Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 20, Number 3 (2024), pp. 567-576
©Research India Publications
http://www.ripublication.com/gjpam.htm

Proving the Riemann Hypothesis with Probability of One

Yoochan Noh

Korea International School, 27, Daewangpangyo-ro, 385 Beon-gil, Bundang-gu, Seongnam-si, Gyeonggi-do-13543, South Korea.

Abstract

This paper provides an attempt to prove the Riemann hypothesis in a weakened conjecture, specifically, by proving it with the probability of 1. In Section 2 we discuss preliminaries needed to prove the Riemann Hypothesis. They include important functions such as the Mobius-μ and the Mertens function, as well as theorems such as the prime number theorem and k-almost primes. In Section 3 we simplify the Riemann Hypothesis by connecting the Mertens function with a probabilistic interpretation. The Law of Iterated logarithm is used to connect the Hypothesis with the example of a fair coin flip. In Section 4 we prove the Riemann Hypothesis with probability of 1 using the given preliminaries, theorems, and lemmas.

keywords: Riemann Hypothesis; k-almost primes; Mertens function; Mobius-μ function

1. INTRODUCTION

The Zeta function, denoted by notation ζ, was first defined in 1737 by Leonard Euler. Later, in 1859, German mathematician Bernhard Riemann proposes the Riemann Hypothesis, stating that the non trivial zeros of the zeta function have real part equal to $\frac{1}{2}$. This hypothesis has been one of the main unsolved issues for number-theorists world wide. There were many papers and approaches formulated since then in order to tackle this unsolved question, but limited research was done towards its connection to k-almost primes. Therefore, this paper will attempt to prove the hypothesis with a probability of 1 using these uninvestigated methods.

2. PRELIMINARIES

To prove the hypothesis, it is necessary to understand certain preliminaries such as the Mertens functions, prime number theorem, and almost primes.

2.1. Mertens function

Before discussing the Mertens function, it is essential to understand the Mobius-μ function, upon which it is built.

Definition 2.1. For any positive integer s, the Mobius-μ function is defined as:

$$\mu(s) = \begin{cases} 0 & \text{if } \exists p \text{ prime, } p^2 \mid s, \\ 1 & \text{if } s = p_1 p_2 \cdots p_k \text{ with } k \text{ even,} \\ -1 & \text{if } s = p_1 p_2 \cdots p_k \text{ with } k \text{ odd.} \end{cases}$$

The Mobius-μ function is a multiplicative function important in number theory, particularly in the study of arithmetic function, as its values encode the structure of the prime factors of integers.

Definition 2.2. The Mertens function is defined for all positive integers n as:

$$M(n) = \sum_{k=1}^{n} \mu(k)$$

The Mertens function provides a cumulative measure of the arithmetic properties captured by the Mobius-μ function up to n. It has been subject to deep studies in number theory, partly due to its connection with the Riemann Zeta function.

This function contributes to the main part of my paper, as it will be used to generate a statement that is equivalent to the Hypothesis, and therefore provides an approach to which can be taken in order to give a proof to it.

2.2. Almost Primes

Understanding the distribution of almost primes—numbers with only a small number of prime factors—is crucial in analytic number theory for simplifying problems about prime numbers.

Definition 2.3. The number of prime numbers less than or equal to x is defined as:

$$\pi(x) = \sum_{p \le x} 1$$

Definition 2.4. The number of k-almost prime numbers less than or equal to x is defined as

$$\pi_k(n) = \sum_{s=1}^{n} [\Omega(s) = k]$$

where if

$$n = \prod p_i^{a_i}$$

then

$$\Omega(n) := \sum a_i.$$

Almost primes are particularly interesting because they bridge prime numbers and general integers. The function $\pi_k(n)$ helps us understand how integers are constructed by the multiplication of prime numbers.

Theorem 2.5.

$$\pi(y) = \mathrm{Li}(y) + O(ye^{-c\sqrt{\log y}})$$

This theorem is known as the prime number theorem, which helps approximate the number of prime numbers less than or equal to y.

