A SHORT TELESCOPING PROOF OF HATA'S FORMULA FOR THE EULER-MASCHERONI CONSTANT

Nikita Kalinin
Mathematics and Computer Science Department, Guangdong Technion Israel Institute of Technology, Shantou, China
nikita.kalinin@gtiit.edu.cn

Received:7/17/25 , Revised: 8/21/25, Accepted: 9/17/25, Published: 11/5/25

Abstract

This note presents an elementary proof of a formula for Euler's constant γ, originally due to Masayoshi Hata. We obtain Hata's formula by a telescoping argument over Farey intervals.

1. Introduction

Recall that *Euler's constant* (also known as the Euler–Mascheroni constant) is defined as

$$\gamma = \lim_{n\to\infty} \left(1 + \tfrac{1}{2} + \cdots + \tfrac{1}{n} - \ln n\right).$$

Alternative series representations include

$$\gamma = \sum_{k=2}^{\infty} \frac{(-1)^k \zeta(k)}{k} = \log\frac{4}{\pi} + \sum_{m=2}^{\infty} \frac{(-1)^m}{m2^m} \sum_{k=1}^{m} \binom{m}{k}.$$

Euler's constant is also related to the Riemann ζ-function through the identity

$$\gamma = \log 4\pi + \sum_{\rho} \frac{1}{\rho} - 2,$$

where the sum is over the non-trivial zeros ρ of the Riemann zeta function. Further formulas and connections can be found in the survey [1] and in a more elaborate survey [4].

However, these and many other presentations of γ offer little hope for proving that γ is an irrational number (which is widely believed), so new formulas and methods are highly desirable. In this note, we investigate a lesser-known series for γ, due to Masayoshi Hata.

DOI: 10.5281/zenodo.17535156

Definition 1. An interval $\left[\frac{a}{b}, \frac{c}{d}\right]$ such that

$$a, b, c, d \in \mathbb{Z}_{\geq 0}, 0 \leq \frac{a}{b} < \frac{c}{d} \leq 1, \text{and } ad - bc = -1$$

is called a *Farey interval.* In the literature, Farey intervals are also often called *Stern-Brocot intervals.* Let $\mathcal{F}$ denote the set of Farey intervals. Let

$$\mathcal{F}^* = \left\{ \left[\frac{0}{1}, \frac{1}{n}\right] : n \in \mathbb{N} \right\} \subset \mathcal{F}.$$

Masayoshi Hata proved the following theorem.

Theorem 1 ([2])**.** *In the above notation*

$$\gamma = \frac{1}{2} + \frac{1}{2} \sum_{\left[\frac{a}{b}, \frac{c}{d}\right] \in \mathcal{F} \setminus \mathcal{F}^*} \frac{1}{abcd(a+c)(b+d)}.$$

Hata's proof is very nice. He studies presentations of functions in a certain Schauder basis associated with Farey intervals. Then, using a Parseval-type identity for the function $\psi(t) = t\{\frac{1}{t}\}\left(1 - \{\frac{1}{t}\}\right)$, he derives the above theorem. In subsequent work, Haynes–Vaaler [3] showed that the derivatives of these piecewise-linear functions from that Schauder basis form an orthonormal basis and a complete system of martingale differences in $L^2([0, 1])$, with applications to metric properties of continued fractions.

The goal of this note is to give a short and elementary proof of the above theorem using the telescoping structure over Farey intervals.

2. A Telescoping Proof of Hata's Theorem

We begin with the following lemma, which is proved by direct computation.

Lemma 1. *Let* $I = \left[\frac{a}{b}, \frac{c}{d}\right]$ *be a Farey interval. Then*

$$\frac{ad+bc}{abcd} - \frac{a(b+d)+b(a+c)}{ab(a+c)(b+d)} - \frac{(a+c)d+(b+d)c}{(a+c)(b+d)cd}$$

$$= \frac{(bc-ad)^2}{abcd(a+c)(b+d)} = \frac{1}{abcd(a+c)(b+d)}.$$

The identity in Lemma 1 suggests to associate the term $\frac{ad+bc}{abcd}$ to the interval I, so denote

$$f\left(\left[\frac{a}{b}, \frac{c}{d}\right]\right) = f(I) = \frac{ad+bc}{abcd} = \frac{1}{bc} + \frac{1}{ad}.$$

The other two terms in Lemma 1 correspond to the values of $f(I_{\text{left}})$ and $f(I_{\text{right}})$ for two Farey intervals I_{left} and I_{right} from the subdivision of I by the *mediant* $\frac{a+c}{b+d}$ of two fractions $\frac{a}{b}$ and $\frac{c}{d}$:

$$I_{\text{left}} = \left[\frac{a}{b}, \frac{a+c}{b+d}\right] \text{ and } I_{\text{right}} = \left[\frac{a+c}{b+d}, \frac{c}{d}\right].$$

