A bound for the error term in the Brent-McMillan algorithm

Richard P. Brent * 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

$$\gamma = \frac{S_0(2n) - K_0(2n)}{I_0(2n)} - \ln(n), \tag{1}$$

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

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

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 γ can be computed in essentially optimal time $O(d^{1+\varepsilon})$.

*Mathematical Sciences Institute, Australian National University, Canberra, Australia [gamma@rpbrent.com]; supported by Australian Research Council grant DP140101417.

†RISC, Johannes Kepler University, 4040 Linz, Austria [fredrik.johansson@risc.jku.at]; supported by the Austrian Science Fund (FWF) grant Y464-N18.

This approach has been used for all recent record calculations of γ, 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 γ via (1). The most efficient, B3, approximates $K_0(2n)$ using the asymptotic expansion

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

where one should take $m \approx 4n$. The expansion (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 γ, 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 γ.

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) (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 $I_0(2n)K_0(2n)$ using the asymptotic expansion (2), and then the term $K_0(2n)/I_0(2n)$ in (1) is computed from $I_0(2n)K_0(2n)/I_0(2n)^2$.

Sections 2–3 contain bounds on the size of various error terms that are needed for the main result. For example, Lemma 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 (2) can be obtained formally by multiplying the asymptotic expansions (see (3)–(4) below) for K_0 and I_0. To obtain m terms in the asymptotic expansion, we multiply the polynomials $P_m(-1/z)$ and $P_m(1/z)$ occurring in (3)–(4), 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 4 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 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 2 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

$$K_0(x) = e^{-x}\left(\frac{\pi}{2x}\right)^{1/2}(P_m(-x) + R_m(x)) \tag{3}$$

and

$$I_0(x) = \frac{e^x}{(2\pi x)^{1/2}}\left(P_m(x) + Q_m(x)\right), \tag{4}$$

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

$$P_m(x) = \sum_{k=0}^{m-1} a_k x^{-k}, \text{ and } a_k = \frac{[(2k)!]^2}{(k!)^3 32^k}\,. \tag{5}$$

For $n \geq 1$,

$$\sqrt{2\pi}n^{n+1/2}e^{-n} \leq n! \leq en^{n+1/2}e^{-n}, \tag{6}$$

so the coefficients a_k in (5) satisfy

$$a_k \leq \frac{e^2}{\pi^{3/2}2^{1/2}}\frac{1}{k^{1/2}}\left(\frac{k}{2e}\right)^k < \frac{1}{k^{1/2}}\left(\frac{k}{2e}\right)^k \tag{7}$$

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

For $x > 0$, we also have the global bounds

$$0 < K_0(x) < e^{-x}\left(\frac{\pi}{2x}\right)^{1/2} \tag{8}$$

and

$$I_0(x) > \frac{e^x}{(2\pi x)^{1/2}}\,. \tag{9}$$

Observe that the bound on $K_0(x)$ and equation (3) imply that

$$|P_m(-x) + R_m(x)| < 1. \tag{10}$$

For $x > 0$, the series (3) for $K_0(x)$ is alternating, and the remainder satisfies

$$|R_m(x)| \leq \frac{a_m}{x^m} < \frac{1}{m^{1/2}}\left(\frac{m}{2e}\right)^m\frac{1}{x^m}\,. \tag{11}$$

The series (4) for $I_0(x)$ is not alternating. The following lemma bounds the error $Q_m(x)$.

Lemma 1. *Let $Q_m(x)$ be defined by (4). Then for $m \geq 1$ and real $x \geq 2$ we have*

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

Proof. The identity $I_0(x) = i(K_0(-x) - K_0(x))/\pi$ gives

$$Q_m(x) = R_m(-x) - \frac{i}{\pi}\frac{(2\pi x)^{1/2}}{e^x}K_0(x). \tag{12}$$

According to Olver [8, p. 269],

$$|R_m(-x)| \le 2\chi(m)\exp(\tfrac{1}{8}\pi x^{-1})a_m x^{-m}, \tag{13}$$

where

$$\chi(m) = \pi^{1/2}\frac{\Gamma(m/2+1)}{\Gamma(m/2+1/2)} \le \frac{\pi}{2}\,m^{1/2} \tag{14}$$

