Article

Note for the Millennium Prize Problems

Frank Vega [1]

[1] GROUPS PLUS TOURS INC., 9611 Fontainebleau Blvd, Miami, FL, 33172, USA; vega.frank@gmail.com

Abstract: The Riemann hypothesis and the P versus NP problem are two of the most important unsolved Millennium Prize Problems. Let $\Psi(n) = n \cdot \prod_{q|n}\left(1 + \frac{1}{q}\right)$ denote the Dedekind Ψ function where $q \mid n$ means the prime q divides n. Define, for $n \geq 3$; the ratio $R(n) = \frac{\Psi(n)}{n \cdot \log\log n}$ where log is the natural logarithm. Let $N_n = 2 \cdot \ldots \cdot q_n$ be the primorial of order n. We prove if the inequality $R(N_{n+1}) < R(N_n)$ holds for all primes q_n (greater than some threshold), then the Riemann hypothesis is true. In this note, we show that the previous inequality always holds for all large enough prime numbers. A precise statement of the P versus NP problem was introduced independently by Stephen Cook and Leonid Levin. Since that date, all efforts to find a proof for this problem have failed. Another major complexity class is NP-complete. It is well-known that P is equal to NP under the assumption of the existence of a polynomial time algorithm for some NP-complete. We show that the Monotone Weighted Xor 2-satisfiability problem ($MWX2SAT$) is NP-complete and P at the same time.

Keywords: Elementary number theory; Computational complexity; Riemann hypothesis; prime numbers; complexity classes; polynomial time

1. Introduction

1.1. The Riemann hypothesis

The Riemann hypothesis is a conjecture that the Riemann zeta function has its zeros only at the negative even integers and complex numbers with real part $\frac{1}{2}$. It is considered by many to be the most important unsolved problem in pure mathematics. The hypothesis was proposed by Bernhard Riemann (1859). The Riemann hypothesis belongs to the Hilbert's eighth problem on David Hilbert's list of twenty-three unsolved problems. This is one of the Clay Mathematics Institute's Millennium Prize Problems. In recent years, there have been several developments that have brought us closer to a proof of the Riemann hypothesis. There are many approaches to the Riemann hypothesis based on analytic number theory, algebraic geometry, non-commutative geometry, etc.

The Riemann zeta function $\zeta(s)$ is a function under the domain of complex numbers. It has zeros at the negative even integers: These are called the trivial zeros. The zeta function is also zero for other values of s, which are called nontrivial zeros. The Riemann hypothesis is concerned with the locations of these nontrivial zeros. Bernhard Riemann conjectured that the real part of every nontrivial zero of the Riemann zeta function is $\frac{1}{2}$.

The Riemann hypothesis's importance remains from its deep connection to the distribution of prime numbers, which are essential in many computational and theoretical aspects of mathematics. Understanding the distribution of prime numbers is crucial for developing efficient algorithms and improving our understanding of the fundamental structure of numbers. Besides, the Riemann hypothesis stands as a testament to the power and allure of mathematical inquiry. It challenges our understanding of the fundamental structure of numbers, inspiring mathematicians to push the boundaries of their field and seek ever deeper insights into the universe of mathematics.

1.2. The P versus NP problem

P versus NP is one of the most important and challenging problems in computer science [1]. It asks whether every problem whose solution can be quickly verified can also be quickly solved. The informal term "quickly" here refers to the existence of an algorithm

that can solve the task in polynomial time [1]. The general class of problems for which such an algorithm exists is called P or "class P" [1].

Another class of problems called NP, which stands for "nondeterministic polynomial time", is defined by the property that if an input to a problem is a solution, then it can be quickly verified [1]. The P versus NP problem asks whether P equals NP. If it turns out that $P \neq NP$, which is widely believed to be the case, it would mean that there are problems in NP that are harder to compute than to verify [1]. This would have profound implications for various fields, including cryptography and artificial intelligence [2].

Solving the P versus NP problem is considered to be one of the greatest challenges in computer science [1]. A solution would have a profound impact on our understanding of computation and could lead to the development of new algorithms and techniques that could solve many of the world's most pressing problems [1]. The problem is so difficult that it is considered to be one of the seven Millennium Prize Problems, which are a set of seven unsolved problems that have been offered a 1 million prize for a correct solution [1].

2. Materials and methods

2.1. The Riemann hypothesis

In mathematics, the Chebyshev function $\theta(x)$ is given by

$$\theta(x) = \sum_{q \leq x} \log q$$

with the sum extending over all prime numbers q that are less than or equal to x, where log is the natural logarithm. We know the following inequalities:

Proposition 1. *For $r \geq 0$ and $-1 \leq x < \frac{1}{r}$ [3, pp. 1]:*

$$(1+x)^r \leq \frac{1}{1-r \cdot x}.$$

Proposition 2. *For $x > -1$ [3, pp. 1]:*

$$\frac{x}{1+x} \leq \log(1+x) \leq x.$$

Leonhard Euler studied the following value of the Riemann zeta function (1734) [4].

