Prime numbers are a concept that have intrigued mathematicians and scholars alike since the dawn of mathematics. A prime number is a natural number greater than 1 that has no positive divisors besides 1 and itself. In more simpler terms, it is a positive number that can not be factored. One of the characteristics of prime numbers that has been the focus of many mathematicians attention throughout time has been the distribution of prime numbers across the number line. Many people have found it fascinating that the distribution of prime numbers do not seem to follow a specific noticable pattern. Prime numbers are abundant at the beginning of the number line, but they appear much less often amoung larger numbers. For example, forty percent of the first 10 numbers are prime- 2, 3, 5, 7. However, among 10 digit numbers, prime numbers only account for 4 percent of the numbers. Among large numbers, the average gap between prime numbers is about 2.3 times the number of digits; so among 100 digit numbers, the expected gap is about 230 between primes. This is only an estimate. Often times, the gap is either much closer or way further than the average predicts. For example prime numbers often don’t appear for hundreds and hundreds of digits. On the other end of the spectrum, “twin primes” or primes that differ by only 2 (e.g. 3 and 5 or 11 and 13) appear randomly throughout the number line. Twin primes become much more scarce among larger numbers, but they never seem to completely disappear. It is believed that there infinitely many twin primes, with the largest known set being 3,756,801,695,685 x 2^{666,669} − 1 and 3,756,801,695,685 x 2^{666,669} + 1. In 1849, French mathematician Alphonse de Poliagnac extended this theory to the idea that there exists infinitely many prime pairs for any finite gap, not just 2. In May 2013, Dr. Yitang Zhang, a relatively unknown mathematician from the math department at Purdue University, came out with a proof using number sieves that proved that there existed a gap, N, of any distance up to 70 million between any two prime numbers.

Dr. Zhang’s ground breaking proof is only the most recent in the long, extensive history of dealing with prime numbers in mathematics. The roots of the study of prime numbers date back all the way to 300 BC, when the prominent Greek mathematician Euclid, in his elements, proved that there is no largest prime, implying that there are infinitely many prime numbers. Over the course of the following centuries, mathematicians unsuccessfully sought to find a formula that could produce an infinite sequence of primes. Soon, after many failed attempts at finding this formula, mathematicians began so search for an equation that could describe the general distribution of primes. Thus, the search for the prime number theorem began. The prime number theorem is perhaps one of the most sought after proofs in the history of mathematics. The prime number theorem gives an asymptotic form of counting the number of prime numbers less than some integer x. This is expressed in the form:

$$\pi(x)$$

which represents the number of primes less than or equal to x. This can also be expressed in the form

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

One of the earliest forms of the prime number theorem appeared in 1798 as a hypothesis by the French mathematician Adrien-Marie Legendre. Based on his study of a table of prime numbers up to 1,000,000, Legendre stated if x ≤ 1,000,000 then:
Carl Friedrich Gauss, a great German mathematician, also made a similar claim about prime numbers during the same time. In 1792, when Gauss was only 16 years old, he noticed that the density of primes decreased to approximately $\frac{1}{\log x}$. This led to his proposition that:

$$\pi(x) = \frac{x}{\ln(x) - 1.08366}$$

Given that the two functions are equivalent, the above statement can also be written as

$$\lim_{x \to \infty} \frac{\pi(x)}{\int_2^x \frac{dt}{\log(t)}} = 1$$

The first person to show that $\pi(x) \sim \frac{x}{\log(x)}$ was Pafnuty Chebyshev in 1852. He used manipulations of the Riemann zeta function for real values of $\Re(s)$, similar to the earlier work of Euler. He proved a slightly weaker form of the statement, that if

$$\lim_{x \to \infty} \frac{\pi(x)}{\int_2^x \frac{dt}{\log(t)}}$$

exists at all, then it has to be equal to 1.

Although he never able to actually prove the Prime Number Theorem using his method, he was able to prove Bertand’s posulate, that there exists a prime number between $n$ and $2n$ for $p \leq 2$. Many consider Chebyshev the father of the Prime Number Theorem because he paved the way and contributed a lot to our understanding of the theorem and many mathematicians were able to add onto his contributions to eventually prove the Prime Number Theorem.