Hence, Lemma 1 can be reformulated as

$$\frac{1}{abcd(a+c)(b+d)} = f(I) - f(I_{\text{left}}) - f(I_{\text{right}}). \tag{1}$$

The idea of the proof of Theorem 1 is to sum the Equation (1) over all $I \in \mathcal{F}\backslash\mathcal{F}^*$. To illustrate the telescoping property, we consider the following example, where we sum Equation (1) over $I = \left[\frac{1}{2}, \frac{1}{1}\right]$ with $I_{\text{left}} = \left[\frac{1}{2}, \frac{2}{3}\right]$ and $I_{\text{right}} = \left[\frac{2}{3}, \frac{1}{1}\right]$:

$$\begin{aligned}
f\left(\left[\frac{1}{2}, \frac{1}{1}\right]\right) &- f\left(\left[\frac{1}{2}, \frac{2}{3}\right]\right) - f\left(\left[\frac{2}{3}, \frac{1}{1}\right]\right) \\
&+ f\left(\left[\frac{1}{2}, \frac{2}{3}\right]\right) - f\left(\left[\frac{1}{2}, \frac{3}{5}\right]\right) - f\left(\left[\frac{3}{5}, \frac{2}{3}\right]\right) \\
&+ f\left(\left[\frac{2}{3}, \frac{1}{1}\right]\right) - f\left(\left[\frac{2}{3}, \frac{3}{4}\right]\right) - f\left(\left[\frac{3}{4}, \frac{1}{1}\right]\right) \\
&= f\left(\left[\frac{1}{2}, \frac{1}{1}\right]\right) - f\left(\left[\frac{1}{2}, \frac{3}{5}\right]\right) - f\left(\left[\frac{3}{5}, \frac{2}{3}\right]\right) - f\left(\left[\frac{2}{3}, \frac{3}{4}\right]\right) - f\left(\left[\frac{3}{4}, \frac{1}{1}\right]\right),
\end{aligned}$$

which is the *seed* term $f\left(\left[\frac{1}{2}, \frac{1}{1}\right]\right)$ minus the sum of *boundary* terms corresponding to the partition $\frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1}$ of $\left[\frac{1}{2}, \frac{1}{1}\right]$.

To compute the limit of finite sums as above, in the following lemma we show that the sums of boundary terms converge to the integral of $\frac{2}{t}$ as the mesh of partition tends to zero.

Lemma 2. *Let* $I = \left[\frac{a}{b}, \frac{c}{d}\right]$ *be a Farey interval with a partition*

$$\frac{a}{b} = \frac{a_1}{b_1} = x_1 < \frac{a_2}{b_2} = x_2 < \cdots < \frac{a_m}{b_m} = \frac{c}{d} = x_m,$$

such that $[\frac{a_k}{b_k}, \frac{a_{k+1}}{b_{k+1}}]$ *is a Farey interval for each* $k = 1, \ldots, m-1$ *and* $a > 0$. *Then*

$$\left| \sum_{k=1}^{m-1} f\left(\left[\frac{a_k}{b_k}, \frac{a_{k+1}}{b_{k+1}}\right]\right) - \int\limits_{a/b}^{c/d} \frac{2}{t} dt \right| \leq \left|\frac{c}{d} - \frac{a}{b}\right| \cdot \frac{b^2}{a^2} \cdot \max_{k=1,\ldots,m-1} \frac{1}{b_k b_{k+1}}. \tag{2}$$

Proof. The bound on the right of Inequality (2) will follow from the standard estimate

$$\left| (y-x)\left(\frac{g(x)+g(y)}{2}\right) - \int\limits_x^y g(t)dt \right| \leq \frac{|y-x|^2}{2} \max_{t \in [x,y]} |g'(t)|, \tag{3}$$

for a C^1-function g on an interval $[x, y]$. We consider $g(t) = 1/t$. Each Farey interval $[x_k, x_{k+1}] = \left[\frac{a_k}{b_k}, \frac{a_{k+1}}{b_{k+1}}\right]$ has length $x_{k+1} - x_k = \frac{1}{b_k b_{k+1}}$. The function f can be rewritten as

$$f\left(\left[\frac{a_k}{b_k}, \frac{a_{k+1}}{b_{k+1}}\right]\right) = \frac{1}{b_k b_{k+1}}\left(\frac{b_{k+1}}{a_{k+1}} + \frac{b_k}{a_k}\right) = (x_{k+1} - x_k) \cdot (g(x_k) + g(x_{k+1})).$$