Proposition 3. *We define [4, (1) pp. 1070]:*

$$\zeta(2) = \prod_{k=1}^{\infty} \frac{q_k^2}{q_k^2 - 1} = \frac{\pi^2}{6},$$

where q_k is the kth prime number (We also use the notation q_n to denote the nth prime number). By definition, we have

$$\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2},$$

where n denotes a natural number. Leonhard Euler proved in his solution to the Basel problem that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \prod_{k=1}^{\infty} \frac{q_k^2}{q_k^2 - 1} = \frac{\pi^2}{6},$$

where $\pi \approx 3.14159$ is a well-known constant linked to several areas in mathematics such as number theory, geometry, etc.

https://doi.org/10.33774/coe-2024-xjsk1-v5 **ORCID:** https://orcid.org/0000-0001-8210-4126

The number $\gamma \approx 0.57721$ is the Euler-Mascheroni constant which is defined as

$$\begin{aligned}\gamma &= \lim_{n\to\infty}\left(-\log n + \sum_{k=1}^{n}\frac{1}{k}\right)\\ &= \int_1^{\infty}\left(-\frac{1}{x}+\frac{1}{\lfloor x\rfloor}\right)dx.\end{aligned}$$

Here, $\lfloor\ldots\rfloor$ represents the floor function. In number theory, $\Psi(n) = n\cdot\prod_{q|n}\left(1+\frac{1}{q}\right)$ is called the Dedekind Ψ function, where $q \mid n$ means the prime q divides n.

Definition 1. *We say that* $\mathsf{Dedekind}(q_n)$ *holds provided that*

$$\prod_{q\le q_n}\left(1+\frac{1}{q}\right) \ge \frac{e^{\gamma}}{\zeta(2)}\cdot\log\theta(q_n).$$

A natural number N_n is called a primorial number of order n precisely when,

$$N_n = \prod_{k=1}^{n} q_k.$$

We define $R(n) = \frac{\Psi(n)}{n\cdot\log\log n}$ for $n \ge 3$. $\mathsf{Dedekind}(q_n)$ holds if and only if $R(N_n) \ge \frac{e^{\gamma}}{\zeta(2)}$ is satisfied.

Proposition 4. *Unconditionally on Riemann hypothesis, we know that [5, Proposition 3 pp. 3]:*

$$\lim_{n\to\infty} R(N_n) = \frac{e^{\gamma}}{\zeta(2)}.$$

The well-known asymptotic notation Ω was introduced by Godfrey Harold Hardy and John Edensor Littlewood [6]. In 1916, they also introduced the two symbols Ω_R and Ω_L defined as [7]:

$$\begin{aligned}f(x) = \Omega_R(g(x)) \text{ as } x\to\infty \text{ if } \limsup_{x\to\infty}\frac{f(x)}{g(x)} > 0;\\ f(x) = \Omega_L(g(x)) \text{ as } x\to\infty \text{ if } \liminf_{x\to\infty}\frac{f(x)}{g(x)} < 0.\end{aligned}$$

After that, many mathematicians started using these notations in their works. From the last century, these notations Ω_R and Ω_L changed as Ω_+ and Ω_-, respectively. There is another notation: $f(x) = \Omega_{\pm}(g(x))$ (meaning that $f(x) = \Omega_+(g(x))$ and $f(x) = \Omega_-(g(x))$ are both satisfied). Nowadays, the notation $f(x) = \Omega_+(g(x))$ has survived and it is still used in analytic number theory as:

$$f(x) = \Omega_+(g(x)) \text{ if } \exists k > 0\,\forall x_0\,\exists x > x_0 \colon f(x) \ge k\cdot g(x)$$

which has the same meaning to the Hardy and Littlewood older notation. For $x \ge 2$, the function f was introduced by Nicolas in his seminal paper as [8, Theorem 3 pp. 376], [9, (5.5) pp. 111]:

$$f(x) = e^{\gamma}\cdot\log\theta(x)\cdot\prod_{q\le x}\left(1-\frac{1}{q}\right).$$

Finally, we have the Nicolas Theorem:

https://doi.org/10.33774/coe-2024-xjsk1-v5 **ORCID:** https://orcid.org/0000-0001-8210-4126

Proposition 5. *If the Riemann hypothesis is false then there exists a real b with* $0 < b < \frac{1}{2}$ *such that, as* $x \to \infty$ *[8, Theorem 3 (c) pp. 376], [9, Theorem 5.29 pp. 131]:*

$$\log f(x) = \Omega_{\pm}(x^{-b}).$$

Putting all together yields a proof for the Riemann hypothesis.

