An Introduction to The Twin Prime Conjecture

Allison Berke

December 12, 2006

Abstract

Twin primes are primes of the form $(p, p+2)$. There are many proofs for the infinitude of prime numbers, but it is very difficult to prove whether there are an infinite number of pairs of twin primes. Most mathematicians agree that the evidence points toward this conclusion, but numerous attempts at a proof have been falsified by subsequent review. The problem itself, one of the most famous open problems in mathematics, has yielded a number of related results, including Brun's conjecture, Mertens' theorems, and the Hardy-Littlewood Conjecture. Along with these conjectures, there are a number of results which are easier to arrive at, but nevertheless help mathematicians think about the infinitude of primes, and the special properties of twin primes. This paper will introduce the aforementioned conjectures associated with the twin prime conjecture, and work through some exercises that illuminate the difficulties and intricacies of the twin prime conjecture.

1 Introduction: The Original Conjecture and Failed Proofs

The term *twin prime* was coined by Paul Stackel in the late nineteenth century. Since that time, mathematicians have been interested in the properties of related primes, both in relation to number theory as a whole, and as specific, well-defined problems. One of the first results of looking at twin primes was the discovery that, aside from $(3, 5)$, all twin primes are of the form $6n \pm 1$. This comes from noticing that any prime greater than 3 must be of the form $6n \pm 1$. To show this, note that any integer can be written as $6x + y$, where x is any integer, and y is $0, 1, 2, 3, 4$ or 5. Now consider each y value individually. When $y = 0$, $6x + y = 6x$ and is divisible by 6. When $y = 1$ there are no immediately recognizable factors, so this is a candidate for primacy. When $y = 2$, $6x + 2 = 2 \cdot (3x + 1)$, and so is not prime. For the case when $y = 3 : 6x + 3 = 3 \cdot (2x + 1)$ and is not prime. When $y = 4 : 6x + 4 = 2 \cdot (3x + 2)$ and is not prime. When $y = 5, 6x + 5$ has no immediately recognizable factors, and is the second candidate for primacy. Then all primes can be represented as either $6n + 1$ or $6n - 1$, and twin primes, since they are separated by two, will have to be $6n - 1$ and $6n + 1$.

Further research into the conjecture has been concerned with finding expressions for a form of the prime counting function $\pi(x)$ that depend on the twin prime constant. The prime counting function is defined as

$$\pi(x) = \{N(p)|p \leqq x\}$$

where $N(p)$ denotes the number of primes, p. One motivation for defining the prime counting function is that it can be used to determine a formula for the size of the intervals between primes, as well as giving us an indication of the rate of decay by which primes thin out in higher numbers. It has been shown algebraically that the prime counting function increases asymptotically with the logarithmic integral [12]. In the following expression, $\pi_2(x)$ refers to the number of primes of the form p and $p+2$ greater than x, and $\prod_2$ is the twin prime constant, which is defined by the expression $\prod(1 - \frac{1}{9p-1)^2})$ over primes $p \geqq 2$. The term $O(x)$, meaning "on the order of x," is defined as follows: if $f(x)$ and $g(x)$ are two functions defined on the same set, then $f(x)$ is $O(g(x))$ as x goes to infinity if and only if there exists some x_0 and some M such that $|f(x)| \leqq M|g(x)|$ for x greater than x_0. This expression for the twin prime counting function is

$$\pi_2(x) \leqq c\Pi_2 \frac{x}{(\ln(x))^2}[1 + O(\frac{\ln(\ln(x))}{\ln(x)})] \tag{1}$$

which is the best that has been proven thus far. The constant c in (1) has been reduced to 6.8325, down from previous values as high as 9 [12]. The formation of this inequality involves two of Merten's theorems which will be discussed in the following section. Hardy and Littlewood [3] have conjectured that $c = 2$, and using this assumption have formulated what is now called the Strong Twin Prime Conjecture. In the following expression, $a \sim b$ means that $\frac{a}{b}$ approaches 1 at the limits of the expressions a and b. In this case, the limit is as x approaches infinity.

$$\pi_2(x) \sim 2\Pi_2 \int_2^x \frac{dx}{(\ln(x))^2}. \tag{2}$$

A necessary condition for the strong conjecture (2) is that the prime gaps constant, $\Delta \equiv \limsup_{n\to\infty} \frac{p_{n+1}-p_n}{p_n}$ be equal to zero. The most recent attempted proof of the twin prime conjecture was that of Arenstorf, in 2004 [1], but an error was found shortly after its publication, and it was withdrawn, leaving the conjecture open to this day.