Proof. This is a well-known result, see [2]. □

Theorem 2.6.

$$\pi_k(s) \sim \left(\frac{s}{\log s}\right) \frac{(\log\log s)^{k-1}}{(k-1)!}$$

$\pi_k(s)$ helps approximate the number of k-almost primes less than or equal to s, which will be used in my paper in order to calculate the number of square-free integers with an odd or even number of prime factors.

Proof. This is also a well known result, see [1]. □

2.3. Asymptotic analysis

Asymptotic analysis is a fundamental tool in many other areas of mathematics, providing insights into the behavior of functions as arguments approach certain limits.

Definition 2.7. Functions $f(x)$ and $g(x)$ are asymptotically equivalent, meaning $f \sim g$, if and only if

$$f(x) = g(x)(1 + o(1))$$

where if a function $h(x) = o(1)$ then:

$$\lim_{x\to\infty} h(x) = 0$$

This definition is crucial for understanding how various number-theoretic functions behave as x becomes large. It allows us to simplify complex expressions and focus on dominant terms in our analyses. For example, the Prime Number Theorem itself can be seen as an asymptotic statement about the ratio of $\pi(x)$ to $\frac{x}{\log x}$.

By applying these preliminary concepts, we set the stage for a deeper exploration of the Riemann Hypothesis in subsequent sections, focusing particularly on its probabilistic aspects and numerical verifications.

3. PROBABILISTIC INTERPRETATION OF THE RIEMANN HYPOTHESIS

In this section we explore a probabilistic interpretation of this hypothesis using concepts from probability theory and statistical behavior.

Theorem 3.1. *The Riemann Hypothesis is equivalent to the statement that for every infinitesimally small positive* ϵ*,* $M(n) = O(n^{\frac{1}{2}+\epsilon})$.

Proof. For a detailed exposition, see [3]. □

3.1. Law of Iterated Logarithm

We can use the Law of Iterated Logarithm in order to show that the number of heads subtracted from the number of tails when a fair coin is flipped n number of times grows slower than a certain rate with probability of 1.

Lemma 3.2. *Suppose* $\{X_n\}$ *are independent, identically distributed random variables with a mean value of* 0 *and variance of* 1*. Let* $\{S_n\} = X_1 + \cdots + X_n$*. Then:*

$$\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\log\log n}} = 1 \quad \textit{a.s.},$$

where $a.s$ *means true with probability of* 1.

Proof. This theorem is cited from [4]. □

When considering a fair coin flip, we denote the outcome of each coin flip as Y_i where $Y_i = 1$ for heads and $Y_i = -1$ for tails. Since $E[Y_i] = 0$ and $\sigma^2_{Y_i} = 1$ are satisfied, we can directly apply this theorem. S_n, in our case, which denotes the number of heads subtracted from the number of tails, would, according to this theorem, grow at a rate proportional to $\sqrt{n\log\log n}$. Since

$$\lim_{n\to\infty} \frac{\sqrt{n\log\log n}}{n^{\frac{1}{2}+\epsilon}} = 0, \quad \forall\epsilon > 0$$

it can be concluded that S_n grows at a rate less than $n^{\frac{1}{2}+\epsilon}$ with probability of 1.

Similarly, if we show that $\mu(n) = \pm 1$ with equal probability then it is the same to showing that $M(n)$ grows at a rate less than $n^{\frac{1}{2}+\epsilon}$ and therefore proves that the Hypothesis is true with probability of 1. The case where $\mu(n) = 0$ can be ignored since adding 0 does not change the value of $M(n)$, and therefore would not affect the growth rate of $M(n)$.

Therefore, if it is shown that the proportion of the number of integers up to n which have a value of 1 when input into the Mobius-μ function to those which have a value of -1 is 1 as n approaches infinity, then it could be concluded that the Riemann Hypothesis is true with probability of 1.

