A Quadratic Harmonic Approximation

April 21, 2022

1 Introduction

Some eight hundred years ago the French archbishop Nicholas Oresme developed his beautiful proof that the $n$-th harmonic number:

$$H_n := 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$$

satisfies the following growth inequality:

$$H_{2^k} > 1 + \frac{k}{2},$$

and thereby presented the first example in the history of mathematics, and the first seen by countless generations of calculus students, of an infinite series $\sum_{n=1}^{\infty} \frac{1}{n}$ which diverges although its $n$-th term decreases to zero.

Unfortunately $H_n$ has no (known) simple closed formula representation and so its further study demanded that mathematicians find suitable approximation formulas. The great Leonhard Euler applied his famous Euler–Maclaurin sum formula to obtain the following asymptotic formula:

$$H_n \sim \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} - \frac{1}{252n^6} \pm \cdots$$  \hspace{1cm} (1)

where $\gamma \approx 0.577\ldots$ is Euler’s constant. If one truncates this expansion after $n$ terms, then the error $E_n$ one commits in using the truncated series as an approximation to $H_n$ is less than the first term truncated and has the same sign.

There is considerable interest in proving simplified versions of (1) without using the heavy analytical machinery employed by Euler. For example, Robert M. Young used an elegant geometrical argument to prove the linear approximation:

$$H_n = \ln n + \gamma + \frac{1}{2(n + \theta_n)} \quad (0 < \theta_n < 1).$$

In this note we will modify his argument to prove the following quadratic approximation.
Theorem 1.

\[
H_n = \ln n + y + \frac{1}{2n} - \frac{1}{12n^2} + \varepsilon_n \quad \text{where} \quad 0 < \varepsilon_n < \frac{1}{4n^3}.
\] (2)

Admittedly, the error in Euler’s formula satisfies \(0 < E_4 < 1/120n^4\) which is much sharper; but all known proofs require much more difficult analysis than ours, while our method still gives the dominant quadratic term \(-1/12n^2\) and so is not too bad. The interest in our note is the simplicity of method to obtain a rather difficult result.

2 Geometrical proof

We let \(T_n\) be the trapezoid with base the line segment \((n, 0)\) to \((n + 1, 0)\) on the \(x\)-axis, sides the lines \(x = n\) and \(x = n + 1\) and slanted top the line segment joining the point \((n, \frac{1}{n})\) to the point \((n + 1, \frac{1}{n+1})\). We decompose \(T_n\) into three parts:

- The rectangle \(r_n\), with vertices \((n, 0), (n + 1, 0), (n + 1, \frac{1}{n+1}), (n, \frac{1}{n+1})\) and area \(\frac{1}{n+1}\).
- The curvilinear right-angled triangle with base the top of the rectangle \(r_n\) and side the segment joining \((n, \frac{1}{n+1})\) to \((n, \frac{1}{n})\) and curved “hypotenuse” the portion of the curve \(y = \frac{1}{x}\) joining the point \((n, \frac{1}{n})\) to the point \((n + 1, \frac{1}{n+1})\). We call its area \(\delta_n\).
- The “sliver” bounded below by the arc of \(y = \frac{1}{x}\) and above by the top of the trapezoid. We call its area \(\sigma_n\).

We define

\[
y_n := H_n - \ln n.
\] (3)

Then, as is well known [11] (see also [23]),

\[
\sum_{p=n}^{\infty} \delta_p = y_n - y.
\]
In the interest of completeness we reproduce Young’s nice proof:

$$\sum_{p=n}^{N} \delta_p = \left[ \int_{n}^{n+1} \frac{1}{x} \, dx - \frac{1}{n+1} \right] + \left[ \int_{n+1}^{n+2} \frac{1}{x} \, dx - \frac{1}{n+2} \right] + \cdots + \left[ \int_{N-1}^{N} \frac{1}{x} \, dx - \frac{1}{N} \right]$$

$$= \int_{n}^{N} \frac{1}{x} \, dx - \sum_{r=1}^{N-n} \frac{1}{n+r} = \int_{n}^{N} \frac{1}{x} \, dx - \left[ \sum_{r=1}^{N} \frac{1}{r} - \sum_{r=1}^{n} \frac{1}{r} \right]$$

$$= \left[ \ln N - \sum_{r=1}^{N} \frac{1}{r} \right] - \left[ \ln n - \sum_{r=1}^{n} \frac{1}{r} \right]$$