2.2. The P versus NP problem

NP-complete problems are a class of computational problems that are at the heart of many important and challenging problems in computer science. They are defined by the property that they can be quickly verified, but there is no known efficient algorithm to solve them. This means that finding a solution to an NP-complete problem can be extremely time-consuming, even for relatively small inputs. In computational complexity theory, a problem is considered NP-complete if it meets the following two criteria:

1. **Membership in NP**: A solution to an NP-complete problem can be verified in polynomial time. This means that there is an algorithm that can quickly check whether a proposed solution is correct [10].
2. **Reduction to NP-complete problems**: Any problem in NP can be reduced to an NP-complete problem in polynomial time. This means that any NP-problem can be transformed into an NP-complete problem by making a small number of changes [10].

If it were possible to find an efficient algorithm for solving any one NP-complete problem, then this algorithm could be used to solve all NP problems in polynomial time. This would have a profound impact on many fields, including cryptography, artificial intelligence, and operations research [2]. Here are some examples of NP-complete problems:

- **Boolean satisfiability problem (SAT)**: Given a Boolean formula, determine whether there is an assignment of truth values to the variables that makes the formula true [11].
- **K-CLOSURE problem**: Given a directed graph $G = (V, A)$ (V is the set of vertices and A is the set of edges) and positive integer k, determine whether there is a set V' of at most k vertices such that for all $(u, v) \in A$ either $u \in V'$ or $v \notin V'$ (see reference [Queyranne, 1976] from the Johnson and Garey book) [11]. Note that in this problem the statement "either $u \in V'$ or $v \notin V'$" does mean the same as: ($u \in V'$ or $v \in V'$) or ($u \notin V'$ or $v \notin V'$) since the logical implication of the word "**Either**" indicates that at least one of the following statements must be true, but not necessarily both.

These are just a few examples of the many NP-complete problems that have been studied and have a close relation with our current result. On the one hand, a vertex cover (sometimes called a node cover) of a graph G is a subset of its vertices, denoted by V', such that every edge in G has at least one endpoint in V'. On the other hand, an independent set V' is a subset of vertices in a graph G where no two vertices in the set are connected by an edge.

Definition 2. ***Vertex Cover and Independent Set***

INSTANCE: An undirected graph $G = (V, E)$ *and a positive integer k.*

QUESTION: Is there set V' *of at most k vertices such that* V' *is both a vertex cover and an independent set in G?*

REMARKS: This problem can be easily solved in polynomial time [11].

In this work, we show there is an NP-complete problem that can be solved in polynomial time using the previous problem. Consequently, we prove that P is equal to NP.

3. Results

3.1. The Riemann hypothesis

Several analogues of the Riemann hypothesis have already been proved. Many authors expect (or at least hope) that it is true. Nevertheless, there exist some implications in case of the Riemann hypothesis could be false. The following is a key Lemma.

Lemma 1. *If the Riemann hypothesis is false, then there exist infinitely many prime numbers q_n such that* Dedekind(q_n) *fails (i.e.* Dedekind(q_n) *does not hold).*

Proof. The function g is defined as [5, Theorem 4.2 pp. 5]:

$$g(x) = \frac{e^{\gamma}}{\zeta(2)} \cdot \log\theta(x) \cdot \prod_{q \leq x} \left(1 + \frac{1}{q}\right)^{-1}.$$

We claim that Dedekind(q_n) fails whenever there exists some real number $x_0 \geq 5$ for which $g(x_0) > 1$ or equivalent $\log g(x_0) > 0$ and q_n is the greatest prime number such that $q_n \leq x_0$. It was proven the following bound [5, Theorem 4.2 pp. 5]:

$$\log g(x) \geq \log f(x) - \frac{2}{x}.$$

By Proposition 5, if the Riemann hypothesis is false, then there is a real number $0 < b < \frac{1}{2}$ such that there exist infinitely many numbers x for which $\log f(x) = \Omega_{+}(x^{-b})$. Actually Nicolas proved that $\log f(x) = \Omega_{\pm}(x^{-b})$, but we only need to use the notation Ω_{+} under the domain of the real numbers. According to the Hardy and Littlewood definition, this would mean that

$$\exists k > 0, \forall y_0 \in \mathbb{R}, \exists y \in \mathbb{R}\ (y > y_0)\colon\ \log f(y) \geq k \cdot y^{-b}.$$

