GENERALIZATIONS ON BERTRAND'S POSTULATE

KEVIN XU

1. Introduction

In this paper, we will discuss and delve into the intricacies of Bertrand's postulate, a well-known theorem that bounds the gaps between prime numbers. First, we will use Ramanujan's method to prove Bertrand's postulate, which involves basic analytic number theory and Stirling approximations. Then, we will go into some generalizations of Bertrand's postulate, including when there are at least two, three, and some arbitrary number of primes between n and $2n$. This paper assumes one has knowledge of basic analytic number theory.

I give special thanks to Simon Rubinstein-Salzedo and the Euler Circle for contributing to the creation of this paper.

Date: October 27, 2022.

2. Definitions

Before we jump into Ramanujan's proof, we need to establish some fundamental functions that are widely used in analytic number theory. For this paper, we will make the convention that $\log x$ denotes the natural log of x and that $x!$ is the factorial of x, or $\prod_{k\leq x} k$.

Definition 2.1. The *gamma function*, denoted as $\Gamma(n)$, is defined as

$$\Gamma(n) = \int_0^\infty x^{n-1}e^{-x}dx.$$

Corollary 2.2. *The gamma function satisfies the following properties:*

(1) $\Gamma(1) = 1$
(2) $\Gamma(n+1) = n\Gamma(n)$
(3) *if n is a positive integer, then* $\Gamma(n) = (n-1)!$.

All of these properties can be proven by just substituting values in the integral. A harder result to prove is that $\Gamma(1/2) = \sqrt{\pi}$, but we will not need it in this paper.

Definition 2.3. The *first Chebyshev function*, denoted as $\vartheta(x)$, is defined to be

$$\vartheta(x) = \sum_{\substack{p \text{ prime} \\ p\leq x}} \log p.$$

Definition 2.4. The *von Mangoldt function* $\Lambda(x)$ satisfies

$$\Lambda(x) = \begin{cases} 1 & x = 1 \\ \log x & \text{if } x \text{ is a power of a prime greater than 1} \\ 0 & \text{otherwise.} \end{cases}$$

Definition 2.5. The *second Chebyshev function*, denoted as $\psi(x)$, is defined to be

$$\psi(x) = \sum_{n\leq x} \Lambda(n).$$

However, the summand is only nonzero when n is a power of a prime, so we also have

$$\psi(x) = \sum_{\substack{p \text{ prime} \\ p\leq x}} k \log p$$

where k is the greatest power of p such that $p^k \leq x$.

Definition 2.6. The *Bernoulli numbers*, where the nth Bernoulli number is denoted as B_n, are a sequence of rational numbers satisfying

$$\frac{x}{e^x - 1} = \sum_{n=0}^\infty \frac{B_n x^n}{n!}$$

In addition to these, we will also introduce the concept of *little o notation*:

Definition 2.7. If $f(x), g(x)$ are two real-valued functions, then $f(x) = o(g(x))$ if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0.$$

In other words, $g(x)$ grows much faster than $f(x)$.

3. Bertrand's Postulate

To start with the proof, we find a relationship between the Chebyshev functions.

Lemma 3.1. *We have*

$$\psi(x) = \sum_{n=1}^{\infty} \vartheta(x^{\frac{1}{n}}).$$

Proof. This is rather simple. Let k be defined as the greatest power of p with $p^k \le x$, and note

$$\psi(x) = \sum_{n=1}^{\infty} \vartheta(x^{\frac{1}{n}}) = \sum_{n=1}^{\infty} \sum_{\substack{p \text{ prime} \\ p \le x^{1/n}}} \log p = \sum_{\substack{p \text{ prime} \\ p \le x}} \sum_{n=1}^{x^{1/n}} \log p = \sum_{\substack{p \text{ prime} \\ p \le x}} k \log p$$

as desired. ■

Lemma 3.2. *Another property of the second Chebyshev function is the following:*

$$\sum_{n=1}^{\infty} \psi\left(\frac{x}{n}\right) = \log \lfloor x \rfloor !$$

Proof. Similar to the last proof, we have

$$\sum_{n=1}^{\infty} \psi\left(\frac{x}{n}\right) = \sum_{n=1}^{\infty} \sum_{mn \le x} \Lambda(m) = \sum_{n \le x} \sum_{mn \le x} \Lambda(m).$$

