

Definitive Proof of the Twin-Prime Conjecture

Kenneth A. Watanabe, PhD

December 27, 2018

1 Abstract

A twin prime is defined as a pair of prime numbers (p_1, p_2) such that $p_1 + 2 = p_2$. The Twin Prime Conjecture states that there are an infinite number of twin primes. The first mention of the Twin Prime Conjecture was in 1849, when de Polignac made the more general conjecture that for every natural number k , there are infinitely many primes p such that $p + 2k$ is also prime. The case where $k = 1$ is the Twin Prime Conjecture [1]. In this document, I derive a function that corresponds to the number of twin primes less than n for large values of n . Then by proof by induction, it is shown that as n increases indefinitely, the function also increases indefinitely thus proving the Twin Prime Conjecture. Using the same methodology, de Polignac's conjecture is also shown to be true.

2 Functions

Before we get into the proof, let me define a couple of functions that are necessary.

Let the function $l(x)$ represent the largest prime number less than x . For example, $l(10.5) = 7$, $l(20) = 19$ and $l(19) = 17$.

Let the function $g(x)$ represent the next higher prime number greater than x . For example, $g(10.5) = 11$, $g(20) = 23$ and $g(23) = 29$.

Let capital P represent all pairs (x, y) such that $x + 2 = y$ and x is an odd number > 1 and $y \leq n$. The values of x or y need not be prime.

3 Background

To prove the Twin Prime conjecture, I must first find a function that determines the number of twin primes less than n . Then by proof by mathematical induction, prove that the number of primes increases with increasing n .

To find all the twin primes less than n , I first start with the set of pairs of odd numbers less than or equal to odd integer n , and pair them (x, y) such that each pair $x + 2 = y$. We eliminate the pairs $(1, 3)$ since 1 is not considered a prime number. For a given odd integer n , we see that there are $(n - 3)/2$ pairs. This give us the following set:

$\{(3,5), (5,7), (7,9), (9,11), (11,13), (13,15), (15,17), (17,19), (19,21), (21,23), (23,25), (25,27), (27,29) \dots (n-4,n-2),(n-2,n)\}$

Next we must eliminate the pairs where the x or y coordinate is evenly divisible by 3 but not equal to 3. Then we eliminate pairs divisible by 5, 7, 11 etc until we reach $l(\sqrt{n})$. There are no prime numbers greater than $l(\sqrt{n})$ that could evenly divide the x or y coordinate that is not already divisible by a lower prime. The remaining pairs will be the twin primes.

We start by eliminating the pairs where the x or y coordinate is divisible by 3, but x or y is not equal to 3. It is easy to see that every third pair starting with $(9,11)$ has an x coordinate that is divisible by 3 (yellow) and that every third pair starting with $(7,9)$ has a y coordinate that is divisible by 3 (orange). There is no instance where both x and y are divisible by 3.

$\{(3,5), (5,7), (7,9), (9,11), (11,13), (13,15), (15,17), (17,19), (19,21), (21,23), (23,25), (25,27), (27,29) \dots (n-4,n-2),(n-2,n)\}$

There are $\lfloor (P - 1)/3 \rfloor$ pairs where the x coordinate is divisible by 3 and $x \neq 3$. There are $\lfloor P/3 \rfloor$ pairs where the y coordinate is divisible by 3. Therefore, in total, there are $\lfloor (P - 1)/3 \rfloor + \lfloor P/3 \rfloor$ pairs where either the x or y coordinates are divisible by 3 but not equal to 3. As P gets very large, the value of $P - 1$ approaches P and the number of pairs divisible by 3 approaches $(2/3)P$.

The number of pairs divisible by 3 $\lim_{n \rightarrow \infty} = (2/3) \times P$.

Next, we eliminate the pairs where the x or y coordinate is divisible by 5, and x or y is not equal to 5. It is easy to see that every fifth pair starting with $(15,7)$ has an x coordinate that is divisible by 5 (yellow) and that every fifth pair starting with $(13,15)$ has a y coordinate that is divisible by 5 (orange).

