepsilon_2)}{(I + \varepsilon_2)^2I^2} \\
\leq \frac{|\varepsilon_3|}{I_0(2n)^2} + T \varepsilon_2 \frac{3}{I_0(2n)^3} \\
< 7\pi e^{-8n} + e^{-8n} \\
< 23e^{-8n}.
\]
Thus, the total error \( |\tilde{\gamma} - \gamma| \) is bounded by \( e^{-8n} + 23e^{-8n} = 24e^{-8n} \). \( \square \)

Remark 4.2. We did not try to obtain the best possible constant in (4.3). A more detailed analysis shows that we can reduce the constant 24 by a factor greater than two if \( n \) is large. See also Remark 4.5.

Since the condition on \( N \) in Theorem 4.1 is rather complicated, we give the following corollary.

Corollary 4.3. Let \( \alpha \approx 4.970625759544 \) be the unique positive real solution of \( \alpha (\ln \alpha - 1) = 3 \). If \( n \geq 138 \) and \( N \geq \alpha n \) are integers, then the conclusion of Theorem 4.1 holds.
Proof. For $138 \leq n \leq 214$ we can verify by direct computation that conditions (4.1)–(4.2) of Theorem 4.1 hold. Hence, in the following we assume that $n \geq 215$. Since $N \geq \alpha n$, this implies that $N \geq \lceil 215\alpha \rceil = 1069$.

Let $\beta = N/n$. Then $\beta \geq \alpha$, so $\beta(\ln \beta - 1) \geq 3$. Thus $2n(\beta\ln \beta - \beta - 3) \geq 0$. Taking exponentials and using $\beta = N/n$, we obtain

$$N^{2N} \geq e^{2N+6n} n^{2N}. \quad (4.4)$$

Define the real analytic function $h(x) := \ln x + \gamma + 1/(2x)$. The upper bound $H_N \leq h(N)$ follows from the Euler-Maclaurin expansion

$$H_N - \ln(N) - \gamma \sim \frac{1}{2N} - \sum_{k=1}^{\infty} \frac{B_{2k}}{2k} N^{-2k},$$

since the terms on the right-hand-side alternate in sign.

Using our assumption that $N \geq 1069$, it is easy to verify that

$$\sqrt{\pi \alpha N} \geq 2h(N)(h(N) + 1). \quad (4.5)$$

Since $\beta \geq \alpha$, it follows from (4.5) that

$$\sqrt{\pi \beta N} \geq 2h(N)(h(N) + 1). \quad (4.6)$$

Substituting $\beta = N/n$ in (4.6), it follows that

$$\pi N > 2h(N)(h(N) + 1)(\pi n)^{1/2}. \quad (4.7)$$

Using (4.4), this gives

$$\pi N^{2N+1} > 2n^{2N} h(N)(h(N) + 1)(\pi n)^{1/2} e^{2N+6n}. \quad (4.8)$$

From the first inequality of (2.4) we have $(N!)^2 \geq 2\pi N^{2N+1} e^{-2N}$. Using this and $h(N) \geq H_N$, we see that (4.7) implies

$$\quad (N!)^2 > 4n^{2N} H_N(1 + H_N)(\pi n)^{1/2} e^{6n}. \quad (4.8)$$

However, it is easy to see that (4.8) is equivalent to conditions (4.1)–(4.2) of Theorem 4.1. Hence, the conclusion of Theorem 4.1 holds. \qed

Remark 4.4. If $0 < n < 138$ then Corollary 4.3 does not apply, but a numerical computation shows that it is always sufficient to take $N \geq \alpha n + 1$.

Remark 4.5. As illustrated in Table 1 the bound in (4.3) is close to optimal for large $n$. Our bound $24e^{-8n}$ overestimates the true error, but by a factor which is inconsequential for high-precision computation of $\gamma$.

| $n$ | $N$  | $|\tilde{\gamma} - \gamma|$ | $24e^{-8n}$ |
|-----|------|-----------------------------|-------------|
| 10  | 50   | $7.68 \cdot 10^{-38}$       | $4.34 \cdot 10^{-34}$ |
| 100 | 498  | $5.32 \cdot 10^{-349}$      | $8.81 \cdot 10^{-347}$ |
| 1000| 4971 | $1.96 \cdot 10^{-3476}$     | $1.06 \cdot 10^{-3473}$ |
| 10000| 49706| $2.85 \cdot 10^{-34746}$    | $6.64 \cdot 10^{-34743}$ |

Table 1. The error $|\tilde{\gamma} - \gamma|$ compared to the bound (4.3).
REFERENCES


MATHEMATICAL SCIENCES INSTITUTE, AUSTRALIAN NATIONAL UNIVERSITY, CANBERRA, ACT 0200, AUSTRALIA

E-mail address: rpb256@rpbrent.com

RISC, JOHANNES KEPLER UNIVERSITY, 4040 LINZ, AUSTRIA

E-mail address: fredrik.johansson@risc.jku.at