The first proof of the Prime Number Theorem appeared in 1896 when the French mathematicians Jacques-Salomon Hadamard and Charles de la Valee Poussin independently used complex analysis to apply Hadamard’s theory of integral functions applied to the zeros of the Riemann zeta function. The Riemann zeta, or Euler-Riemann zeta function is a extremely complex function with variable $s$ that continues to the sum of the infinite series. The proper definition of this function can be stated as $\zeta(s)$ is a function of a complex variables $s = \sigma + it$. This can be shown as:

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

This proof in no way was considered elementary. Hadamard himself, in a lecture he was delivering to the Mathematical Society of Copenhagen in 1921 stated:

"No elementary proof of the prime number theorem is known, and one may ask whether it is reasonable to expect one.... if anyone produces an elementary proof of the prime number theorem, he will show that the subject does not hang together in the way we have supposed."
We then fast forward more than 25 years to the year 1948. The mathematical community was flipped upside
down when Paul Erdös announced that he and Atle Selberg had together discovered an elementary proof of the
Prime Number theorem. He also announced that he and Selberg would together be publishing their discovery
in a joint proof.

**Selberg vs. Erdös Controversy**

During their time, both Selberg and Erdös were prominent mathematicians. Selberg was born in Norway
on June 14, 1917. He was the son of a teacher and a mathematician. Academics was an important part of
Selbergs life growing up, he along with all of his brothers, attended college and pursued careers in the math and
sciences. One of his brothers also became a mathematician while the other became an engineering professor.
Selberg was also a war veteran. During World War II, he fought for his homeland against the German invasion
of Norway and was even imprisoned several times. During that time, he also worked on mathematics, but
had to work in isolation, due to the German occupation of his country. After the war, his contributions and
discoveries in mathematics became well known, boosting his reputation.

Erdös was a Hungarian mathematician born on March 26, 1913. He was a very prolific mathematician,
known for both his mathematical accomplishments and his eccentric lifestyle. Erdós was a a very social
mathematician and had collaborated with over 500 mathematicians during his career. Like Selberg, both of
Erdös’s parents were well educated teachers. Unfortunately, his father, along with much of his family died,
died during the holocaust. Many people who knew him very well described him as a simplisic person with a
passion for what he did.

The fact that both Erdös and Selberg were very well known and respected in their field became one of the major
reasons the dispute between them escalated so quickly (along with the fact that they had both solved one of
the most sought after proof in mathematics). Although details of what happened between them have become
blurred over time, there is still much we do know. The dispute began after the announcement that both of them
had come up with an elementary proof of the prime number theorem, something that mathematicians had
sought after for centuries. Selberg and Erdös first met in 1948 through Paul Turan, a hungarian mathematician
with close ties to the two. It was at one of his discussions, where he lectured on Selberg’s sieve method and
fundamental formula, where they met. Erdös approached Selberg and told him about his discovery of how

\[
\lim_{n \to \infty} \frac{P(n+1)}{P(n)} = 1
\]

To this, Selberg replied that Erdös’s discovery was essential in Selberg’s efforts to create an elementary proof of
the prime number theorem. That is when the two began working together. The two eventually collaborated to
create a rudimentary proof of the prime number theorem, which relied solely on an extensive use of logarithms.
During the creation of this proof, many discrepancies arose between the two. Selberg eventually became the
one who became against the idea of a joint proof and pushed for the idea of the two publishing their work
separately. However, Erdös was against this idea because it would ultimately give credit to Selberg for the
elementary proof of the prime number theorem. Selberg has also been accused of mistreating Erdös and
undermining his contributions. Although, according to a letter from Selberg to Erdös, he would give him
credit in his paper, Selberg was still going to be the one who was going to get the credit for the Prime Number
Theorem. He also refused to allow Erdös to mention the prime Number Theorem in his publication.

Conflict continued between the two as Selberg refused the notion of the two publishing a joint proof. Erdös
claimed that if he were to keep his above limit to himself, he would’ve eventually reached the Prime Number
Theorem, while Selberg would have still been left alone with his Fundamental inequality. This irritated Erdös
as he believed that he was entitled to a joint proof where he would recieve more credit than Selberg wanted
to give him.
Erdős eventually gave into the Selberg’s demands, and Selberg was able to publish his own proof of the prime number theorem. And he received credit for coming up with the first elementary proof of the prime number theorem. Selberg would go on to win the Fields medal in 1950 for his work in the Prime number theorem, the Riemann zeta functions, along with his work with his Selberg sieve method. Erdős also won the Frank Nelson Cole Prize in 1951 for his work in number theory.

Preliminaries

Although this proof of the prime number theorem is advertised as elementary, and the most difficult mathematics in this paper is only the simplest usage of logarithms, background information is still needed to understand all the conversions and substitutions incorporated into this proof.