$\{(3,5), (5,7), (7,9), (9,11), (11,13), (13,15), (15,17), (17,19), (19,21),$

(21,23), (23,25), (25,27), (27,29), (29,31), (31,33), (33,35), (35,37) ...
 $(n-4, n-2), (n-2, n)$

There are $\lfloor (P-2)/5 \rfloor$ pairs where x coordinate is divisible by 5 and $x \neq 5$. There are $\lfloor (P-1)/5 \rfloor$ pairs where y is divisible by 5 and $y \neq 5$. So there are $\lfloor (P-2)/5 \rfloor + \lfloor (P-1)/5 \rfloor$ pairs where either the x or y coordinates are divisible by 5 but not equal to 5. As P gets very large, the values of $P-2$ and $P-1$ approach P and the number of pairs divisible by 5 approaches $(2/5)P$.

Notice however, that every third pair (green) where the x coordinate is divisible by 5, the x coordinate is also divisible by 3.

(5,7), (15,17), (25,27), (35,37), (45,47), (55,57), (65,67), (75,77), (85,87) ...

Likewise, every third pair where the y coordinate is divisible by 5, is also divisible by 3.

(3,5), (13,15), (23,25), (33,35), (43,45), (53,55), (63,65), (73,75), (83,85) ...

So to avoid double counting, the number of pairs divisible by 5 but not by 3 approaches the following equation as n gets very large.

$$\text{Number of pairs divisible by only 5 } \lim_{n \rightarrow \infty} = (1/3)(2/5) \times P.$$

Next, we eliminate the pairs where the x or y coordinate is divisible by 7, and x or y is not equal to 7. For pairs where the x or y coordinate is divisible by 7, it is easy to see that every seventh pair starting with (21,23) has an x coordinate that is divisible by 7 (yellow)

(7,9), (21,23), (35,37), (49,51), (63,65), (77,79), (91,93), (105,107) ...

Likewise, every seventh pair starting with (19,21) has a y coordinate that is divisible by 7 (orange).

(5,7), (19,21), (33,35), (47,49), (61,63), (75,77), (89,91), (103,105) ...

Note that every third pair is divisible by 3 and every fifth pair is divisible by 5. So to avoid double counting, the number of pairs divisible by 7 and not by 3 or 5, approaches the following equation as n gets very large.

$$\text{Number of pairs divisible by only 7 } \lim_{n \rightarrow \infty} = (1/3)(3/5)(2/7) \times P.$$

The general formula for number of pairs divisible by prime number p is as follows

Number of pairs divisible by only p $\lim_{n \rightarrow \infty} = (1/3)(3/5)(5/7)\dots(l(p)-2)/l(p)(2/p) \times P$.

or

Number of pairs divisible by only p $\lim_{n \rightarrow \infty} = P \times (2/p) \prod_{q=3}^{l(p)} ((q-2)/q)$.

where the product is over prime numbers only.

To find the total number of non-prime pairs, we must sum up all the pairs evenly divisible by a prime number. The total number of non-prime pairs less than or equal to n can be defined as follows

Total number of non-prime pairs $\lim_{n \rightarrow \infty} = P \times \sum_{p=3}^{l(\sqrt{n})} (2/p) \prod_{q=3}^{l(p)} ((q-2)/q)$

where the sum and products are over prime numbers only.

Subtracting the number of non-prime pairs from the total number of pairs gives the number of twin primes less than or equal to n . We will denote the number of twin primes less than n as $\pi_2(n)$.

$$\begin{aligned} \pi_2(n) \lim_{n \rightarrow \infty} &= P - P \times \sum_{p=3}^{l(\sqrt{n})} (2/p) \prod_{q=3}^{l(p)} ((q-2)/q) \\ &\text{or} \\ \pi_2(n) \lim_{n \rightarrow \infty} &= P[1 - \sum_{p=3}^{l(\sqrt{n})} (2/p) \prod_{q=3}^{l(p)} ((q-2)/q)] \end{aligned}$$

Let us define the function $W(x)$ equal the following:

$$\begin{aligned} W(x) &= (1/3) + \\ & (1/5) \times (1/3) + \\ & (1/7) \times (1/3) \times (3/5) + \\ & (1/11) \times (1/3) \times (3/5) \times (5/7) + \\ & (1/13) \times (1/3) \times (3/5) \times (5/7) \times (9/11) + \\ & \dots \\ & (1/x) \times (1/3) \times (3/5) \times (5/7) \times (9/11) \times \dots \times (l(x)-2)/l(x) \end{aligned}$$