2 Mertens' Theorems

A number of important results about the spacing of prime numbers were derived by Franz Mertens, a German mathematician of the late nineteenth and early twentieth century. The following proofs of Mertens' conjectures lead up to the result that the sum of the reciprocals of primes diverges, which will contrast with Brun's conjecture, that the sum of the reciprocals of twin primes converges. First, we should briefly show that the primes are infinite, for otherwise the implications of Mertens' theorems are not obvious. Euclid's proof of this postulate, his second theorem, is as follows.

Let $2, 3, 5, ..., p$ be an enumeration of all prime numbers up to p, and let $q = (2 \cdot 3 \cdot 5 \cdot ... \cdot p) + 1$. Then q is not divisible by any of the primes up to and including p. Therefore, it is either prime or divisible by a prime between p and q. In the first case, q is a prime greater than p. In the second case, the divisor of q between p and q is a prime greater than p. Then for any prime p, this construction gives us a prime greater than p. Thus, the number of primes must be infinite [4]. Now we can resume with Mertens' theorems.

Mertens Theorem 1:

For any real number $x \geq 1$,

$$0 \leq \sum_{n \leq x} \ln(\frac{x}{n}) < x. \tag{3}$$

Proof:

The function $f(t) = \ln(\frac{x}{t})$ is decreasing on the interval $[1, x]$, so

$$\sum_{1 \leq n \leq x} \ln(\frac{x}{n}) < \ln(x) + \int_1^x \ln(\frac{x}{t})dt$$

We can rewrite the right-hand side of the inequality as the following:

$$\ln(x) + \int_1^x \ln(\frac{x}{t})dt = x\ln(x) - \int_1^x \ln(t)dt.$$

Similarly, we can rewrite this:

$$x\ln(x) - \int_1^x \ln(t)dt = x\ln(x) - (x\ln(x) - x + 1) < x.$$

For Mertens' second theorem, we introduce the Von Mangoldt's function, $\Lambda(n)$, where

$$\Lambda(n) = \ln(p) \text{ if } n = p^m \text{ is a prime power, and zero otherwise.}$$

Then the psi function of the prime number theorem is defined as follows

$$\Psi(x) = \sum_{1 \leq m \leq x} \Lambda(m).$$

Mertens Theorem 2:
For any real number $x \geq 1$,

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} = \ln(x) + O(1). \tag{4}$$

Proof:
Let $N = [x]$. Then

$$0 \leq \sum_{n \leq x} \ln(\frac{x}{n}) = N \ln(x) - \sum_{n=1}^{N} \ln(n) = x \ln(x) - \ln(N!) + O(\ln(x)) < x$$

$$\ln(N!) = x \ln(x) + O(x).$$

The proof of this equation, which is integral to the theorem, is shown on the next page. Let $v_p(n)$ denote the highest power of p, a prime, that divides n. Then

$$\ln(N!) = \sum_{p \leq N} v_p(N) \ln(p)$$

We can rewrite this as a single summation, by combining the limits on p and k:

$$\ln(N!) = \sum_{p \leq N} \sum_{k=1}^{[\frac{\ln(N)}{\ln(p)}]} [\frac{N}{p^k}] \ln(p).$$

For ease of notation, we replace N with x, and substitute Von Mangoldt's function:

$$\ln(x!) = \sum_{n \leq x} [\frac{x}{n}] \Lambda(n).$$

Now we can remove the floor function brackets by introducing an error term on the order of one, since the maximum error obtained by removing the floor function is less than one:

$$\ln(x!) = \sum_{n \leq x} (\frac{x}{n} + O(1)) \Lambda(n).$$

We can distribute this term, forming two sums, one in the error term:

$$\ln(x!) = x \sum_{n \leq x} \frac{\Lambda(n)}{n} + O(\sum_{n \leq x} \Lambda(n)).$$