1. A prime number is a natural number greater than 1 that has no positive divisors other than 1 and itself. A natural number greater than 1 that is not prime is called a composite number. The first 10 primes are relatively easy to calculate. They are: 2, 3, 5, 7, 11, 13, 17, 19, 23, 29. The prime number theorem is based on determining the number of primes in between a group of numbers.

2. The term log is used quite extensively in this proof. Any place where the notation log is written, it will actually indicate the natural log, or ln(x) function.

3. n, k, x, y, y' = Variables for any number: This theorem incorporates a large amount of non-specific specific numbers, variables such as those shown above are used to represent different natural numbers.

4. p = Prime Number Variables: Similar to the variables shown above, the Prime number theorem doesn’t focus on one specific prime number, so other variables are set aside to represent prime numbers, these include p, r and q.

5. δ = Delta: In this paper, delta is seen as a very small number less than one.

\[ \delta < 1 \]

6. |x| = Absolute value: The absolute value symbol represents the distance of a number away from zero on the number line. This means that |−25| = 25 = 25. The absolute value is used often in the proof due to the fact that R(x) can be either positive or negative.

7. \( \sum \) = Sum: The summation notation is used to represent the sum of all quantities in a given set of parameters. The data given under and above the symbol are called the limits, The lower limit describes the starting point and the upper limit indicates the ending point for which the sum should be accounted for. For example, in the notation:

\[ \sum_{b}^{c} f = f(b) + f(b + 1) + f(b + 2) + f(b + 3) + \ldots + f(c) \]

ex: \[ \sum_{i}^{3} x^2 + 1 = 2 + 5 + 10 = 17 \]

8. O(x) = Landau Notation: This symbol, named after its inventor, the German mathematician Edmund Landau, is used to describe the rate of growth of a given function. O(x) is used in Selbergs proof as an error term. An important concept explained in the proof is that O(x) does not increase faster than x.

9. Another important concept is the Chebyshev theta function, denoted by \( \vartheta \). The first Chebyshev function states:
\[ \vartheta(x) = \sum_{p \leq x} \log p \]

with the sum extending over all prime numbers \( p \) that are less than or equal to \( x \).

The Chebyshev function calculates the sum of the natural logarithms of all prime numbers less than or equal to that desired number, \( x \). Let us look at Chebyshev function using the example integer, 10, for \( x \):

\[ \vartheta(x) = \sum_{p \leq 10} \log p = \log 2 + \log 3 + \log 5 + \log 7 \]

The second Chebyshev function

\[ \psi(x) = \sum_{p^k \leq x} \log p \]

is defined similarly, with the sum extending over all prime powers not exceeding \( x \).

The second Chebyshev function is also sometimes written as

\[ \psi(x) = \sum_{p^k \leq x} \Lambda(n) \]

where \( \Lambda(n) \) is the von Mangoldt function which states:

\[ \Lambda(n) = \begin{cases} 
\log(p) & \text{if } n = p^k \text{ for some prime } p \text{ and integer } k \leq 1 \\
0 & \text{otherwise} 
\end{cases} \]

The Chebyshev function is influential because it allows a summation to be written in a much more concise manner than otherwise possible with summation notation. Though the theta function denotes a very specific summation, it becomes a key tool in Selberg’s proof since the summation of the natural logarithms of prime numbers appears often in the proof of the prime number theorem.

10. Here we will become familiar with Selberg’s \( R \) notation. Selberg creates the \( R \) notation just for his proof to make his proof more concise by denoting a recurrent, and otherwise longer operation. The \( R \) notation denotes the following:

\[ R(x) = \vartheta(x) - x \]

Hence, the \( R(x) \) denotes the subtraction of some real number denoted by variable \( x \) from the sum created when \( x \) is applied to the Chebyshev formula described in preliminary 9. We write the following example to further expound on the function:

\[ R(10) = \vartheta(10) - 10 = \sum_{p \leq 10} \log p - 10 = \log 2 + \log 3 + \log 5 + \log 7 - 10 \]
Selberg’s $R$ notation is unique to his proof of the Prime Number Theorem, and it has not become widely spread as common notation in mathematics. The $R$ notation is utilized to rewrite the Prime Number Theorem as well.

**Outline of Selberg’s Elementary Proof of the Prime Number Theorem**

Selberg’s Proof of the Prime number theorem consists of four main parts.

*Part A*

In Part A, Selberg begins by giving an explanation and the proof of his formula:

$$\vartheta(x) \log x + \sum_{p \leq x} \vartheta \left( \frac{x}{p} \right) \log p = 2x \log x + O(x)$$

This relation is shown to be the breakthrough discovery that Selberg uses as the foundation for the rest of his proof of the Prime Number theorem.