Now we let \( N \to \infty \) in the last equality and use the definitions of \( \gamma_n \) and \( \gamma \) to obtain

$$\sum_{p=n}^{\infty} \delta_p = -\gamma + \gamma_n = \gamma_n - \gamma$$

which was to be proved. But the area of the right-angled triangle at the top of the trapezoid equals

$$\frac{1}{2} \left( \frac{1}{n} - \frac{1}{n+1} \right) = \delta_n + \sigma_n,$$

and summing from \( n \) to infinity we obtain

$$\frac{1}{2n} = H_n - \ln n - \gamma + \sum_{p=n}^{\infty} \sigma_p,$$

that is,

$$H_n = \ln n + \gamma + \frac{1}{2n} - \sum_{p=n}^{\infty} \sigma_p.$$ 

Since \( \sigma_n \) is the area of the trapezoid decreased by the area under the curve \( y = \frac{1}{x} \). we obtain

$$\sigma_n = \frac{1}{2} \left( \frac{1}{n} + \frac{1}{n+1} \right) - \int_{n}^{n+1} \frac{1}{x} \, dx = \frac{1}{2n} + \frac{1}{2n(1+1/n)} - \ln \left( 1 + \frac{1}{n} \right)$$

$$= \frac{1}{2n} + \left[ \frac{1}{2n} - \frac{1}{2n^2} + \frac{1}{2n^3} - \frac{1}{2n^4} + \cdots \right] - \left[ \frac{1}{n} + \frac{1}{2n^2} - \frac{1}{3n^3} + \frac{1}{4n^4} + \cdots \right]$$

$$= \frac{1}{2} - \left( \frac{1}{3} - \frac{1}{n} - \frac{1}{4} \right) \frac{1}{n^3} + \left( \frac{1}{2} - \frac{1}{5} \right) \frac{1}{n^5} - \left( \frac{1}{2} - \frac{1}{6} \right) \frac{1}{n^6} + \cdots$$

$$= \frac{1}{6n^3} - \frac{1}{4n^4} + \frac{3}{10n^5} - \frac{1}{3n^6} + \cdots$$

which is an alternating series whose terms decrease monotonically to zero. A well-known theorem due to Leibniz states that if

$$S := a_1 - a_2 + a_3 - a_4 \pm \cdots$$
is an alternating series such that $a_n \geq 0$ and $a_n$ decreases monotonically to zero, then the series converges to a sum $S$ and if
\[ S_n := a_1 - a_2 + a_3 - a_4 \pm \cdots + (-1)^{n-1} a_n \]
is the $n$-th partial sum, then the absolute value of the remainder $R_n$ satisfies:
\[ |R_n| := |S - S_n| \leq a_{n+1} \]
and the sign of $R_n$ is $(-1)^n$. Therefore, by the Leibniz error estimate,
\[ \frac{1}{6n^3} - \frac{1}{4n^4} < \sigma_n < \frac{1}{6n^3}. \tag{4} \]
The standard estimate for the remainder from the integral test is:
\[ \int_{n+1}^{\infty} f(x) \, dx < R_n < \int_{n}^{\infty} f(x) \, dx \]
where $R_n$ is the remainder
\[ R_n := f(n+1) + f(n+2) + \cdots \]
in the series $\sum_{n=1}^{\infty} f(n)$. If we apply it to the series $\sum_{n=1}^{\infty} \frac{1}{6n^3}$ and $\sum_{n=1}^{\infty} \frac{1}{4n^4}$ we obtain
\[ \frac{1}{12(n+1)^2} - \frac{1}{12n^3} < \sum_{n=1}^{\infty} \sigma_n < \frac{1}{12n^2}. \tag{5} \]
But,
\[ \frac{1}{12(n+1)^2} - \frac{1}{12n^3} = \frac{1}{12n^2} - 2 \frac{1}{12n^3} + 3 \frac{1}{12n^4} - 4 \frac{1}{12n^5} + 5 \frac{1}{12n^6} \mp \cdots - \frac{1}{12n^3} \]
\[ = \frac{1}{12n^2} - 3 \frac{1}{12n^3} + 3 \frac{1}{12n^4} - 4 \frac{1}{12n^5} + 5 \frac{1}{12n^6} \mp \cdots \]
\[ = \frac{1}{12n^2} - \frac{1}{4n^3} + 3 \frac{1}{12n^4} - 4 \frac{1}{12n^5} + 5 \frac{1}{12n^6} \mp \cdots \]
\[ > \frac{1}{12n^2} - \frac{1}{4n^3} \]
since the series is alternating and the terms converge monotonically to zero. Therefore, if we define
\[ \epsilon_n := \frac{1}{12n^2} - \sum_{n=1}^{\infty} \sigma_n \]
we conclude that
\[ 0 < \epsilon_n < \frac{1}{4n^3} \tag{6} \]
as stated in the theorem. This completes the proof.
3 Concluding remark

Our method does not lead to an error term $O(1/n^4)$ since the terms of order $1/n^3$ for $\sigma_n$ do not cancel. It would be desirable to modify this geometric reasoning to achieve such a cancellation (perhaps using telescopic cancellation if necessary).

References

[1] R. M. Young, “Euler’s constant”, Math. Gaz. 75 (1991), 187–190.

[2] J. Havil, Gamma: Exploring Euler’s Constant, Princeton Univ. Press, Princeton, NJ, 2003: p. 74.

Mark B. Villarino
Escuela de Matemática, Universidad de Costa Rica,
10101 San José, Costa Rica
mark.villarino@ucr.ac.cr