This can be expressed as the following equation:

$$W(x) = \sum_{p=3}^x (1/p) \prod_{q=3}^{l(p)} ((q-2)/q)$$

Using this function, the expression for number of pairs that contain a non-prime number can be simplified to

$$\text{Number of non-twin-primes} = 2P \times W(l(\sqrt{n}))$$

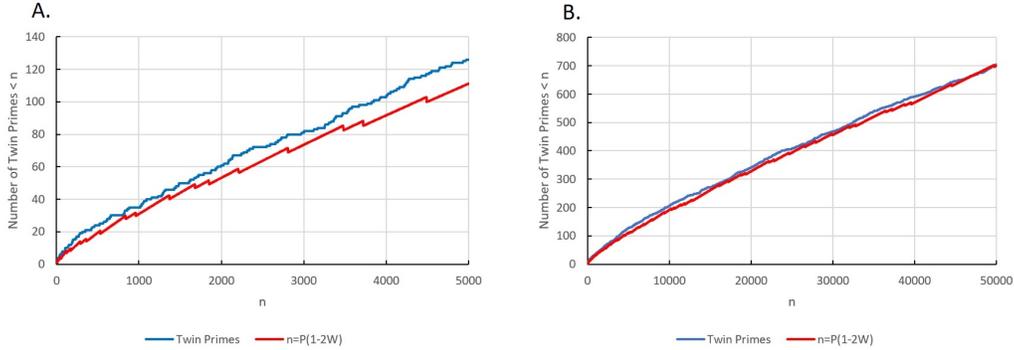


Figure 1: The actual number of twin primes (blue line) is underestimated by the equation $\pi_2(n) = P(1 - 2W(\sqrt{n}))$ (red line) for values of $n < 5,000$. But as n gets larger, the equation $\pi_2(n) = P(1 - 2W(\sqrt{n}))$ approaches the actual number of twin primes.

$$\text{Number of twin-primes} = \pi_2(n) = P - 2P \times W(l(\sqrt{n}))$$

$$\text{Equation 1: } \pi_2(n) = P[1 - 2W(l(\sqrt{n}))]$$

To verify that the derivation of equation 1 was correct and to determine at what point the equation begins to accurately determine the number of twin primes, I plotted the actual number of twin primes less than n (blue line) and equation 1 (orange line) (Figure 1) for all values of n up to 50,000. As can be seen in the graph, the actual number of twin primes is underestimated by equation 1 for values of $n < 5,000$. This is not a problem since this errs on the side of caution. But as n increases, equation 1 very closely estimates the number of twin primes. For large values of n , the lines lie almost directly on top of each other, indicating that the number of twin primes less than n can be accurately predicted by equation 1.

4 The Proof of the Twin Prime Conjecture

I will use proof by mathematical induction to prove the Twin Prime Conjecture is true. To do this we must first get $(1 - 2W(p_{i+1}))$ in terms of $W(p_i)$. To do this, we must look at the actual values of $2W(p_i)$.

$$2 \times W(3) = (2/3)$$

$$\begin{aligned}
2 \times W(5) &= (2/3) + (2/5) \times (1/3) \\
2 \times W(7) &= (2/3) + (2/5) \times (1/3) + (2/7) \times (1/3) \times (3/5) \\
2 \times W(11) &= (2/3) + (2/5) \times (1/3) + (2/7) \times (1/3) \times (3/5) + (2/11) \times (1/3) \times \\
&\quad (3/5) \times (5/7) \\
&\text{Etc . . .}
\end{aligned}$$

Therefore, the values of $1 - 2W(p_i)$ are as follows:

$$\begin{aligned}
1 - 2W(3) &= 1 - (2/3) = 1/3 \\
1 - 2W(5) &= [1 - (2/3)] - (2/5)(1/3) = (1/3)(3/5) \\
1 - 2W(7) &= [1 - (2/3) - (2/5)(1/3)] - (2/7)(1/3)(3/5) = (1/3)(3/5)(5/7) \\
1 - 2W(11) &= [1 - (2/3) - (2/5)(1/3) - (2/7)(1/3)(3/5)] - (2/11)(1/3)(3/5)(5/7) \\
&= (1/3)(3/5)(5/7)(9/11)
\end{aligned}$$

Notice the value of $1 - 2W(p_i)$ (yellow) can be substituted into the green part of $1 - 2W(p_{i+1})$. Therefore, these equations can be simplified to:

$$\mathbf{Equation\ 2:} \quad [1 - 2 \times W(p_{i+1})] = [(p_{i+1} - 2)/p_{i+1}] \times [1 - 2 \times W(p_i)]$$

Another way to think about how we get to equation 2 is by cutting away pieces from a pie.