Now we can substitute in the Psi function defined earlier:

$$\ln(x!) = x \sum_{n \leq x} \frac{\Lambda(n)}{n} + O(\Psi(x)).$$

Since the Psi function is of the same order as a linear function in x, we can replace it in the error term, obtaining the following:

$$\ln(x!) = x\sum_{n\leq x}\frac{\Lambda(n)}{n} + O(x).$$

Therefore,

$$x\sum_{n\leq x}\frac{\Lambda(n)}{n} + O(x) = x\ln(x) + O(x)$$

$$\sum_{n\leq x}\frac{\Lambda(n)}{n} = \ln(x) + O(1)$$

Mertens Theorem 3:

For any real number $x \geq 1$,

$$\sum_{p\leq x}\frac{\ln(p)}{p} = \ln(x) + O(1). \tag{5}$$

Proof:
From the previous theorem,

$$0 \leq \sum_{n\leq x}\frac{\Lambda(n)}{n} - \sum_{p\leq x}\frac{\ln(p)}{p} = \sum_{p^k\leq x, k\geq 2}\frac{\ln(p)}{p^k}$$

$$\leq \sum_{p\leq x}\ln(p)\sum_{k=2}^{\infty}\frac{a}{p^k} \leq \sum_{p\leq x}\frac{\ln(p)}{p(p-1)}$$

$$\leq 2\sum_{p\leq x}\frac{\ln(p)}{p^2} \leq 2\sum_{n=1}^{\infty}\frac{\ln(n)}{n^2} = O(1)$$

It then follows from the previous theorem that

$$\sum_{p\leq x}\frac{\ln(p)}{p} = \sum_{n\leq x}\frac{\Lambda(n)}{n} + O(1) = \ln(x) + O(1).$$

Finally, the last and longest of Mertens' theorems presented here brings us to our desired conclusion.

Mertens Theorem 4:

There exists a constant $b_1 > 0$ such that

$$\sum_{p\leq x}\frac{1}{p} = \ln(\ln(x)) + b_1 + O(\frac{1}{\ln(x)}), x \geq 2. \tag{6}$$

Proof:
We can write

$$\sum_{p\leq x}\frac{1}{p}=\sum_{p\leq x}\frac{\ln(p)}{p}\frac{1}{\ln(p)}=\sum_{n\leq x}u(n)f(n)$$

where $u(n)=\frac{\ln(p)}{p}$ if $n=p$, and 0 otherwise, and $f(t)=\frac{1}{\ln(t)}$. We define new functions $U(t)$ and $g(t)$ as follows

$$U(t)=\sum_{n\leq t}u(n)=\sum_{p\leq t}\frac{\ln(p)}{p}=\ln(t)+g(t)$$

Then $U(t)=0$ for $t<2$ and $g(t)=O(1)$ by the previous theorem. Then the integral $\int_2^\infty\frac{g(t)}{(t\ln(t)^2)}dt$ converges absolutely, and

$$\int_x^\infty\frac{g(t)dt}{t(\ln(t))^2}=O(\frac{1}{\ln(x)}).$$

We know that $f(t)$ is continuous and $U(t)$ is increasing, so we can express $\sum_{p\leq x}\frac{1}{p}$ as a Riemann integral. $U(t)=0$ for $t<2$, so by partial sums, Nathanson [8] shows the conclusion of this theorem:

$$\sum_{p\leq x}\frac{1}{p}=\sum_{n\leq x}u(n)f(n)=\frac{1}{2}+\int_2^x f(t)dU(t)$$

Integrating by parts, we obtain the following:

$$\frac{1}{2}+\int_2^x f(t)dU(t)=f(x)U(x)-\int_2^x U(t)df(t)=\frac{\ln(x)+g(x)}{\ln(x)}-\int_2^x U(t)f'(t)dt.$$

Now we can simplify the term outside the integral, and substitute in for $U(t)$:

$$\frac{1}{2}+\int_2^x f(t)dU(t)=1+O(\frac{1}{\ln(x)})+\int_2^x\frac{\ln(t)+g(t)}{t(\ln(t))^2}dt.$$