Therefore

$$\sum_{k=1}^{m-1} f\left(\left[\frac{a_k}{b_k}, \frac{a_{k+1}}{b_{k+1}}\right]\right) = \sum_{k=1}^{m-1} (x_{k+1} - x_k) \cdot (g(x_k) + g(x_{k+1})), \tag{4}$$

which converges to $2 \int\limits_{a/b}^{c/d} g(t)dt$ as $\max_k |x_{k+1} - x_k| \to 0$, since it is a Riemann sum. Applying Inequality (3) with $g(t) = 1/t$ on each subinterval $[x_k, x_{k+1}]$ and summing over k, then using Equation (4), we obtain

$$\left|\sum_{k=1}^{m-1} f\left(\left[\frac{a_k}{b_k}, \frac{a_{k+1}}{b_{k+1}}\right]\right) - \int\limits_{a/b}^{c/d} \frac{2}{t} dt\right| \le \sum_{k=1}^{m-1} \left|\frac{a_{k+1}}{b_{k+1}} - \frac{a_k}{b_k}\right|^2 \cdot \max_{t \in [\frac{a}{b}, \frac{c}{d}]} |g'(t)|.$$

Finally, to establish Inequality (2), we compute $\max_{t \in [\frac{a}{b}, \frac{c}{d}]} |g'(t)| = \frac{b^2}{a^2}$ and

$$\sum_{k=1}^{m-1} \left|\frac{a_{k+1}}{b_{k+1}} - \frac{a_k}{b_k}\right|^2 \le \max_{k=1,\ldots,m-1} \left|\frac{a_{k+1}}{b_{k+1}} - \frac{a_k}{b_k}\right| \cdot \left|\frac{c}{d} - \frac{a}{b}\right| = \max_{k=1,\ldots,m-1} \frac{1}{b_k b_{k+1}} \cdot \left|\frac{c}{d} - \frac{a}{b}\right|.$$

□

We recall some classical properties of Farey intervals. Let

$$\mathcal{F}_0 = \left\{\left[\frac{0}{1}, \frac{1}{1}\right]\right\}, \mathcal{F}_1 = \left\{\left[\frac{0}{1}, \frac{1}{2}\right], \left[\frac{1}{2}, \frac{1}{1}\right]\right\},$$

$$\mathcal{F}_2 = \left\{\left[\frac{0}{1}, \frac{1}{3}\right], \left[\frac{1}{3}, \frac{1}{2}\right], \left[\frac{1}{2}, \frac{2}{3}\right], \left[\frac{2}{3}, \frac{1}{1}\right]\right\}, \ldots$$

where $\mathcal{F}_{n+1}$ is obtained from $\mathcal{F}_n$ by replacing each interval $I = [\frac{a}{b}, \frac{c}{d}] \in \mathcal{F}_n$ by two intervals $[\frac{a}{b}, \frac{a+c}{b+d}]$ and $[\frac{a+c}{b+d}, \frac{c}{d}]$. Using induction by n, we see that for each $[\frac{a}{b}, \frac{c}{d}] \in \mathcal{F}_n$ we have

$$\max(b, d) \ge n + 1. \tag{5}$$

Since $[\frac{0}{1}, \frac{1}{1}]$ is a Farey interval and $a(b + d) - b(a + c) = ad - bc$, it follows that $\mathcal{F}_n$ consists of Farey intervals for each $n \ge 0$. Since the length of $I = [\frac{a}{b}, \frac{c}{d}] \in \mathcal{F}_n$ is $\frac{1}{bd}$, using Inequality (5) we get the following corollary.

Corollary 1. *The length of each interval $I \in \mathcal{F}_n$ is at most $\frac{1}{n+1}$.*

We recall the proof that $\bigcup_{n=0}^{\infty} \mathcal{F}_n = \mathcal{F}$. It follows from the construction that all $\mathcal{F}_n$ are disjoint. It remains to show that each Farey interval $[\frac{a}{b}, \frac{c}{d}]$ belongs to a certain $\mathcal{F}_n$. We proceed by induction on $\max(b, d)$. If $\max(b, d) = 1$ then $b = d = 1, a = 0$, and $c = 1$, so $[\frac{a}{b}, \frac{c}{d}] = [\frac{0}{1}, \frac{1}{1}] \in \mathcal{F}_0$.

Assume that the claim holds for $\max(b, d) < k$. Consider a Farey interval $[\frac{a}{b}, \frac{c}{d}]$ with $\max(b, d) = k$. Since $ad - bc = -1$, it is not possible that $b = d = k > 1$. Without loss of generality, suppose that $b < d$. Then $a < c$. Thus $\frac{c}{d}$ is the mediant of fractions $\frac{a}{b}$ and $\frac{c-a}{b-d}$; also $a(b - d) - b(c - a) = ad - bc = -1$; so $\left[\frac{a}{b}, \frac{c-a}{b-d}\right]$ is a Farey interval. Since $\max(b, b - d) < k$, by induction hypothesis we have that $[\frac{a}{b}, \frac{c-a}{b-d}] \in \mathcal{F}_{n-1}$ for a certain $n \geq 1$. Hence $[\frac{a}{b}, \frac{c}{d}] \in \mathcal{F}_n$ by the definition.