(the bound on $\chi(m)$ follows as $\chi(m)/m^{1/2}$ is monotonic decreasing for $m \ge 1$). Since $x \ge 2$, applying (7) gives

$$|R_m(-x)| \le \pi e^{\pi/16}\left(\frac{m}{2e}\right)^m \frac{1}{x^m} < 4\left(\frac{m}{2ex}\right)^m. \tag{15}$$

Combined with the global bound (8) for $K_0(x)$, we obtain

$$|Q_m(x)| \le |R_m(-x)| + \frac{1}{\pi}\frac{(2\pi x)^{1/2}}{e^x}K_0(x) \le 4\left(\frac{m}{2ex}\right)^m + e^{-2x}. \tag{16}$$

□

Corollary 1. *For $x \ge 2$, 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 (4) and (16) with $m = 1$ gives

$$I_0(x) \le \frac{e^x}{(2\pi x)^{1/2}}(1 + e^{-1} + e^{-4}),$$

so from (8) we have

$$I_0(x)K_0(x) \le \frac{1+e^{-1}+e^{-4}}{2x} < \frac{1}{x}\,.$$

□

Lemma 2. *If $R_m(x)$ and $Q_m(x)$ are defined by* (3) *and* (4) *respectively, then*

$$|R_{4n}(2n)| \le \frac{e^{-4n}}{2n^{1/2}} \;\text{ and }\; |Q_{4n}(2n)| \le 5e^{-4n}. \tag{17}$$

Proof. Taking $x = 2n$ and $m = 4n$, the inequality (11) gives the first inequality, and Lemma 1 gives the second inequality. □

We also need the following lemma.

Lemma 3. *If $P_m(x)$ is defined by* (5), *then*

$$|P_{4n}(2n)| < 2 \;\; \textit{and} \;\; |P_{4n}(-2n)| < 1. \tag{18}$$

Proof. Using (5) and (7), we have

$$\begin{aligned} P_{4n}(2n) &= 1 + \sum_{k=1}^{4n-1} \frac{a_k}{(2n)^k} \\ &\le 1 + \sum_{k=1}^{4n-1} k^{-1/2} \left(\frac{k}{4en}\right)^k \\ &\le 1 + \sum_{k=1}^{4n-1} e^{-k} < \frac{e}{e-1} < 2. \end{aligned}$$

The right inequality in (18) can be proved in a similar manner, taking the sign alternations into account. □

3 Bounds for the product

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

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

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

$$\begin{aligned} 2xI_0(x)K_0(x) &= (P_m(-x) + R_m(x))(P_m(x) + Q_m(x)) \\ &= P_m(x)P_m(-x) + \\ &\quad\; [(P_m(-x) + R_m(x))Q_m(x) + P_m(x)R_m(x)] \,. \end{aligned} \tag{19}$$

It follows from (10), (17) and (18) that the expression $[\cdots]$ in (19), evaluated at $x = 2n$, $m = 4n$, is bounded in absolute value by

$$5e^{-4n} + e^{-4n}/n^{1/2} \le 6e^{-4n}. \tag{20}$$

Next, we rewrite

$$P_m(x)P_m(-x) = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} (-1)^i a_i a_j x^{-(i+j)}$$

as $L + U$, where

$$L = \sum_{k=0}^{m-1} \left(\sum_{j=0}^{k} (-1)^j a_j a_{k-j} \right) x^{-k} \tag{21}$$

and

$$U = \sum_{k=m}^{2m-2} \left(\sum_{j=k-(m-1)}^{m-1} (-1)^j a_j a_{k-j} \right) x^{-k}. \tag{22}$$

The "lower" sum L is precisely $\sum_{k=0}^{m/2-1} b_k x^{-2k}$. Replacing k by $2k$ in (21) (as the odd terms vanish by symmetry), we have to prove

$$\sum_{j=0}^{2k} \frac{(-1)^j [(2j)!]^2 [(4k-2j)!]^2}{(j!)^3 [(2k-j)!]^3 32^{2k}} = \frac{[(2k)!]^3}{(k!)^4 8^{2k}}. \tag{23}$$

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 [6] can be used. The command

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

outputs the recurrence equation

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

matching the right-hand side of (23), together with a telescoping certificate. Since the summand in (23) 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.[1]