We can split the integral in order to simplify the result:

$$\frac{1}{2}+\int_2^x f(t)dU(t)=\int_2^x\frac{1}{t\ln(t)}dt+\int_2^\infty\frac{g(t)}{t(\ln(t))^2}dt-\int_x^\infty\frac{g(t)}{t(\ln(t))^2}dt+1+O(\frac{1}{\ln(x)}).$$

Now we can evaluate two of the integrals:

$$\int_2^x\frac{1}{t\ln(t)}dt+\int_2^\infty\frac{g(t)}{t(\ln(t))^2}dt=\ln(\ln(x))-\ln(\ln(2))$$

Finally, we can simplify this result in terms of a variable b_1:

$$\ln(\ln(x))-\ln(\ln(2))+\int_2^\infty\frac{g(t)}{t(\ln(t))^2}dt+1+O(\frac{1}{\ln(x)})=\ln(\ln(x))+b_1+O(\frac{1}{\ln(x)})$$

where
$$b_1 = 1 - \ln(\ln(2)) + \int_2^{\infty} \frac{g(t)}{t(\ln(t))^2} dt.$$
Now thanks to Mertens, we not only know that the sum of reciprocals of primes diverges, we can see that it diverges like the function $\ln(\ln(x))$.

3 Brun's Conjecture

An important step towards proving the twin prime conjecture is the realization, first made by Brun in 1919, that the sum of the reciprocals of odd twin primes converges to a definite number. While this fact does not limit the number of twin primes, it shows that they are distributed infrequently among the real numbers.

Brun's Conjecture:

Let p_1, p_2, ... be the sequence of prime numbers p such that $p+2$ is also prime. Then
$$\sum_{n=1}^{\infty} (\frac{1}{p_n} + \frac{1}{p_n+2}) = (\frac{1}{3} + \frac{1}{5}) + (\frac{1}{5} + \frac{1}{7}) + (\frac{1}{11} + \frac{1}{13}) + \ldots < \infty \qquad (7)$$

Proof:
The proof of the theorem depends on a result which will be shown in the exercises and assumed here, namely that $\pi_2(x) << \frac{x}{(\ln(x))^{\frac{3}{2}}}$ for all $x \geq 2$. Then
$$n = \pi_2(p_n) < \frac{p_n}{(\ln(p_n))^{\frac{3}{2}}} \leq \frac{p_n}{(\ln(n))^{\frac{3}{2}}}$$
for $n \geq 2$. Then
$$\frac{1}{p_n} < \frac{1}{n(\ln(n))^{\frac{3}{2}}}.$$
It follows that the series defined above is convergent:
$$\sum_{n=1}^{\infty} \frac{1}{p_n} \leq \frac{1}{3} + \sum_{n=2}^{\infty} \frac{1}{p_n} << \frac{1}{3} + \sum_{n=2}^{\infty} \frac{1}{n(\ln(n))^{\frac{3}{2}}}.$$
Then we have a stronger bound for $\pi_2(x)$ than the one discussed earlier: we can now write that $\pi_2(x) = O(\frac{x}{\ln(\ln(x))})$, as shown in Nathanson [8] The exact constant to which Brun's series converges is unknown, but its value has been calculated to be approximately 1.9021605824 [5]. The constant's value gives an idea of how infrequent twin primes actually are. Had Brun's series diverged, it would have indicated immediately that there are an infinite number of twin primes. The fact that the series converges does not allow one to reach a conclusion about the infinitude of twin primes, but it makes the problem more difficult.

4 The Hardy-Littlewood Conjecture

The first Hardy-Littlewood Conjecture, also known as the $k - tuple$ conjecture, states that the number of prime constellations can be computed. A *prime constellation* is a sequence of primes $p_1, p_2, ..., p_n$ where the difference between the first and last primes, $p_n - p_1$, is s. This s is the smallest number such that there exist n integers in an interval of length s, and, for every prime q, at least one of the residues $p_i \pmod q$ is not one of these n integers. Hardy and Littlewood have then conjectured that

$$\pi_2(x) \sim 2\Pi_2 \int_2^x \frac{dt}{ln(t)^2} \tag{8}$$

Here Π_2 refers to the twin prime constant introduced before, which is

$$\Pi_2 = \prod_{p>3} \frac{p(p-2)}{(p-1)^2} \approx 0.66016181584686957392781211001455\ldots [12]$$