Corollary 2. *For each interval* $I = [\frac{a}{b}, \frac{c}{d}] \in \mathcal{F} \setminus \mathcal{F}^*$ *there exists* $n \geq 1$ *such that* $I \subset [\frac{1}{n+1}, \frac{1}{n}]$.

Note that I cannot be $[\frac{0}{1}, \frac{1}{1}]$, so I must be contained in $[\frac{0}{1}, \frac{1}{2}]$ or in $[\frac{1}{2}, \frac{1}{1}] = [\frac{1}{n+1}, \frac{1}{n}]$ for $n = 1$. If I is contained in $[\frac{0}{1}, \frac{1}{2}]$, then, since I cannot be equal to $[\frac{0}{1}, \frac{1}{2}]$, I must be contained in $[\frac{0}{1}, \frac{1}{3}]$ or $[\frac{1}{3}, \frac{1}{2}] = [\frac{1}{n+1}, \frac{1}{n}]$ for $n = 2$, etc.

Lemma 3. *Fix* $n \geq 1$. *Consider the interval* $J_n = \left[\frac{1}{n+1}, \frac{1}{n}\right] \in \mathcal{F}_n$. *Then the sum of expressions in Equation* (1) *over all Farey intervals* $I \subset J_n$ *equals*

$$\frac{1}{n} + \frac{1}{n+1} - 2\ln(\frac{n+1}{n}).$$

Proof. Sum Equation (1) over the intervals $I = \left[\frac{1}{n+1}, \frac{1}{n}\right]$, $\left[\frac{1}{n+1}, \frac{2}{2n+1}\right]$, $\left[\frac{2}{2n+1}, \frac{1}{n}\right]$, etc., each time subdividing all intervals by the mediant of its endpoints. After N steps we have a partition of J_n as in Lemma 2. Therefore

$$\sum_{\substack{I \in \bigcup_{i=0}^{N} \mathcal{F}_{n+i}, \\ I \subset J_n}} \big(f(I) - f(I_{\text{left}}) - f(I_{\text{right}})\big) = f\left(\left[\frac{1}{n+1}, \frac{1}{n}\right]\right) - \sum_{k=1}^{2^N} f\left(\left[\frac{a_k}{b_k}, \frac{a_{k+1}}{b_{k+1}}\right]\right), \tag{6}$$

where $\{\left[\frac{a_k}{b_k}, \frac{a_{k+1}}{b_{k+1}}\right]\}_{k=0}^{2^N} = \{I \subset F_n, I \in \mathcal{F}_{n+k}\}$. Due to Inequality (2) and Corollary 1, as $N \to \infty$, Equation (6) converges to

$$f\left(\left[\frac{1}{n+1}, \frac{1}{n}\right]\right) - \int_{\frac{1}{n+1}}^{\frac{1}{n}} \frac{2}{t} dt = \frac{1}{n} + \frac{1}{n+1} - 2\ln(\frac{n+1}{n}).$$

□

Proof of Theorem 1. Since each interval in $\mathcal{F} \setminus \mathcal{F}^*$ is contained in one of the intervals $J_n = [\frac{1}{n+1}, \frac{1}{n}]$, using the telescoping identity from Lemma 1 and Lemma 3 for each of the intervals J_n, we can rearrange the summands so that for each k, we group all terms corresponding to the intervals contained in $[\frac{1}{k+1}, 1]$, compute their sum, and then let $k \to \infty$. Therefore, we split the sum from Theorem 1 as

$$\frac{1}{2} + \frac{1}{2} \lim_{k\to\infty} \left(\sum_{n=1}^{k} \left(\frac{1}{n} + \frac{1}{n+1} \right) - 2\ln\left(\frac{n+1}{n} \right) \right),$$

which evaluates to

$$\frac{1}{2} \cdot 2 \lim_{k\to\infty} \left(\sum_{n=1}^{k+1} \frac{1}{n} - \ln(k+1) - \frac{1}{2(k+1)} \right) = \gamma.$$

$\square$

Acknowledgement. I would like to thank the reviewer whose suggestions improved this article.

References

[1] T. P. Dence and J. B. Dence, A survey of Euler's constant, *Math. Mag.* **82** (4) (2009), 255–265.

[2] M. Hata, Farey fractions and sums over coprime pairs, *Acta Arith.* **70** (2) (1995), 149–159.

[3] A. K. Haynes and J. D. Vaaler, Martingale differences and the metric theory of continued fractions, *Illinois J. Math.* **52** (1) (2008), 213–242.

[4] J. Lagarias, Euler's constant: Euler's work and modern developments, *Bull. Amer. Math. Soc. (N.S.)* **50** (4) 2013, 527–628.