It remains to bound the "upper" sum U given by (22). The coefficients of $U = \sum_{k=m}^{2m-2} c_k x^{-k}$ can also be written as

$$c_k = \sum_{j=1}^{2m-k-1} (-1)^{j+k+m} a_{k-m+j} a_{m-j}. \tag{24}$$

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 $1 \le i < j$, then $a_i a_j \ge a_{i+1} a_{j-1}$. This can be seen by observing that the ratio satisfies

$$\frac{a_i a_j}{a_{i+1} a_{j-1}} = \frac{(i+1)(2j-1)^2}{j(2i+1)^2} \ge 1. \tag{25}$$

Thus, after adding the duplicated terms, c_k can be written as an alternating sum in which the terms decrease in magnitude, e.g.

$$-2a_1 a_{11} + 2a_2 a_{10} - \ldots + 2a_5 a_7 - a_6 a_6, \tag{26}$$

and its absolute value can be bounded by that of the first term, $2a_{1+k-m} a_{m-1}$, giving

$$\left| \sum_{k=m}^{2m-2} \frac{c_k}{x^k} \right| \le \sum_{k=m}^{2m-2} t_k, \quad t_k = \frac{2a_{1+k-m} a_{m-1}}{x^k}. \tag{27}$$

[1] Curiously, the built-in `Sum` function in Mathematica 9.0.1 computes a closed form for the sum (23), but returns an answer that is wrong by a factor 2 if the factor $[(4k-2j)!]^2$ in the summand is input as $[(2(2k-j))!]^2$.

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

$$\frac{t_{k+1}}{t_k} = \frac{(3 + 2k - 8n)^2}{16n(2 + k - 4n)} \tag{28}$$

is bounded by 1 when $4n \le k \le 8n - 2$. Therefore, using (7),

$$\sum_{k=m}^{2m-2} t_k \le (m-1)t_m \le e^{-4n} \frac{(4n-1)^{4n-1/2}}{2^{8n-1} n^{4n}} < e^{-4n}. \tag{29}$$

Adding (20) and (29), we find that $|T_{4n}(2n)| < 7e^{-4n}$. □

4 A complete error bound

We are now equipped to justify Algorithm B3. The algorithm computes an approximation $\widetilde{\gamma}$ to γ. Theorem 1 bounds the error $|\widetilde{\gamma} - \gamma|$ 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 1. *Given an integer $n \ge 1$, let $N \ge 4n$ be an integer such that*

$$\frac{2n^{2N} H_N}{(N!)^2} < \varepsilon_0, \tag{30}$$

where

$$\varepsilon_0 = \frac{e^{-6n}}{(4\pi n)^{1/2}(1 + H_N)} \,. \tag{31}$$

Let

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

and

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

Then

$$|\widetilde{\gamma} - \gamma| < 24e^{-8n}. \tag{32}$$

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 ε_1 and ε_2 are bounded by the left side of (30). Using (9) to bound $1/I_0(2n)$, it follows that

$$\begin{aligned}
\left|\frac{S+\varepsilon_1}{I+\varepsilon_2} - \frac{S}{I}\right| &= \left|\frac{\varepsilon_1 I - \varepsilon_2 S}{(I+\varepsilon_2)I}\right| \\
&\leq \frac{\varepsilon_0(I+S)}{(I+\varepsilon_2)I} \\
&= \varepsilon_0 \left(\frac{1}{I_0(2n)}\right)\left(1+\frac{S}{I}\right) \\
&< \frac{e^{-6n}}{(4\pi n)^{1/2}(1+H_N)} \left(\frac{(4\pi n)^{1/2}}{e^{2n}}\right)(1+H_N) \\
&= e^{-8n}.
\end{aligned}$$

We have $T + \varepsilon_3 = I_0(2n)K_0(2n)$ where, from Lemma 4, $|\varepsilon_3| < 7e^{-4n}/(4n)$. Thus, from Corollary 1,

$$T \leq \frac{1}{2n} + \frac{7e^{-4n}}{4n} < \frac{1}{n}\,\cdot$$

Therefore, using (9) again,

$$\begin{aligned}
\left|\frac{T+\varepsilon_3}{(I+\varepsilon_2)^2} - \frac{T}{I^2}\right| &= \left|\frac{\varepsilon_3 I^2 - T\varepsilon_2(2I+\varepsilon_2)}{(I+\varepsilon_2)^2 I^2}\right| \\
&\leq \frac{|\varepsilon_3|}{(I+\varepsilon_2)^2} + T\varepsilon_2 \frac{(2I+\varepsilon_2)}{(I+\varepsilon_2)^2 I^2} \\
&\leq \frac{|\varepsilon_3|}{I_0(2n)^2} + T\varepsilon_2 \frac{3}{I_0(2n)^3} \\
&< 7\pi e^{-8n} + e^{-8n} \\
&< 23e^{-8n}.
\end{aligned}$$