Each $\Lambda(m)$ occurs once each time there exists an n with $mn \le x$, so we can also express this sum as

$$\sum_{m \le x} \Lambda(m) \sum_{mn \le x} 1 = \sum_{m \le x} \Lambda(m) \left\lfloor \frac{x}{m} \right\rfloor .$$

Because $\Lambda(m) = 0$ if m is not a power of a prime, let $p_1^{e_1} p_2^{e_2} \dots p_r^{e_r}$ be the prime factorization of $\lfloor x \rfloor !$, so our sum is

$$\sum_{i=1}^{r} \left(e_i \log(p_i) \sum_{k=1}^{e_i} \left\lfloor \frac{x}{p_i^k} \right\rfloor \right) = \sum_{i=1}^{r} \left(\log(p_i^{e_i}) \sum_{k=1}^{e_i} \left\lfloor \frac{x}{p_i^k} \right\rfloor \right)$$

which is just $\log(\lfloor x \rfloor !)$ because of Legendre's formula. ■

With these, we begin our proof of Bertrand's postulate. Define a new function $P(x)$ satisfying $P(x) = \psi(x) - 2\psi(\sqrt{x})$. Because of Lemma 3.1, we find that

$$P(x) = \sum_{n=1}^{\infty} \vartheta(x^{\frac{1}{n}}) - 2 \sum_{n=1}^{\infty} \vartheta(x^{\frac{1}{2n}}) = \sum_{n=1}^{\infty} (-1)^{n+1} \vartheta(x^{\frac{1}{n}})$$

so $P(x)$ is the alternating sum of $\vartheta(x)$. Now because $\vartheta(x)$ is a steadily increasing function, we have $\vartheta(x^{\frac{1}{k}}) \ge \vartheta(x^{\frac{1}{k+1}})$ for positive integral k, so in fact we have $P(x) \le \vartheta(x)$. Next, we define another function $Q(x)$ satisfying $Q(x) = \log \lfloor x \rfloor ! - 2 \log \lfloor \frac{x}{2} \rfloor !$, which satisfies (by Lemma 3.2)

$$Q(x) = \sum_{n=1}^{\infty} \psi\left(\frac{x}{n}\right) - 2\sum_{n=1}^{\infty} \psi\left(\frac{x}{2n}\right) = \sum_{n=1}^{\infty} (-1)^{n+1} \psi\left(\frac{x}{n}\right)$$

so $Q(x)$ is the alternating sum of $\psi(x)$. Again, as $\psi(x)$ is a steadily increasing function, we actually have $\psi(x) - \psi\left(\frac{x}{2}\right) \leq Q(x) \leq \psi(x) - \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right)$.

Remembering that $\Gamma(x) \geq \Gamma(\lfloor x \rfloor) = \lfloor x - 1 \rfloor !$, we can also express this bound as

$$\log[\Gamma(x)] - 2\log\left[\Gamma\left(\frac{x+1}{2}\right)\right] \leq Q(x) \leq \log[\Gamma(x+1)] - 2\log\left[\Gamma\left(\frac{x+1}{2}\right)\right].$$

There are more malleable bounds of $Q(x)$, as we will prove in the following theorem:

Theorem 3.3. *$Q(x) < \frac{3}{4}x$ for $x \geq 8$ and $Q(x) > \frac{2}{3}x$ for $x \geq 570$.*

This is quite hard to prove, so we will only do so after we have established Stirling approximations.