The previous inequality is also $\log f(y) \geq \left(k \cdot y^{-b} \cdot \sqrt{y}\right) \cdot \frac{1}{\sqrt{y}}$, but we notice that

$$\lim_{y \to \infty} \left(k \cdot y^{-b} \cdot \sqrt{y}\right) = \infty$$

for every possible values of $k > 0$ and $0 < b < \frac{1}{2}$. Now, this implies that

$$\forall y_0 \in \mathbb{R}, \exists y \in \mathbb{R}\ (y > y_0)\colon\ \log f(y) \geq \frac{1}{\sqrt{y}}.$$

Note that, the value of k is not necessary in the statement above. In this way, if the Riemann hypothesis is false, then there exist infinitely many wide apart numbers x such that $\log f(x) \geq \frac{1}{\sqrt{x}}$. Since $\frac{1}{\sqrt{x_0}} > \frac{2}{x_0}$ for $x_0 \geq 5$, then it would be infinitely many wide apart real numbers x_0 such that $\log g(x_0) > 0$. In addition, if $\log g(x_0) > 0$ for some real number $x_0 \geq 5$, then $\log g(x_0) = \log g(q_n)$ where q_n is the greatest prime number such that $q_n \leq x_0$. The reason is because of the equality of the following terms:

$$\prod_{q \leq x_0} \left(1 + \frac{1}{q}\right)^{-1} = \prod_{q \leq q_n} \left(1 + \frac{1}{q}\right)^{-1}$$

and

$$\theta(x_0) = \theta(q_n)$$

according to the definition of the Chebyshev function. □

This is a new Criterion for the Riemann hypothesis.

Lemma 2. *The Riemann hypothesis is true whenever for each large enough prime number q_n, there exists another prime $q_{n'} > q_n$ such that*

$$R(N_{n'}) \leq R(N_n).$$

https://doi.org/10.33774/coe-2024-xjsk1-v5 **ORCID:** https://orcid.org/0000-0001-8210-4126

Proof. By Lemma 1, if the Riemann hypothesis is false and the inequality

$$R(N_{n'}) \leq R(N_n)$$

is satisfied for each large enough prime number q_n, then there exists an infinite subsequence of natural numbers n_i such that

$$R(N_{n_{i+1}}) \leq R(N_{n_i}),$$

$q_{n_{i+1}} > q_{n_i}$ and $\mathsf{Dedekind}(q_{n_i})$ fails. By Proposition 4, this is a contradiction with the fact that

$$\liminf_{n\to\infty} R(N_n) = \lim_{n\to\infty} R(N_n) = \frac{e^\gamma}{\zeta(2)}.$$

By definition of the limit inferior for any positive real number ε, only a finite number of elements of $R(N_n)$ are less than $\frac{e^\gamma}{\zeta(2)} - \varepsilon$. This contradicts the existence of such previous infinite subsequence and thus, the Riemann hypothesis must be true. □

This is the main insight.

Theorem 1. *The inequality $R(N_n) > R(N_{n+1})$ holds for all primes q_n (greater than some threshold).*

Proof. By Lemma 2, the Riemann hypothesis is true if for all primes q_n (greater than some threshold), the inequality

$$R(N_{n'}) < R(N_n)$$

is satisfied for some prime $q_{n'} > q_n$. In particular, we will consider the case of $n' = n + 1$. That is the same as

$$\frac{\prod_{q\leq q_{n'}}\left(1+\frac{1}{q}\right)}{\log\theta(q_{n'})} < \frac{\prod_{q\leq q_n}\left(1+\frac{1}{q}\right)}{\log\theta(q_n)}$$

and

$$\log\log\theta(q_{n'}) > \log\log\theta(q_n) + \sum_{q_n<q\leq q_{n'}} \log\left(1+\frac{1}{q}\right)$$

after of applying the logarithm to the both sides and distributing the terms. That is equivalent to

$$1 > \frac{\log\log\theta(q_n)}{\log\log\theta(q_{n'})} + \frac{\sum_{q_n<q\leq q_{n'}} \log\left(1+\frac{1}{q}\right)}{\log\log\theta(q_{n'})}$$

after dividing both sides by $\log\log\theta(q_{n'})$. This is possible because of the prime number $q_{n'}$ is large enough and thus, the real number $\log\log\theta(q_{n'})$ would be greater than 0. We can apply the exponentiation to the both sides in order to obtain that

$$e > \exp\left(\frac{\log\log\theta(q_n)}{\log\log\theta(q_{n'})}\right) \cdot \left(\prod_{q_n<q\leq q_{n'}}\left(1+\frac{1}{q}\right)\right)^{\frac{1}{\log\log\theta(q_{n'})}}.$$