*Part B*

The main outcome derived from Part A was Selberg’s relation:

$$\vartheta(x) \log x + \sum_{p \leq x} \vartheta \left( \frac{x}{p} \right) \log p = 2x \log x + O(x)$$

This relation yields for the function $R(x)$:

$$R(x) \log x = - \sum_{pq \leq x} R \left( \frac{x}{pq} \right) \log p + O(x)$$

By applying a couple partial summations to Selberg’s relation, we can also obtain:

$$\sum_{p \leq x} \log p + \sum_{pq \leq x} \left( \frac{\log \log q}{\log p} \right) = 2x + O \left( \frac{x}{\log x} \right)$$

This relation becomes useful for two reasons. The first is that by applying Mertens relation we get:

$$R(x) \log x = \sum_{pq \leq x} \frac{\log \log q}{\log(pq)} R \left( \frac{x}{pq} \right) + O(x \log \log x)$$

The second is that when we combine the two expressions for $R(x) \log x$, we get:

$$2 |R(x)| \log x \leq \sum_{p \leq x} \log p \left|R \left( \frac{x}{p} \right) \right| + \sum_{pq \leq x} \frac{\log \log q}{\log(pq)} \left|R \left( \frac{x}{pq} \right) \right| + O(x \log \log x)$$

From there, after applying one more partial summation and simplifying, we derive the inequality:
\[ |R(x)| \log x \leq \sum_{n \leq x} \left| R\left(\frac{x}{n}\right) \right| + O(x\log \log x) \]

from Selbergs formula.

Which can also be rewritten as:

\[ |R(x)| \leq \frac{1}{\log x} \sum_{n \leq x} \left| R\left(\frac{x}{n}\right) \right| + O\left(\frac{x\log \log x}{\log x}\right) \]

**Part C**

After acquiring the following new inequality, we can establish the existence of a constant \( K > 0 \) so that for any \( \delta > 0 \) and any \( x > e^{\frac{k}{\delta}} \), the interval \((x, xe^{\frac{k}{\delta}})\) contains a subinterval of the form \((y, ye^{\frac{\delta}{2}})\) so that for every \( z \) in this last interval, we have:

\[ \left| R(z) \right| < 4\delta z \]

**Part D**

Showing that if \( a < 8 \) is a positive number then the inequality \( \left| R(z) \right| < ax \) for \( x \) large enough, leads to the inequality:

\[ \left| R(z) \right| < a \left(1 - \frac{a^2}{300K}\right)x \]

for \( x \) large enough.

**Focus on Part One of Selberg’s Proof**

In this elementary exploration of the Prime Number Theorem (PNT) we are going to explain a specific section of Selberg’s PNT. The overall point of the PNT is to prove that

\[ \lim_{x \to \infty} \frac{\vartheta(x)}{x} = 1 \]

where for all \( x > 0 \), \( \vartheta(x) \) is defined as:

\[ \vartheta(x) = \sum_{p \leq x} \log p \]

with \( p \) defining all prime numbers.

The focus of this section is to show that
If we look at the number line, we see that the number of prime numbers as we approach infinity is always increasing. However, as we search further, the number of primes we find become fewer and more spread apart. If you were to graph this, having the number of primes on the vertical axis, and the search size on the x-axis, you would see that the curve of this line, which represents the number of primes, never flattens out. It’s slope does decrease, but it always remains positive, and its always gradually increasing. You can calculate the density of the graph by dividing the number of primes, denoted as \( \pi(x) \), over the search size, which can be represented as \( x \).

The following statement for determining the density of primes is shown below:

\[
\frac{\text{number of primes}}{\text{size}} = \frac{\pi(x)}{x}
\]

For example if we were to look at the first 100 integers, we would find 25 primes, showing that 25 percent are prime. For the first 1,000 integers we find 1229 primes, showing that 12.29 percent are prime. And finally, for the first 100 million integers, 5.76 percent are prime. As we approach infinity, the density of prime numbers drop, however, this drop is asymptotic in nature because the rate at which it drops also decreases.

This behavior relates to another well known mathematical concept, and that is the behavior of logarithms. The graph of natural logarithms is can be written as \( y = \ln(x) \). If we look at the graph of natural logarithms, as \( x \) approaches infinity, the natural logarithm of \( x \) increases, but the rate at which it increases becomes smaller and smaller, similar to the behavior of the density of primes. If we were to invert the natural log equation, and write it as:

\[
y = \frac{1}{\log(x)}
\]

we find the same curve created as when we plot the density of primes.