4. RIEMANN HYPOTHESIS WITH PROBABILITY OF 1

Theorem 4.1. *The Riemann Hypothesis is true with Probability of* 1

Proof. Let us define the function $F_1(m)$ as:

$$F_1(m) := \sum_{n=1}^{m} [\mu(n) = -1]$$

Then, $F_1(m)$ can also be written as:

$$\begin{aligned} F_1(m) &= \pi(m) + (\pi_3(m) - S_3(m)) + (\pi_5(m) - S_5(m)) + (\pi_7(m) - S_7(m)) + \dots \\ &= (\pi(m) + \pi_3(m) + \pi_5(m) + \pi_7(m) + \dots) - (S_3(m) + S_5(m) + S_7(m) + \dots) \end{aligned} \tag{1}$$

Where $S_k(m)$ is the number of non square-free k-almost primes less than m.
Since S_k include numbers that can be divided by the square of a prime number, we can define S_k as:

$$\begin{aligned} S_k(m) := \sum_{p<\sqrt{\frac{m}{2}}} \pi_{k-2}\left(\frac{m}{p^2}\right) - \sum_{p<q<\sqrt{\frac{m}{2}}} \pi_{k-4}\left(\frac{m}{p^2q^2}\right) \\ + \sum_{p<q<r<\sqrt{\frac{m}{2}}} \pi_{k-6}\left(\frac{m}{p^2q^2r^2}\right) - \cdots \end{aligned} \tag{2}$$

Now, by using (1), (2), and 2.6 we get:

$$\begin{aligned}
F_1(m) =& \Bigg(\bigg(\frac{m}{\log m} + \left(\frac{m}{\log m}\right)\left(\frac{(\log\log m)^2}{2!}\right) + \left(\frac{m}{\log m}\right)\left(\frac{(\log\log m)^4}{4!}\right)\\
& + \left(\frac{m}{\log m}\right)\left(\frac{(\log\log m)^6}{6!}\right) + \dots\bigg) -\\
& \left(\sum_{p<\sqrt{\frac{m}{2}}} \frac{\frac{m}{p^2}}{\log(\frac{m}{p^2})} + \sum_{p<\sqrt{\frac{m}{2}}} \frac{\frac{m}{p^2}}{\log(\frac{m}{p^2})}\left(\frac{(\log\log\frac{m}{p^2})^2}{2!}\right)\right.\\
& \left. + \sum_{p<\sqrt{\frac{m}{2}}} \frac{\frac{m}{p^2}}{\log(\frac{m}{p^2})}\left(\frac{(\log\log\frac{m}{p^2})^4}{4!}\right) + \dots\right) +\\
& \left(\sum_{p<q<\sqrt{\frac{m}{2}}} \frac{\frac{m}{p^2q^2}}{\log(\frac{m}{p^2q^2})} + \sum_{p<q<\sqrt{\frac{m}{2}}} \frac{\frac{m}{p^2q^2}}{\log(\frac{m}{p^2q^2})}\left(\frac{(\log\log\frac{m}{p^2q^2})^2}{2!}\right)\right.\\
& \left. + \sum_{p<q<\sqrt{\frac{m}{2}}} \frac{\frac{m}{p^2q^2}}{\log(\frac{m}{p^2q^2})}\left(\frac{(\log\log\frac{m}{p^2q^2})^4}{4!}\right) + \dots\right) -\\
& \left(\sum_{p<q<r<\sqrt{\frac{m}{2}}} \frac{\frac{m}{p^2q^2r^2}}{\log(\frac{m}{p^2q^2r^2})} + \sum_{p<q<r<\sqrt{\frac{m}{2}}} \frac{\frac{m}{p^2q^2r^2}}{\log(\frac{m}{p^2q^2r^2})}\left(\frac{(\log\log\frac{m}{p^2q^2r^2})^2}{2!}\right)\right. +\\
& \left.\sum_{p<q<r<\sqrt{\frac{m}{2}}} \frac{\frac{m}{p^2q^2r^2}}{\log(\frac{m}{p^2q^2r^2})}\left(\frac{(\log\log\frac{m}{p^2q^2r^2})^4}{4!}\right) + \dots\right)\Bigg)(1+o(1))
\end{aligned} \tag{3}$$