For large enough prime $q_{n'}$, we have

$$e = (\log\theta(q_{n'}))^{\frac{1}{\log\log\theta(q_{n'})}}$$

since $e = x^{\frac{1}{\log x}}$ for $x > 0$. Hence, it is enough to show that

$$\log\theta(q_{n'}) > \prod_{q_n<q\leq q_{n'}}\left(1+\frac{1}{q}\right).$$

https://doi.org/10.33774/coe-2024-xjsk1-v5 **ORCID:** https://orcid.org/0000-0001-8210-4126

That is equal to

$$\log\theta(q_{n+1}) > 1 + \frac{1}{q_{n+1}}$$

under the assumption that $n' = n+1$. In addition, the previous inequality is satisfied when

$$\log\theta(q_{n+1}) \geq 2.$$

We would have

$$1 + \epsilon_1 = \exp\left(\frac{\log\log\theta(q_n)}{\log\log\theta(q_{n'})}\right)$$

and

$$e \cdot (1-\epsilon_2) = \left(\prod_{q_n < q \leq q_{n'}} \left(1 + \frac{1}{q}\right)\right)^{\frac{1}{\log\log\theta(q_{n'})}}.$$

We only need to prove that

$$e > (1+\epsilon_1)\cdot e \cdot (1-\epsilon_2)$$

which is

$$\epsilon_2 > \frac{\epsilon_1}{\epsilon_1 + 1}.$$

In addition, we can see that

$$1 - e^{-1} \cdot \left(\prod_{q_n < q \leq q_{n'}} \left(1 + \frac{1}{q}\right)\right)^{\frac{1}{\log\log\theta(q_{n'})}} = \epsilon_2.$$

We have

$$\begin{aligned}
\left(\prod_{q_n < q \leq q_{n'}} \left(1 + \frac{1}{q}\right)\right)^{\frac{1}{\log\log\theta(q_{n'})}} &= \left(1 + \prod_{q_n < q \leq q_{n'}} \left(1 + \frac{1}{q}\right) - 1\right)^{\frac{1}{\log\log\theta(q_{n'})}} \\
&\leq \frac{1}{1 - \frac{\left(\prod_{q_n < q \leq q_{n'}} \left(1+\frac{1}{q}\right) - 1\right)}{\log\log\theta(q_{n'})}} \\
&= \frac{\log\log\theta(q_{n'})}{\log\log\theta(q_{n'}) + 1 - \prod_{q_n < q \leq q_{n'}} \left(1 + \frac{1}{q}\right)}
\end{aligned}$$

by Proposition 1, since

$$-1 \leq \left(\prod_{q_n < q \leq q_{n'}} \left(1 + \frac{1}{q}\right) - 1\right) < \log\log\theta(q_{n'})$$

due to q_n and $q_{n'}$ are large enough. It is a fact that if we take $n' = n+1$, then we obtain

$$\left(\prod_{q_n < q \leq q_{n+1}} \left(1 + \frac{1}{q}\right) - 1\right) = \frac{1}{q_{n+1}} < \log\log\theta(q_{n+1})$$

and thus, whenever we have

$$1 \leq \log\log\theta(q_{n+1}),$$

then that would be quite enough. As a consequence, we obtain that

$$1 - \frac{e^{-1} \cdot \log\log\theta(q_{n'})}{\log\log\theta(q_{n'}) + 1 - \prod_{q_n < q \leq q_{n'}} \left(1 + \frac{1}{q}\right)} \leq \epsilon_2.$$

https://doi.org/10.33774/coe-2024-xjsk1-v5 **ORCID:** https://orcid.org/0000-0001-8210-4126

Putting all together, we show that

$$1-\frac{e^{-1}\cdot\log\log\theta(q_{n'})}{\log\log\theta(q_{n'})+1-\prod_{q_n<q\leq q_{n'}}\left(1+\frac{1}{q}\right)}>\frac{\epsilon_1}{\epsilon_1+1}.$$

That is equal to say that

$$\frac{\epsilon_1+1}{\epsilon_1}-\frac{e^{-1}\cdot\frac{\epsilon_1+1}{\epsilon_1}\cdot\log\log\theta(q_{n'})}{\log\log\theta(q_{n'})+1-\prod_{q_n<q\leq q_{n'}}\left(1+\frac{1}{q}\right)}>1$$

and

$$1>\frac{e^{-1}\cdot(\epsilon_1+1)\cdot\log\log\theta(q_{n'})}{\log\log\theta(q_{n'})+1-\prod_{q_n<q\leq q_{n'}}\left(1+\frac{1}{q}\right)}.$$

where

$$\log\log\theta(q_{n'})+1-\prod_{q_n<q\leq q_{n'}}\left(1+\frac{1}{q}\right)>e^{-1}\cdot(\epsilon_1+1)\cdot\log\log\theta(q_{n'})$$

after making a simple distribution of the terms. If we take $n'=n+1$, then we obtain

$$-\frac{1}{q_{n+1}}>\left(e^{-1}\cdot(\epsilon_1+1)-1\right)\cdot\log\log\theta(q_{n+1}).$$