The formulation of the Hardy-Littlewood conjecture builds upon some of the techniques used to prove Brun's conjecture, namely the Brun sieve techniques. The Brun sieve can be constructed as follows: Let X be a nonempty, finite set of N objects, and let $P_1, \ldots, P_r$ be r different properties that the elements of the set X might have. Let N_0 denote the number of elements of X that have none of these properties. For any subset $I = \{i_1, \ldots, i_k\}$ of $\{1, 2, \ldots, r\}$, let $N(1) = N(i_1, \ldots, i_k)$ denote the number of elements of X that have each of the properties $P_{i_1}, P_{i_2}, \ldots, P_{i_k}$. Let $N(\emptyset) = |X| = N$. If m is a nonnegative even integer, then

$$N_0 \leq \sum_{k=0}^{m} (-1)^k \sum_{|I|=k} N(I). \tag{9}$$

If m is a nonnegative odd integer, then

$$N_0 \geq \sum_{k=0}^{m} (-1)^k \sum_{|I|=k} N(I).[8] \tag{10}$$

The proof given in Nathanson [8] is as follows. Let x be an element of the set X, and suppose that x has exactly l properties P_i. If $l = 0$, then x is counted once in N_0 and once in $N(\emptyset)$, but is not counted in $N(I)$ if I is nonempty. If $l \geq 1$, then x is not counted in N_0. By renumbering the properties, we can assume that x has the properties $P_1, P_2, \ldots, P_l$. Let $I \subseteq \{1, 2, \ldots, l, \ldots, r\}$. If $i \in I$ for some $i > l$, then x is not counted in $N(I)$. If $I \subseteq \{1, 2, \ldots, l\}$ then x contributes 1 to $N(I)$. For each $k = 0, 1, \ldots, l$, there are exactly $\binom{l}{k}$ such subsets with $|I| = k$. If $m \geq l$, then the element x contributes

$$\sum_{k=0}^{l} (-1)^k \binom{l}{k} = 0$$

to the right sides of the inequalities. If $m < l$, then x contributes

$$\sum_{k=0}^{m} (-1)^k \binom{l}{k}$$

to the right sides of the inequalities. This contribution is positive if l is even and negative if l is odd.
The Hardy-Littlewood conjecture makes use of this sieve by considering a set of primes with properties corresponding to the size of the intervals between consecutive members. Then the Brun sieve is used to make generalizations about these properties based on the primes themselves.
The Hardy-Littlewood conjecture is compatible with the twin prime conjecture as originally stated, and is in fact sometimes substituted in as the strong twin prime conjecture, because it is understood that its proof would involve the integral going to infinity, and therefore π_2 going to infinity as well.

5 Exercises

An interesting exercise involving the distribution of primes is determining the number of primes in an interval. To formulate a corresponding expression, we can start out by rewriting $\pi(x + \epsilon x) - \pi(x)$ in terms of Euler's approximation. Then

$$\pi(x + \epsilon x) - \pi(x) = \frac{x + \epsilon x}{\ln(x) + \ln(1 + \epsilon)} - \frac{x}{\ln(x)} + O(\frac{x}{\ln(x)}).$$

We can rewrite the right-hand side as

$$\frac{\epsilon x}{\ln(x)} + O(\frac{x}{\ln(x)})$$

Then if we let $\epsilon = 1$,

$$\pi(2x) - \pi(x) = \frac{x}{\ln(x)} + O(\frac{x}{\ln(x)})$$

Where $O(\frac{x}{\ln(x)}) \sim \pi(x)$. At first, it appears that this is saying that the number of primes in an interval is equal to the number of primes in a sequential interval of equal length, which is misleading, because we know from experience and computation that the primes thin out as we progress through the numbers. However, a better formulation of the result of this exercise would be that

$$\pi(2x) - 2\pi(x) = O(\pi(x))$$

and this result is not inconsistent with known tables of primes.