Thus, the total error $|\widetilde{\gamma} - \gamma|$ is bounded by $e^{-8n} + 23e^{-8n} = 24e^{-8n}$. □

Remark 1. We did not try to obtain the best possible constant in (32). 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 3.

Since the condition on N in Theorem 1 is rather complicated, we give the following corollary.

Corollary 2. *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 1 holds.*

Proof. For $138 \leq n \leq 214$ we can verify by direct computation that conditions (30)–(31) of Theorem 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

$$N^{2N} \geq e^{2N+6n} n^{2N}. \tag{33}$$

Define the real analytic function $h(x) := \ln x + \gamma + 1/(2x)$. The upper bound $H_N \le h(N)$ 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 $N \ge 1069$, it is easy to verify that

$$\sqrt{\pi \alpha N} \ge 2h(N)(h(N)+1). \tag{34}$$

Since $\beta \ge \alpha$, it follows from (34) that

$$\sqrt{\pi \beta N} \ge 2h(N)(h(N)+1). \tag{35}$$

Substituting $\beta = N/n$ in (35), it follows that

$$\pi N > 2h(N)(h(N)+1)(\pi n)^{1/2}. \tag{36}$$

Using (33), this gives

$$\pi N^{2N+1} > 2n^{2N} h(N)(h(N)+1)(\pi n)^{1/2} e^{2N+6n}. \tag{37}$$

From the first inequality of (6) we have $(N!)^2 \ge 2\pi N^{2N+1} e^{-2N}$. Using this and $h(N) \ge H_N$, we see that (37) implies

$$(N!)^2 > 4n^{2N} H_N (1 + H_N)(\pi n)^{1/2} e^{6n}. \tag{38}$$

However, it is easy to see that (38) is equivalent to conditions (30)–(31) of Theorem 1. Hence, the conclusion of Theorem 1 holds. □

Remark 2. If $0 < n < 138$ then Corollary 2 does not apply, but a numerical computation shows that it is always sufficient to take $N \ge \alpha n + 1$.

Remark 3. As indicated in Table 1, the bound in (32) is nearly optimal for large n. Our bound $24e^{-8n}$ appears to overestimate the true error by a factor that grows slightly faster than order $n^{1/2}$, which is inconsequential for high-precision computation of γ.

n	N	$\lvert\widetilde{\gamma} - \gamma\rvert$	$24e^{-8n}$
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 $\lvert\widetilde{\gamma} - \gamma\rvert$ compared to the bound (32).

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York, 1964. http://www.math.sfu.ca/~cbm/aands/.

[2] R. P. Brent. Ramanujan and Euler's Constant. Presented at the CARMA Workshop on Exploratory Experimentation and Computation in Number Theory, Newcastle, Australia, July 2010. http://maths-people.anu.edu.au/~brent/pd/Euler_CARMA_10.pdf.

[3] R. P. Brent and E. M. McMillan. Some new algorithms for high-precision computation of Euler's constant. *Mathematics of Computation*, 34(149):305–312, 1980.

[4] R. P. Brent and P. Zimmermann. *Modern Computer Arithmetic*. Cambridge University Press, Cambridge, 2010.

[5] J. D. Jackson and W. K. H. Panofsky. Edwin Mattison McMillan 1907–1991. *Biographical Memoirs Nat. Acad. Sci. (USA)*, 69:213–237, 1996.

[6] C. Koutschan. HolonomicFunctions (User's Guide). Technical Report 10-01, RISC Report Series, University of Linz, Austria, 2010.

[7] National Institute of Standards and Technology. Digital Library of Mathematical Functions. http://dlmf.nist.gov/, 2013.

[8] F. W. J. Olver. *Asymptotics and Special Functions*. A K Peters, Wellesley, MA, 1997.

[9] A. J. Yee. Euler-Mascheroni Constant. http://www.numberworld.org/digits/EulerGamma/, 2011.