That would be

$$1<q_{n+1}\cdot\left(1-e^{-1}\cdot(\epsilon_1+1)\right)\cdot\log\log\theta(q_{n+1})$$

which is

$$0<\log q_{n+1}+\log\left(1-e^{-1}\cdot(\epsilon_1+1)\right)+\log\log\log\theta(q_{n+1}).$$

That could be rewritten as

$$0<-\frac{e^{-1}\cdot(\epsilon_1+1)}{1-e^{-1}\cdot(\epsilon_1+1)}+\log q_{n+1}+\log\log\log\theta(q_{n+1})$$

and

$$\frac{1}{e\cdot(\epsilon_1+1)^{-1}-1}<\log q_{n+1}+\log\log\log\theta(q_{n+1})$$

by Proposition 2 since $-e^{-1}\cdot(\epsilon_1+1)>-1$. The inequality

$$\frac{1}{e\cdot(\epsilon_1+1)^{-1}-1}<\log q_{n+1}+\log\log\log\theta(q_{n+1})$$

would be

$$\frac{1}{\exp\left(1-\frac{\log\log\theta(q_n)}{\log\log\theta(q_{n+1})}\right)-1}<\log q_{n+1}+\log\log\log\theta(q_{n+1})$$

because of

$$\epsilon_1=\exp\left(\frac{\log\log\theta(q_n)}{\log\log\theta(q_{n+1})}\right)-1.$$

We know that

$$\frac{1}{\exp\left(1-\frac{\log\log\theta(q_n)}{\log\log\theta(q_{n+1})}\right)-1}<\log q_{n+1}+\log\log\log\theta(q_{n+1})$$

https://doi.org/10.33774/coe-2024-xjsk1-v5 **ORCID:** https://orcid.org/0000-0001-8210-4126

holds when

$$\log q_{n+1} + \log\log\log\theta(q_{n+1}) < \exp\left(1 - \frac{\log\log\theta(q_n)}{\log\log\theta(q_{n+1})}\right) \cdot (\log q_{n+1} + \log\log\log\theta(q_{n+1}))$$

also holds. However, we deduce that

$$\log q_{n+1} + \log\log\log\theta(q_{n+1}) < \exp\left(1 - \frac{\log\log\theta(q_n)}{\log\log\theta(q_{n+1})}\right) \cdot (\log q_{n+1} + \log\log\log\theta(q_{n+1}))$$

trivially holds since

$$\exp\left(1 - \frac{\log\log\theta(q_n)}{\log\log\theta(q_{n+1})}\right) > 1$$

under the supposition that q_n and q_{n+1} are large enough. □

This is the main theorem.

Theorem 2. *The Riemann hypothesis is true.*

Proof. By Lemma 2, the Riemann hypothesis is true if for all primes q_n (greater than some threshold), the inequality

$$R(N_{n'}) \leq R(N_n)$$

is satisfied for some prime $q_{n'} > q_n$. Therefore, the Riemann hypothesis is true by Theorem 1. □

3.2. The P versus NP problem

Formally, an instance of **Boolean satisfiability problem (SAT)** is a Boolean formula ϕ which is composed of:

1. Boolean variables: $x_1, x_2, \ldots, x_n$;
2. Boolean connectives: Any Boolean function with one or two inputs and one output, such as $\wedge$(AND), $\vee$(OR), $\neg$(NOT), $\Rightarrow$(implication), $\Leftrightarrow$(if and only if);
3. and parentheses.

A truth assignment for a Boolean formula ϕ is a set of values for the variables in ϕ. A satisfying truth assignment is a truth assignment that causes ϕ to be evaluated as true. A Boolean formula with a satisfying truth assignment is satisfiable. The problem SAT asks whether a given Boolean formula is satisfiable [11].

We define a CNF Boolean formula using the following terms: A literal in a Boolean formula is an occurrence of a variable or its negation [10]. A Boolean formula is in conjunctive normal form, or CNF, if it is expressed as an AND of clauses, each of which is the OR of one or more literals [10]. A Boolean formula is in 2-conjunctive normal form or $2CNF$, if each clause has exactly two distinct literals [10].

For example, the Boolean formula:

$$(x_1 \vee \neg x_1) \wedge (x_3 \vee x_2) \wedge (\neg x_1 \vee \neg x_3)$$

is in $2CNF$. The first of its three clauses is $(x_1 \vee \neg x_1)$, which contains the two literals x_1 and $\neg x_1$.