Earlier, we used a result discovered by the German mathematician Brun, which we can prove here. Brun's conjecture is as follows:

$$\pi_2(x) << \frac{x(\ln(\ln(x)))^2}{(\ln(x))^2} \tag{11}$$

To prove this, we use the Brun sieve to find an upper bound for $N_0(x, y)$, where here N_0 is the number of integers such that $n(n+2)$ is prime, and x is a prime less than y but greater than 5. Then by the Brun sieve, with m an even integer such that $1 \leq m \leq r$, it is shown in Narkiewicz [7] that

$$N_0(x, y) \leq \sum_{k=0}^{m} (-1)^k \sum_{|I|=k} N(I)$$

$$\leq \sum_{k=0}^{m} (-1)^k \sum_{\{i_1, \dots i_k\}} \left(\frac{2^k y}{p_{i_1}, \dots, p_{i_k}} + O(2^k)\right)$$

$$\leq y \sum_{k=0}^{m} \sum_{\{i_1, \dots i_k\}} \frac{(-2)^k}{p_{i_1}, \dots, p_{i_k}} + \sum_{k=0}^{m} (-1)^k \binom{r}{k} O(2^k)$$

$$\leq y \sum_{k=0}^{r} \sum_{\{i_1, \dots i_k\}} \frac{(-2)^k}{p_{i_1}, \dots, p_{i_k}} - y \sum_{k=m+1}^{r} \sum_{\{i_1, \dots i_k\}} \frac{(-2)^k}{p_{i_1}, \dots, p_{i_k}} + O\left(\sum_{k=0}^{m} \binom{r}{k} (2^k)\right).$$

By another of Brun's theorems,
Theorem:

$$y \sum_{k=0}^{r} \sum_{\{i_1, \dots i_k\}} \frac{(-2)^k y}{p_{i_1}, \dots, p_{i_k}} << \frac{y}{(\ln(x))^2}$$

Nathanson [8] supplies the bound for the second term: Let $s_k(x_1, \dots, x_r)$ be the elementary symmetric polynomial of degree k in r variables. For any nonnegative real numbers $x_1, \dots, x_r$, we have

$$s_k(x_1, \dots, x_r) = \sum_{\{i_1, \dots i_k\}} x_{i_1} \dots x_{i_k}$$

$$\leq \frac{(x_1 + \dots + x_r)^k}{k!} = \frac{(s_1(x_1, \dots, x_r))^k}{k!}$$

$$< \frac{e}{k}^k s_1(x_1, \dots, x_r)^k.$$

Therefore,

$$|y \sum_{k=m+1}^{r} \sum_{\{i_1, \dots i_k\}} \frac{(-2)^k}{p_{i_1}, \dots, p_{i_k}}| \leq y \sum_{k=m+1}^{r} \sum_{\{i_1, \dots i_k\}} \frac{(-2)^k}{p_{i_1}, \dots, p_{i_k}}$$

$$\leq y \sum_{k=m+1}^{r} \sum_{\{i_1, \dots i_k\}} \left(\frac{2}{p_{i_1}}\right) \dots \left(\frac{2}{p_{i_k}}\right)$$

$$= y \sum_{k=m+1}^{r} s_k\left(\frac{2}{p_1}, \dots, \frac{2}{p_r}\right)$$

$$< y \sum_{k=m+1}^{r} (\frac{e}{k})^k s_1(\frac{2}{p_1}, \dots, \frac{2}{p_r})^k$$

$$< y \sum_{k=m+1}^{r} (\frac{2e}{m})^k (\sum_{p \le x} \frac{1}{p})^k$$

$$< y \sum_{k=m+1}^{r} (\frac{c \cdot \ln(\ln(x))}{m})^k$$

where c is a positive constant. If we choose m such that $m > 2c \cdot \ln(\ln(x))$, then

$$y \sum_{k=m+1}^{r} (\frac{c \cdot \ln(\ln(x))}{m})^k \le x \sum_{k=m+1}^{r} \frac{1}{2^k} < \frac{x}{2^m}$$