Theorem 3.4 (Euler-Maclaurin)**.** *Let $f(x)$ be a continuous function in the closed interval $[a, b]$. Then*

$$\sum_{n=a}^{b} f(n) = \int_a^b f(x)dx + \frac{f(a) + f(b)}{2} + \sum_{n=2}^{\infty} \left[\frac{B_n}{n!}\left(f^{(n-1)}(b) - f^{(n-1)}(a)\right)\right] + R_{a,b}$$

where $f^{(n)}(x)$ denotes the nth derivative of $f(x)$ and $R_{a,b}$ is the small error term defined as the difference of the summation and integral of $f(x)$ over $[a, b]$.

Proof. To make things clearer, we let the operator of differentiation, $\frac{d}{dx}$, be an operator called D, and also let $D^n = \frac{d^n}{dx^n}$. Next, define T to be the exponential function of the differential, or $T = e^D$. Note that

$$T = e^D = \sum_{n=0}^{\infty} \frac{D^n}{n!}$$

and when T is applied onto a function $f(x)$,

$$Tf(x) = \sum_{n=0}^{\infty} \frac{D^n f(x)}{n!} = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x) = f(x+1)$$

where the last step follows because of the Taylor series for $f(x)$. More generally, $T^k f(x) = f(x+k)$ for positive integral k. Now, notice that

$$\begin{aligned}\sum_{n=0}^{\infty} f(x+n) &= f(x) + Tf(x) + T^2 f(x) + \dots \\ &= \frac{1}{1-T} f(x) \\ &= \frac{1}{1-e^D} f(x) \\ &= -\frac{1}{D}\frac{D}{e^D-1} f(x) \\ &= -D^{-1} f(x)\left(1 - \frac{1}{2}D + \sum_{n=2}^{\infty}\frac{B_n}{n!}D^n\right).\end{aligned}$$

But because $D = \dfrac{d}{dx}$, then D^{-1} is the inverse of $\dfrac{d}{dx}$, or the integral. Thus

$$-D^{-1} f(x) = \int_x^{\infty} f(u)du$$

where the limits of integration are found by noting when the integral becomes close to the summation on the LHS. Thus, we can see that

$$\sum_{n=0}^{\infty} f(x+n) = \int_x^{\infty} f(u)du + \frac{1}{2}f(x) - \sum_{n=2}^{\infty}\frac{B_n}{n!}D^{n-1}f(x).$$

Using this identity,

$$\begin{aligned}\sum_{n=a}^{b} f(n) &= \sum_{n=a}^{\infty} f(n) - \sum_{n=b+1}^{\infty} f(n) \\ &= \sum_{n=0}^{\infty} f(a+n) - \sum_{n=0}^{\infty} f(b+n) + f(b) \\ &= \int_a^{\infty} f(x)dx - \int_b^{\infty} f(x)dx + R_{a,b} + \frac{f(a)-f(b)}{2} + f(b) - \sum_{n=2}^{\infty}\left[\frac{B_n}{n!}D^{n-1}(f(a)-f(b))\right] \\ &= \int_a^b f(x)dx + \frac{f(a)+f(b)}{2} - \sum_{n=2}^{\infty}\left[\frac{B_n}{n!}(f^{(n-1)}(a) - f^{(n-1)}(b))\right] + R_{a,b}\end{aligned}$$

as desired. ■

Proof of Theorem 3.3. Setting $f(x) = \log\lfloor x\rfloor!$, we have $f^{(n)}(x) = (-1)^{n+1}(n-1)!x^{-n}$. Now using the Euler-Maclaurin formula,

$$\begin{aligned}\log\lfloor x\rfloor! = \sum_{n\le x}\log n &= \int_1^x \log(u)du + \frac{1}{2}\log x - \sum_{n=2}^{\infty}\left[\frac{(-1)^n B_n}{n(n-1)}(1-x^{1-n})\right] + R_{1,x} \\ &= x\log x - x + 1 + \frac{1}{2}\log x + R_{1,x} - \frac{1}{12}(1-x^{-1}) + \frac{1}{360}(1-x^{-3}) - + \dots\end{aligned}$$