This gives a formula that shows the density of primes up to some integer \( x \).

\[
\text{Density of primes} = \frac{\text{number of primes}}{x} = \frac{1}{\log(x)}
\]

If one were to want to calculate the density of number of primes from 1 to 1 trillion, one could write the following equation:

\[
\frac{\text{number of primes}}{x} = \frac{1}{\log(1000000000000)} = \frac{1}{32.23} = 3.1\text{percent}
\]

When compared to the actual result, the actual density of the number of primes from 1 to 1 trillion is approximately 3.2 percent. As \( x \) increases, the current error of 0.1 percent approaches zero, giving more accurate estimates.

By manipulating the equation for the density of primes, one can rewrite the previous equation as:
number of primes = size \times \text{density}

in mathematical notation, this is expressed as:

\[
\pi(x) = \frac{x}{\log(x)}
\]

giving us one of the fundamental equations for the prime number theorem for calculating the number of primes less than or equal to some integer \(x\).

The actual equation for the Prime number theorem proved in Selbergs paper is:

\[
\lim_{x \to \infty} \frac{\pi(x)}{\frac{x}{\log x}} = 1
\]

In the paper, Selberg states that it is sufficient to prove the theorem in the form:

\[
\lim_{x \to \infty} \frac{\vartheta(x)}{x} = 1
\]

If we take the above statement that

\[
\lim_{x \to \infty} \frac{\pi(x)}{\frac{x}{\log x}} = 1
\]

, using simple algebra, we can rewrite it as

\[
\lim_{x \to \infty} \frac{\pi(x) \log(x)}{x} = 1
\]

If we also take into consideration as shown earlier that \(\pi(x) = \sum_{p \leq x} 1\) we get

\[
\lim_{x \to \infty} \frac{\sum_{p \leq x} \log p}{x} = 1
\]

Finally, taking the Chebyshev function shown in preliminary 9, we get

\[
\lim_{x \to \infty} \frac{\vartheta(x)}{x} = 1
\]

as desired.

From here it is relatively simple to check the claim that the Prime Number theorem is also equivalent to the statement
\[
\lim_{x \to \infty} \frac{\psi(x)}{x} = 1
\]
given in preliminary 9 as well.

We know that \( \psi(x) = \sum_{p^n \leq x} \log p \). If we want to know how many powers of \( p \) are less than or equal to \( x \) for a given \( p \leq x \), we can suppose that \( p^k \) is the largest such power. In mathematical terms, we can suppose that \( p^k \leq x < p^{k+1} \). If we take the logarithms of all sides, we get:

\[
k \log p \leq \log x < (k + 1) \log p
\]

After dividing the entire inequality by \( \log p \), we get

\[
k \leq \frac{\log x}{\log p} < k + 1
\]

implying that \( k = \left\lfloor \frac{\log x}{\log p} \right\rfloor \). This indicates that the term \( \log p \) appears \( \left\lfloor \frac{\log x}{\log p} \right\rfloor \) times in the sum that define \( \psi(x) \). We can then rewrite:

\[
\psi(x) = \sum_{p \leq x} \left\lfloor \frac{\log x}{\log p} \right\rfloor \log p
\]

After know this, we can immediately get

\[
\psi(x) \leq \sum_{p \leq x} \left\lfloor \frac{\log x}{\log p} \right\rfloor \log p = \sum_{p \leq x} \log x = \pi(x) \log x
\]

When simplified, the above inequality implies:

\[
\vartheta(x) \leq \psi(x) \leq \pi(x) \log x
\]

which when expanded gives us

\[
\frac{\vartheta(x)}{x} \leq \frac{\psi(x)}{x} \leq \frac{\pi(x) \log x}{x}
\]

Finally, from the above inequality, we can assume their limits as \( x \) become arbitrarily large are equivalent as well, and we get:

\[
\lim_{x \to \infty} \frac{\vartheta(x)}{x} \leq \lim_{x \to \infty} \frac{\psi(x)}{x} \leq \lim_{x \to \infty} \frac{\pi(x) \log x}{x}
\]

as desired.
Conclusion

The Elementary Proof of the Prime Number theorem is one of the most influential pieces of mathematical work of modern times. Although Selberg and Erdös are credited with the proof of this theorem, it took enormous effort and time from multiple mathematicians from different eras to eventually come up with it. The Prime Number Theorem is a prime (no pun intended) example of how mathematics is bounded by no single person, group or time. The expansion and further understanding mathematics relies on contributions from everyone, and benefits society as a whole. The Prime Number Theorem is just one example of how, with enough time and effort, society is able to do anything, and achieve what has previously been thought of as the impossible.