For the last term,

$$\sum_{k=0}^{m} \binom{r}{k} 2^k < \sum_{k=0}^{m} (2r)^k << (2r)^m \le x^m.$$

Combining these three terms, we get

$$\pi_2(y) << x + \frac{y}{(\ln(x))^2} + \frac{y}{2^m} + x^m << \frac{y}{(\ln(x))^2} + \frac{y}{2^m} + x^m$$

where again x is any number less than y but greater than 5, and m is any even integer such that $m > 2c \cdot \ln(\ln(x))$. If we let $c' = max\{2c, (\ln(2)^{-1})\}$, and let

$$x = e^{(\frac{\ln(y)}{3c' \cdot \ln(\ln(y))})} = y^{\frac{1}{3c' \cdot \ln(\ln(y))}}$$

$$m = 2[c' \cdot \ln(\ln(y))]$$

Then since

$$\ln(x) = \frac{\ln(y)}{3c' \cdot \ln(\ln(y))}$$

$$\frac{y}{(\ln(x))^2} << \frac{y(\ln(\ln(y)))^2}{(\ln(y))^2}$$

Since $c' \ge (\ln(2))^{-1}$, and $m = 2[c' \cdot \ln(\ln(y))] > 2c' \cdot \ln(\ln(y)) - 2$,

$$\frac{y}{2^m} < \frac{4y}{2^{2c' \cdot \ln(\ln(y))}} = \frac{4y}{(\ln(y))^{2c' \cdot \ln(2)}} \le \frac{4y}{(\ln(y))^2}$$

Then

$$x^m \le x^{2c' \cdot \ln(\ln(y))} = exp(\frac{2c' \cdot \ln(\ln(y) \ln(y))}{3c' \cdot \ln(\ln(y))}) = y^{\frac{2}{3}}$$

Finally,

$$\pi_2(x) << \frac{x(\ln(\ln(x)))^2}{(\ln(x))^2}.$$

6 Conclusion

The twin prime conjecture may never be proven, but studying the properties of twin primes is certainly a rewarding exercise. Recent work on the twin prime conjecture by Dan Goldston and Cem Yilidrim has focused on creating expressions for the gap size between primes, and in particular focusing on the expression

$$\Delta = \liminf_{n\to\infty} \frac{p_{n+1} - p_n}{\ln(p_n)} = 1$$

Research into better expressions for the interval between consecutive primes is currently being conducted at Stanford, sponsored by the American Institute of Mathematics [12]. Though number theory has been the foundation of many different branches of higher mathematics, its fundamental problems remain interesting and fruitful for researchers interested in the properties of prime numbers.

References

[1] Arenstorf, R. F. "There Are Infinitely Many Prime Twins." 26 May 2004. http://arxiv.org/abs/math.NT/0405509.

[2] Guy, R. K. "Gaps between Primes. Twin Primes." A8 in Unsolved Problems in Number Theory, 2nd ed. New York: Springer-Verlag, pp. 19-23, 1994.

[3] Hardy, G. H. and Littlewood, J. E. "Some Problems of 'Partitio Numerorum.' III. On the Expression of a Number as a Sum of Primes." Acta Math. 44, 1-70, 1923.

[4] Hardy, G. H. and Wright, E. M. An Introduction to the Theory of Numbers, 5th ed. Oxford, England: Clarendon Press, 1979.

[5] Havil, J. Gamma: Exploring Euler's Constant. Princeton, NJ: Princeton University Press, pp. 30-31, 2003.

[6] Miller, S. J. and Takloo-Bighash, R. An Invitation to Number Theory. Princeton, NJ: Princeton University Press ,pp. 326-328, 2006.

[7] Narkiewicz, W. The Development of Prime Number Theory. Berlin, Germany: Springer Press, 2000.

[8] Nathanson, M. B. Additive Number Theory. New York, New York: Springer Press, 1996.

[9] Ribenboim, P. The New Book of Prime Number Records. New York: Springer-Verlag, pp. 261-265, 1996.

[10] Shanks, D. Solved and Unsolved Problems in Number Theory, 4th ed. New York: Chelsea, p. 30, 1993.

[11] Tenenbaum, G. "Re Arenstorf's paper on the Twin Prime Conjecture." 8 Jun 2004.

[12] Weisstein, Eric W. "Twin Prime Conjecture" http://mathworld.wolfram.com/TwinPrimeConjecture.html, 2006.

[13] Young, R. M. Excursions in Calculus. The Mathematical Association of America, 1992.