Using the Taylor series, 3 simplifies to:

$$\begin{aligned}
F_1(m) = & \left(\frac{m}{2} + \frac{m}{2\log^2(m)} - \sum_{p<\sqrt{\frac{m}{2}}}\left(\frac{m}{2p^2} + \frac{m}{2p^2\log^2\left(\frac{m}{p^2}\right)}\right)\right.\\
& + \sum_{p<q<\sqrt{\frac{m}{2}}}\left(\frac{m}{2p^2q^2} + \frac{m}{2p^2q^2\log^2\left(\frac{m}{p^2q^2}\right)}\right)\\
& \left. - \sum_{p<q<r<\sqrt{\frac{m}{2}}}\left(\frac{m}{2p^2q^2r^2} + \frac{m}{2p^2q^2r^2\log^2\left(\frac{m}{p^2q^2r^2}\right)}\right) + \cdots\right)(1+o(1))
\end{aligned} \tag{4}$$

Similarly using the same method if we define F_2 as:

$$F_2(m) := \sum_{n=1}^{m} [\mu(n) = 1]$$

We get:

$$\begin{aligned} F_2(m) = \Bigg(& \frac{m}{2} - \frac{m}{2\log^2(m)} - \sum_{p<\sqrt{\frac{m}{2}}} \left(\frac{m}{2p^2} - \frac{m}{2p^2 \log^2\left(\frac{m}{p^2}\right)} \right) \\ & + \sum_{p<q<\sqrt{\frac{m}{2}}} \left(\frac{m}{2p^2q^2} - \frac{m}{2p^2q^2 \log^2\left(\frac{m}{p^2q^2}\right)} \right) \\ & - \sum_{p<q<r<\sqrt{\frac{m}{2}}} \left(\frac{m}{2p^2q^2r^2} - \frac{m}{2p^2q^2r^2 \log^2\left(\frac{m}{p^2q^2r^2}\right)} \right) + \cdots \Bigg) (1+o(1)) \end{aligned} \tag{5}$$

So, if we want to prove that $\mu(m) = \pm 1$ with equal probability then we need to prove that:

$$\lim_{m\to\infty} \frac{F_1(m)}{F_2(m)} = 1.$$

Now we can simplify this to:

$$\lim_{m\to\infty} \frac{\left(\left(1 + \frac{1}{\log^2(m)} - \sum_{p<\sqrt{\frac{m}{2}}} \frac{1}{p^2} + \sum_{p<q<\sqrt{\frac{m}{2}}} \frac{1}{p^2q^2} - \sum_{p<q<r<\sqrt{\frac{m}{2}}} \frac{1}{p^2q^2r^2} + \cdots\right) - A(m)\right)(1+o(1))}{\left(\left(1 + \frac{1}{\log^2(m)} - \sum_{p<\sqrt{\frac{m}{2}}} \frac{1}{p^2} + \sum_{p<q<\sqrt{\frac{m}{2}}} \frac{1}{p^2q^2} - \sum_{p<q<r<\sqrt{\frac{m}{2}}} \frac{1}{p^2q^2r^2} + \cdots\right) + A(m)\right)(1+o(1))}. \tag{6}$$

Where

$$A(n) := \sum_{p<\sqrt{\frac{n}{2}}} \frac{1}{p^2 \log^2\left(\frac{n}{p^2}\right)} - \sum_{p<q<\sqrt{\frac{n}{2}}} \frac{1}{p^2q^2 \log^2\left(\frac{n}{p^2q^2}\right)} + \sum_{p<q<r<\sqrt{\frac{n}{2}}} \frac{1}{p^2q^2r^2 \log^2\left(\frac{n}{p^2q^2r^2}\right)} - \cdots$$