which is referred to as *Stirling's approximation*. Then,

$$Q(x) = \log\lfloor x \rfloor! - 2\log\left\lfloor \frac{x+1}{2} \right\rfloor! = x\log x - (x+2)\log\left(\frac{x+1}{2}\right) + \frac{1}{2}\log x + E(x)$$

where $E(x)$ is some error function. We can express this as

$$\begin{aligned} Q(x) &= \frac{2}{3}x + (\log 2 - \frac{2}{3})x + 2\log 2 + x\log x - (x+2)\log(x+1) + \frac{1}{2}\log x + E(x) \\ &= \frac{2}{3}x + A(x) + E(x) \end{aligned}$$

for some function $A(x)$, so it suffices to prove that $A(x) + E(x) > 0$ for $x \geq 570$.

First, note that

$$x\log x - (x+2)\log(x+1) = -x\log\left(1 + \frac{1}{x}\right) - 2\log(x+1) = \sum_{n=1}^{\infty} \frac{(-1)^n}{nx^{n-1}} - 2\log(x+1)$$

from the Taylor expansion for log. But then

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{nx^{n-1}} - 2\log(x+1) > -\sum_{n=1}^{\infty} \frac{1}{x^{n-1}} - 2\log(x+1) = -\frac{x}{x-1} - 2\log(x+1)$$

for $x > 1$. Thus we have

$$\begin{aligned} A(x) &> (\log 2 - \frac{2}{3})x + 2\log 2 - \frac{x}{x-1} - 2\log(x+1) + \frac{1}{2}\log x \\ &> (\log 2 - \frac{2}{3})x - \frac{1}{x} - \frac{3}{2}\log x \\ &\approx (\log 2 - \frac{2}{3})x - \frac{3}{2}\log x \end{aligned}$$

for reasonably large x. Now for $E(x)$, note

$$\begin{aligned} E(x) &= \sum_{n=1}^{\infty} \left[\frac{B_{2n}}{2n(2n-1)} \left(1 - 2\left(\frac{x+1}{2}\right)^{1-2n} + x^{1-2n} \right) \right] + R_{1,x} - 2R_{1,(x+1)/2} \\ &\approx \sum_{n=1}^{\infty} \frac{B_{2n}}{2n(2n-1)} + R_{1,x} - 2R_{1,(x+1)/2} \\ &> \sum_{n=1}^{\infty} \frac{B_{2n}}{2n(2n-1)} - R_{1,(x+1)/2} \end{aligned}$$

because $R_{1,x}$ is an increasing function. But we also have

$$R_{1,(x+1)/2} = \sum_{n=1}^{(x+1)/2} \log n - \int_1^{(x+1)/2} \log u \, du$$

and upon noticing that the summation is a right Riemann sum of the integral, we see that the error term cannot be bigger than the rightmost rectangle in the Riemann sum:

$$R_{1,(x+1)/2} < \log(x+1) - \log 2.$$

Putting it all together,

$$A(x)+E(x)>(\log 2-\frac{2}{3})x-\frac{3}{2}\log x+\sum_{n=1}^{\infty}\frac{B_{2n}}{2n(2n-1)}-\log(x+1)+\log 2$$

but clearly $(\log 2-\frac{2}{3})x$ grows faster than $\frac{3}{2}\log x$ and $\log(x+1)$, so $A(x)+E(x)>0 \Rightarrow Q(x)>\frac{2}{3}x$ for $x\geq 570$.