We define the following problem:

Definition 3. ***Monotone Weighted Xor 2-satisfiability problem (MWX2SAT)***

INSTANCE: An n-variable $2CNF$ *formula with monotone clauses (meaning the variables are never negated) using logic operators* $\oplus$ *(instead of using the operator* $\vee$*) and a positive integer k.*

QUESTION: Is there exists a satisfying truth assignment in which at most k of the variables are true?

https://doi.org/10.33774/coe-2024-xjsk1-v5 **ORCID:** https://orcid.org/0000-0001-8210-4126

The following is key Lemma.

Lemma 3. *$MWX2SAT \in NP$–complete.*

Proof. For any given instance $G = (V, A)$ of the *K-CLOSURE* problem, one can construct an equivalent $MWX2SAT$ problem with a variable for each vertex of a graph and two variables for each edge of a graph. Each edge (u, v) of the graph may be represented by the $2CNF$ clauses $(u \oplus x_{uv}) \wedge (x_{uv} \oplus v) \wedge (x_{vu} \oplus x_{uv})$ where x_{vu} and x_{uv} are two new variables such that for a possible satisfying truth assignment, either both variables u and v are true and belong to a closure V' or both variables u and v are false and belong to $V - V'$. By definition, the k-vertex closure cannot have any outgoing edges pointing to vertices outside the closure. Therefore, no edge can exist where one vertex belongs to the solution and the other does not. Both endpoints of any edge must either be inside the closure or outside it. Then the satisfying instances of the resulting $2CNF$ formula using logic operators $\oplus$ encode solutions to the *K-CLOSURE* problem, and there is a satisfying truth assignment with at most $k + |A|$ true variables if and only if there is a closure with at most k vertices where $|\ldots|$ is the cardinality set function. Therefore, like *K-CLOSURE*, $MWX2SAT$ is NP-complete. □

This is the main insight.

Theorem 3. *$MWX2SAT \in P$.*

Proof. There is a connection between finding a satisfying truth assignment in $MWX2SAT$ with at most k true variables and finding a set of at most k vertices that is both a vertex cover and an independent set in a specific graph construction.

Here's a breakdown of the equivalence:

1. Graph Construction:
   - Each vertex in the original graph represents a variable in the $MWX2SAT$ formula.
   - Edges are created between variables based on the structure of the 2CNF clauses: If two variables appear in a clause (e.g., $(x \oplus y)$), then an edge is drawn between the corresponding vertices in the graph.
2. $MWX2SAT$ and the Graph:
   - A truth assignment in $MWX2SAT$ where at most k variables are true directly translates to a set of at most k vertices in the constructed graph where true variables correspond to the vertices included in the set.
   - The properties of $MWX2SAT$ clauses ensure that:
     - Vertex Cover: The chosen vertices cover all the edges (due to the structure of the clauses and the way edges are formed). This satisfies the vertex cover condition.
     - Independent Set: The chosen vertices don't have any edges connecting them (because the variables are connected in the graph, and only one variable from each clause can be true). This satisfies the independent set condition.

Therefore, finding a satisfying truth assignment with at most k true variables in $MWX2SAT$ is indeed equivalent to finding a set of at most k vertices that fulfills both vertex cover and independent set requirements in the corresponding graph. However, we know the problem of finding a set of at most k vertices that is both a vertex cover and an independent set can be easily solved in polynomial time [11]. Consequently, the instances of the problem $MWX2SAT$ can be solved in polynomial time as well. □

This is the main theorem.

Theorem 4. *There is a quadratic polynomial time algorithm for the problem $MWX2SAT$.*

https://doi.org/10.33774/coe-2024-xjsk1-v5 **ORCID:** https://orcid.org/0000-0001-8210-4126

Proof. Suppose we have the following sequence of variables in a given instance of $MWX2SAT$:

$$x_1, \ldots, x_n.$$

For each variable x_i in the $2CNF$ formula, we define the functions f and g as,

- $f(x_i)$ is the number of variables x_j such that either $(x_i \vee x_j)$ or $(x_j \vee x_i)$ belongs to the $2CNF$ formula whenever $j > i$;
- $g(x_i)$ is the number of variables x_j such that either $(x_i \vee x_j)$ or $(x_j \vee x_i)$ belongs to the $2CNF$ formula whenever $j < i$.

We define a state as a quadruple (i, s, r, t) of integers. This state represents the fact that,

$$\text{“}\textit{the subset of variables } x_1, \ldots, x_i$$

$$\textit{with } s \textit{ true variables}$$

$$\textit{where } -m \leq r \leq m \textit{ and } -m \leq t \leq m''\text{,}$$

where m is the amount of clauses into the $2CNF$ formula. Each state (i, s, r, t) has two next states:

1. $(i+1, s+1, r+f(x_{i+1}), t-g(x_{i+1}))$, implying that x_{i+1} is included in the subset and it is evaluated as true;
2. $(i+1, s, r-g(x_{i+1}), t+f(x_{i+1}))$, implying that x_{i+1} is included in the subset and it is evaluated as false.