And

$$A(n) \leq \sum_{p<\sqrt{\frac{n}{2}}} \frac{1}{p^2 \log^2\left(\frac{n}{p^2}\right)}$$

From 2.5, we can see that

$$d\pi(y) = \frac{dy}{\log y} + d(O(ye^{-c\sqrt{\log y}}))$$

Then:

$$\sum_{p\leq\sqrt{\frac{x}{2}}} \frac{1}{(p\log(\frac{x}{p^2}))^2} = \int_2^{\sqrt{\frac{x}{2}}} \frac{d\pi(y)}{(y\log(\frac{x}{y^2}))^2} = E_1(x) + E_2(x) \tag{7}$$

where

$$E_1(n) := \int_2^{\sqrt{\frac{n}{2}}} \frac{d(y)}{(y\log(\frac{n}{y^2}))^2 \log y}$$

$$E_2(n) := \int_2^{\sqrt{\frac{n}{2}}} \frac{d(O(ye^{-c\sqrt{\log y}}))}{(y\log(\frac{n}{y^2}))^2}$$

Lemma 4.2. $E_1(n) = O(\frac{1}{\log^2(n)})$

Proof. First, we split the integral into two parts:

$$\int_2^{\sqrt{\frac{n}{2}}} \frac{dy}{(y\log(\frac{n}{y^2}))^2\log y} = \int_2^{n^{1/3}} \frac{dy}{(y\log(\frac{n}{y^2}))^2\log y} + \int_{n^{1/3}}^{\sqrt{\frac{n}{2}}} \frac{dy}{(y\log(\frac{n}{y^2}))^2\log y}.$$

For the first integral, we can use the fact that $y^2 \leq \frac{n}{2}$ to show:

$$\int_2^{n^{1/3}} \frac{dy}{(y\log(\frac{n}{y^2}))^2\log y} \leq \frac{9}{(\log n)^2}\int_2^{n^{1/3}} \frac{dy}{y^2\log y} = O(\frac{1}{\log^2 n})$$

For the second integral, we can use $y^2 \geq \frac{n}{2}$ to show:

$$\int_{n^{1/3}}^{\sqrt{\frac{n}{2}}} \frac{dy}{(y\log(\frac{n}{y^2}))^2\log y} < \frac{1}{(\log 2)^2}\int_{n^{1/3}}^{\sqrt{\frac{n}{2}}} \frac{dy}{y^2} = O(n^{-\frac{1}{3}})$$

By combining the results of both integrals:

$$\int_2^{\sqrt{\frac{n}{2}}} \frac{dy}{(y\log(\frac{n}{y^2}))^2\log y} = O\left(\frac{1}{\log^2 n}\right) + O(n^{-\frac{1}{3}}) = O\left(\frac{1}{\log^2 n}\right)$$

□

Lemma 4.3. $E_2(n) = O(\frac{1}{\log^2(n)})$

Proof. Using integration by parts, we set $u = \frac{1}{(y\log(\frac{n}{y^2}))^2}$ and $dv = d(O(ye^{-c\sqrt{\log y}}))$. Through a simple computation, we can show that

$$\frac{d}{dy}\left(\frac{1}{(y\log(\frac{n}{y^2}))^2}\right) = O\left(\frac{1}{y^3\log(\frac{n}{y^2})^2}\right)$$

When $y^2 \leq \frac{n}{2}$. And so,

$$E_2(n) = \left[O(ye^{-c\sqrt{\log y}})\cdot\frac{1}{(y\log(\frac{n}{y^2}))^2}\right]_2^{\sqrt{\frac{n}{2}}} - \int_2^{\sqrt{\frac{n}{2}}} O(ye^{-c\sqrt{\log y}})\cdot O\left(\frac{1}{y^3\log(\frac{n}{y^2})^2}\right)dy$$