Now for the upper bound,

$$\begin{aligned}Q(x)&=\frac{3}{4}x-\left(\frac{3}{4}-\log 2\right)x+2\log 2+x\log x-(x+2)\log(x+1)+\frac{1}{2}\log x+1+E(x)\\&=\frac{3}{4}x-B(x)+E(x)\end{aligned}$$

for some function $B(x)$, so it suffices to prove that $B(x)-E(x)>0$. First, note that

$$\begin{aligned}B(x)&>\left(\frac{3}{4}-\log 2\right)x-2\log 2+x\log\left(1+\frac{1}{x}\right)+\frac{3}{2}\log(x+1)-1\\&>\left(\frac{3}{4}-\log 2\right)x-2\log 2+\frac{3}{2}\log(x+1)-1.\end{aligned}$$

For $E(x)$, we just evaluate to get

$$\begin{aligned}E(x)&=\sum_{n=1}^{\infty}\left[\frac{B_{2n}}{2n(2n-1)}\left(1-2\left(\frac{x+1}{2}\right)^{1-2n}+x^{1-2n}\right)\right]+R_{1,x}-2R_{1,(x+1)/2}\\&\approx\sum_{n=1}^{\infty}\frac{B_{2n}}{2n(2n-1)}+R_{1,x}-2R_{1,(x+1)/2}\\&<\sum_{n=1}^{\infty}\frac{1}{2n(2n-1)}+R_{(x+1)/2,x}\\&=R_{(x+1)/2,x}+1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\dots\\&<\log x-\log(x+1)+2\log 2\end{aligned}$$

where the last step follows like before. As $(\frac{3}{4}-\log 2)x$ grows faster than $\log x$, we must have $B(x)-E(x)>0\Rightarrow Q(x)<\frac{3}{4}x$ for $x\geq 8$. ■

Finally, we can prove Bertrand's postulate:

Theorem 3.5 (Bertrand's Postulate)**.** *There exists a prime between n and $2n$ for all $n>1$.*

Proof. From earlier we have $\psi(x)-\psi\left(\frac{x}{2}\right)\leq Q(x)\leq\psi(x)-\psi\left(\frac{x}{2}\right)+\psi\left(\frac{x}{3}\right)$, so due to Theorem 3.3 $\psi(x)-\psi\left(\frac{x}{2}\right)<\frac{3}{4}x$ for $x\geq 8$. Notice

$$\psi(x)-\psi\left(\frac{x}{2}\right)+\psi\left(\frac{x}{2}\right)-\psi\left(\frac{x}{4}\right)+\dots=\psi(x)<\frac{3}{4}(x+\frac{x}{2}+\frac{x}{4}+\dots)=\frac{3}{2}x$$

for $x \geq 8$ by substituting $x, \frac{x}{2}, \frac{x}{4}, \ldots$ into the above inequality. Also from $P(x) \leq \vartheta(x)$ and $\vartheta(x) \leq \psi(x)$,

$$\begin{aligned}\psi(x) - \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) &\leq \vartheta(x) + 2\psi(\sqrt{x}) - \vartheta\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) \\ &< \vartheta(x) + 3\sqrt{x} - \vartheta\left(\frac{x}{2}\right) + \frac{1}{2}x\end{aligned}$$

But we also have $\frac{2}{3}x < Q(x) \leq \psi(x) - \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right)$ so

$$\vartheta(x) - \vartheta\left(\frac{x}{2}\right) > \frac{1}{6}x - 3\sqrt{x}$$

for $x \geq 570$. But $\frac{1}{6}x - 3\sqrt{x} \geq 0$ for $x \geq 324$, so

$$\vartheta(2x) - \vartheta(x) > 0$$

for $x \geq 285$. But this immediately implies that there exists at least one prime between x and $2x$ for $x \geq 285$. Now we just have to check for $x < 285$, but the primes $3, 5, 7, 13, 23, 41, 79, 157, 313$ work for these values of x, and we are done. ■

4. Generalizations

Ramanujan not only proved Bertrand's postulate, but he generalized it to get bounds for when there are at least $2, 3, 4, \ldots$ primes between n and $2n$. To establish this, we will need the following important function:

Definition 4.1. The *prime-counting function*, denoted as $\pi(x)$, is the number of primes not more than x.

And now we can begin with our theorem, using some theorems we proved in the last section.