Starting from the initial state $(0, 0, 0, 0)$, it is possible to use any graph search algorithm (e.g. **Breadth-first search (BFS)** [10]) to search any state $(n, \varpi, 0, 0)$ such that $0 < \varpi \leq k$. Certainly, we satisfy all the clauses if they contain exactly one true literal just adding 1 by the true literal from the left most position (otherwise adding 1 by the false literal from the left most position) and subtracting 1 by the false literal from the right most position (otherwise subtracting 1 by the true literal from the right most position). The run-time of this algorithm is at most linear in the number of states. The number of states is at bounded by $n^2 \cdot 4 \cdot m^2$ times and therefore, the whole time required is $O((n \cdot m)^2)$. We create our software programming implementation in Python for this algorithm [12]. This is placed into a GitHub repository under my GitHub username “frankvegadelgado” [12]. The last commit was on February 27th of 2024 with a SHA commit **eec2ccb2bba51e3efaf3ad7d722b948b88861ea9** [12]. □

4. Conclusion

On the one hand, the Riemann hypothesis has far-reaching implications for mathematics, with potential applications in cryptography, number theory, and even particle physics. Certainly, a proof of the hypothesis would not only provide a profound insight into the nature of prime numbers but also open up new avenues of research in various mathematical fields. On the other hand, a proof of $P = NP$ will have stunning practical consequences, because it possibly leads to efficient methods for solving some of the important problems in computer science [1]. The consequences, both positive and negative, arise since various NP-complete problems are fundamental in many fields [2]. But such changes may pale in significance compared to the revolution an efficient method for solving NP-complete problems will cause in mathematics itself [1]. Research mathematicians spend their careers trying to prove theorems, and some proofs have taken decades or even centuries to be discovered after problems have been stated [1]. A method that guarantees to find proofs for theorems, should one exist of a “reasonable” size, would essentially end this struggle [1].

References

1. Cook, S.A. The P versus NP Problem, Clay Mathematics Institute. https://www.claymath.org/wp-content/uploads/2022/06/pvsnp.pdf, 2022. Accessed 1 February 2024.
2. Fortnow, L. The status of the P versus NP problem. *Communications of the ACM* **2009**, *52*, 78–86. https://doi.org/10.1145/1562164.1562186.

3. Kozma, L. Useful Inequalities. Kozma's Homepage, Useful inequalities cheat sheet. http://www.lkozma.net/inequalities_cheat_sheet/ineq.pdf, 2011–2024. Accessed 2 February 2024.
4. Ayoub, R. Euler and the Zeta Function. *The American Mathematical Monthly* **1974**, *81*, 1067–1086. https://doi.org/10.2307/2319041.
5. Solé, P.; Planat, M. Extreme values of the Dedekind Ψ function. *Journal of Combinatorics and Number Theory* **2011**, *3*, 33–38.
6. Hardy, G.H.; Littlewood, J.E. Some problems of diophantine approximation: Part II. The trigonometrical series associated with the elliptic ϑ-functions. *Acta mathematica* **1914**, *37*, 193–239.
7. Hardy, G.H.; Littlewood, J.E. Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes. *Acta Mathematica* **1916**, *41*, 119–196.
8. Nicolas, J.L. Petites valeurs de la fonction d'Euler. *Journal of Number Theory* **1983**, *17*, 375–388. https://doi.org/10.1016/0022-314X(83)90055-0.
9. Broughan, K., Euler's Totient Function. In *Equivalents of the Riemann Hypothesis*; Cambridge University Press, 2017; Vol. 1, *Encyclopedia of Mathematics and its Applications*, pp. 94–143. https://doi.org/10.1017/9781108178228.007.
10. Cormen, T.H.; Leiserson, C.E.; Rivest, R.L.; Stein, C. *Introduction to Algorithms*, 3rd ed.; The MIT Press, 2009.
11. Garey, M.R.; Johnson, D.S. *Computers and Intractability: A Guide to the Theory of NP-Completeness*, 1 ed.; San Francisco: W. H. Freeman and Company, 1979.
12. Vega, F. ALMA | MWX2SAT Solver. https://github.com/frankvegadelgado/alma, 2024. Accessed 2 February 2024.

Short Biography of Authors

Frank Vega is essentially a Back-End Programmer and Mathematical Hobbyist who graduated in Computer Science in 2007. In May 2022, The Ramanujan Journal accepted his mathematical article about the Riemann hypothesis. The article "Robin's criterion on divisibility" makes several significant contributions to the field of number theory. It provides a proof of the Robin inequality for a large class of integers, and it suggests new directions for research in the area of analytic number theory.