The boundary term

$$\left[O(ye^{-c\sqrt{\log y}}) \cdot \frac{1}{(y\log(\frac{n}{y^2}))^2}\right]_2^{\sqrt{\frac{n}{2}}}$$

requires us to consider the values at the endpoints $y = 2$ and $y = \sqrt{\frac{n}{2}}$. When $y = \sqrt{\frac{n}{2}}$,

$$O(ye^{-c\sqrt{\log y}}) \cdot \frac{1}{(y\log(\frac{n}{y^2}))^2} = O(n^{-\frac{1}{2}})$$

Similarly, at $y = 2$,

$$O(ye^{-c\sqrt{\log y}}) \cdot \frac{1}{(y\log(\frac{n}{y^2}))^2} = O\left(\frac{1}{(\log n)^2}\right)$$

Therefore,

$$\left[O(ye^{-c\sqrt{\log y}}) \cdot \frac{1}{(y\log(\frac{n}{y^2}))^2}\right]_2^{\sqrt{\frac{n}{2}}} = O(\frac{1}{\log^2 n}) + O(n^{-\frac{1}{2}})$$

For the integral:

$$\int_2^{\sqrt{\frac{n}{2}}} O(ye^{-c\sqrt{\log y}}) \cdot O\left(\frac{1}{y^3\log(\frac{n}{y^2})^2}\right) dy$$

we can split the integral into 2 parts such like 4.2 and using the majorization $e^{-c\sqrt{\log y}} << 1$ we show that :

$$\int_2^{\sqrt{\frac{n}{2}}} O(ye^{-c\sqrt{\log y}}) \cdot O\left(\frac{1}{y^3\log(\frac{n}{y^2})^2}\right) dn = O\left(\frac{1}{(\log n)^2}\right)$$

And therefore, we can conclude that:

$$E_2(n) = O(\frac{1}{\log^2 n}) + O(n^{-\frac{1}{2}}) + O\left(\frac{1}{(\log n)^2}\right) = O\left(\frac{1}{(\log n)^2}\right)$$

$\square$

Corollary 4.4. $\sum_{p\leq\sqrt{\frac{n}{2}}} \frac{1}{(p\log(\frac{n}{p^2}))^2} = O\left(\frac{1}{\log^2 n}\right)$

Proof. Using (7), and lemmas 4.2 and 4.3, we can conclude that:

$$\sum_{p\leq\sqrt{\frac{n}{2}}} \frac{1}{(p\log(\frac{n}{p^2}))^2} = O\left(\frac{1}{\log^2 n}\right)$$

$\square$

Now, we come back to (6) and 4.4 to conclude that:

$$\lim_{n\to\infty}\frac{\left(\left(1+\frac{1}{\log^2(n)}-\sum_{p<\sqrt{\frac{n}{2}}}\frac{1}{p^2}+\sum_{p<q<\sqrt{\frac{n}{2}}}\frac{1}{p^2q^2}-\sum_{p<q<r<\sqrt{\frac{n}{2}}}\frac{1}{p^2q^2r^2}+\cdots\right)-A(n)\right)(1+o(1))}{\left(\left(1+\frac{1}{\log^2(n)}-\sum_{p<\sqrt{\frac{n}{2}}}\frac{1}{p^2}+\sum_{p<q<\sqrt{\frac{n}{2}}}\frac{1}{p^2q^2}-\sum_{p<q<r<\sqrt{\frac{n}{2}}}\frac{1}{p^2q^2r^2}+\cdots\right)+A(n)\right)(1+o(1))}=1,$$

and therefore $\mu(n) = \pm 1$ with equal probability, which concludes that the Riemann hypothesis is true with probability of 1. □

REFERENCES

[1] E. Landau: *Handbuch der Lehre von der Verteilung der Primazahlen,* BG Teubner, 1909.

[2] C.J.d.L.V Poussin: *Sur la fonction $\zeta(s)$ de Riemann et le nombre des nombres premiers inferieurs à une limite donnée,* Hayez, 1899.

[3] H.M. Edwards: *Riemann's zeta function,* Courier Corporation, 2001.

[4] S.R.S Varadhan: *Stochastic processes,* American Mathematical Soc, 2007.