Theorem 4.2. *For every positive integer k, there exists a constant C such that for any positive integer $n > C$, there are at least k primes between n and $2n$ exclusive.*

Proof. First, note that

$$\vartheta(x) - \vartheta\left(\frac{x}{2}\right) = \sum_{\substack{p \text{ prime} \\ x/2 < p \leq x}} \log p \leq \sum_{\substack{p \text{ prime} \\ x/2 < p \leq x}} \log x = \left(\pi(x) - \pi\left(\frac{x}{2}\right)\right) \log x.$$

However, we already know that $\vartheta(x) - \vartheta\left(\frac{x}{2}\right) > \frac{1}{6}x - 3\sqrt{x}$ for $x \geq 570$, so

$$\pi(x) - \pi\left(\frac{x}{2}\right) > \frac{1}{\log x}\left(\frac{1}{6}x - 3\sqrt{x}\right).$$

Then if there are at least k primes between $\frac{x}{2}$ and x, we must have

$$\frac{1}{\log x}\left(\frac{1}{6}x - 3\sqrt{x}\right) > k - 1.$$

The LHS always increases when $x > 324$, so there exists a unique $C \geq 324$ such that

$$\frac{1}{\log C}\left(\frac{1}{6}C - 3\sqrt{C}\right) = k - 1.$$

If $C < 570$, then we can let $C = 570$, which works. We can check that C satisfies our desired conditions, so we are done. ■

We can use this statement along with looking at primes to find that $\pi(x) - \pi\left(\frac{x}{2}\right) \geq 1, 2, 3, 4, 5, \ldots$ when $x \geq 2, 11, 17, 29, 41, \ldots$. However, what about the other direction? To clarify, what is the minimum value of k such that there is at least one prime between x and kx for all sufficiently large n?

Theorem 4.3. *For any $\epsilon > 0$, there exists a number x_0 such that at least one prime lies between the closed interval $(x, x + \epsilon x)$ for all $x > x_0$,*

In order to prove this, we need the Prime Number Theorem, stated below:

Theorem 4.4 (Prime Number Theorem)**.**

$$\pi(x) = \frac{x}{\log x} + o\left(\frac{x}{\log x}\right)$$

where $\pi(x)$ is defined as before.

The proof of this theorem goes beyond what is talked about in this paper, so for our purposes we will assume that the result is true.

Proof of Theorem 4.3. With the Prime Number Theorem, we have

$$\pi(x + \epsilon x) - \pi(x) = \frac{x + \epsilon x}{\log x + \log(1 + \epsilon)} - \frac{x}{\log x} + o\left(\frac{x}{\log x}\right)$$
$$\pi(x + \epsilon x) - \pi(x) = \frac{\epsilon x}{\log x} + o\left(\frac{x}{\log x}\right)$$

or expressed as a limit,

$$\lim_{x \to \infty} \frac{\pi(x + \epsilon x) - \pi(x)}{x/\log x} = \epsilon.$$

Thus, $\pi(x + \epsilon x) - \pi(x)$ goes to infinity as x increases, which means for sufficiently large x_0, $\pi(x + \epsilon x) - \pi(x) > 0$ for all $x > x_0$. ■

References

[1] S. Sadov. "Euler's derivation of the Euler-Maclaurin formula." `http://www.math.mun. ca/~sergey/Winter03/m3132/Asst8/eumac.pdf`, 2003.
[2] Zyymat. "Ramanujan's proof of Bertrand's Postulate." `https://www.zyymat.com/ramanujans-proof-of-bertrands-postulate.html`, 2016.
[3] Zander. "Stirling's Approximation and Ramanujan's Proof of Bertrand's Postulate." `https://math.stackexchange.com/questions/342520`, StackExchange, 2013.
[4] G. H. Hardy, E. M. Wright. *An Introduction to the Theory of Numbers, 6th Edition.* Oxford University Press, 2008.

Euler Circle, Palo Alto, CA 94306
E-mail address: `kevinxu144@gmail.com`