The pie has a value of 1. We cut away $2/3^{rds}$ from the pie leaving $1/3$.

Now from **this piece**, we cut $2/5^{ths}$ away leaving $3/5^{ths}$ of $1/3$.

Now from **this piece**, we cut $2/7^{ths}$ away leaving $5/7^{ths}$ of the last piece.

Now from **this piece**, we cut $2/11^{ths}$ away leaving $9/11^{ths}$ of the last piece.

For each iteration, we cut away $2/p^{ths}$ leaving $(p-2)/p$ of the previous piece, thus resulting in equation 2.

We must first prove that equation 1 is always greater than 0. Since we know P is greater than 0, we must prove that $(1 - 2W(p_i)) > 0$ for any prime number p_i .

The base case $p_0 = 3$.

$$1 - 2W(p_0) = 1 - 2W(3) = 1 - 2/3 = 1/3 \text{ which is greater than 0.}$$

Now assuming that Equation 1 is true $1 - 2W(p_i) > 0$, we must prove that $1 - 2W(p_{i+1}) > 0$.

Assumption: $1 - 2W(p_i) > 0$

Prove: $1 - 2W(p_{i+1}) > 0$

Substituting p_{i+1} into Equation 2 gives:

$$1 - 2W(p_{i+1}) = [(p_{i+1} - 2)/p_{i+1}] \times [1 - 2W(p_i)]$$

Since we assumed that $[1 - 2W(p_i)]$ is greater than 0, and $[(p_{i+1} - 2)/p_{i+1}]$ is greater than 0, then the product of $[(p_{i+1} - 2)/p_{i+1}]$ and $[1 - 2W(p_i)]$ must be greater than 0. Therefore, equation 1 is greater than 0 for p_{i+1} and thus greater than 0 for all p_i .

Now we can prove the number of twin primes increases indefinitely with increasing values of n . Let's say we have P pairs of adjacent odd numbers less than n . If we approximate $P = n/2$ for large n , we get the following equation:

$$\pi_2(n) = P[1 - 2W(l(\sqrt{n}))] = (n/2)[1 - 2W(l(\sqrt{n}))]$$

For prime number $p_i = l(\sqrt{n})$, for large n we can approximate $n \approx p_i^2$. Actually, n will be at least p_i^2 , but approximating $n \approx p_i^2$ errs on the side of caution.

Substituting p_i for $l(\sqrt{n})$ and p_i^2 for n in the above equation we get:

$$\pi_2(p_i^2) = (p_i^2/2)[1 - 2W(p_i)]$$

The number of twin primes less than p_{i+1}^2 is

$$\begin{aligned} \pi_2(p_{i+1}^2) &= (p_{i+1}^2/2)[1 - 2W(p_{i+1})] = \\ &= (p_{i+1}^2/2)[(p_{i+1} - 2)/p_{i+1}][1 - 2W(p_i)] = \qquad \text{Using equation 2} \\ &= [p_{i+1}(p_{i+1} - 2)/2][1 - 2W(p_i)] \end{aligned}$$

Let $\Delta\pi_2(p_i)$ represent the difference between the number of twin primes less than p_i^2 and the number of twin primes less than p_{i+1}^2 . Subtracting $\pi_2(p_i^2)$ from $\pi_2(p_{i+1}^2)$ gives us the following expression:

$$\Delta\pi_2(p_i) = [p_{i+1}(p_{i+1} - 2)/2][1 - 2W(p_i)] - (p_i^2/2)[1 - 2W(p_i)]$$

or

$$\textbf{Equation 3: } \Delta\pi_2(p_i) = [1 - 2W(p_i)]/2 \times \{[p_{i+1}(p_{i+1} - 2)] - (p_i^2)\}$$

Since we have proven that $[1 - 2W(p_i)] > 0$, and p_{i+1} is greater than p , the value of $[p_{i+1}(p_{i+1} - 2)] - (p_i^2)$ is always greater than 0. Thus, the product

is greater than 0.

Since the product is greater than 0, this proves that the number of twin primes less than $n = p_{i+1}^2$ is greater than the number of twin primes less than $n = p_i^2$. However, the term $[1 - 2W(p)]$ approaches 0 as p gets very large and it may be the case that $\Delta\pi_2(p_i)$ may approach 0. If this was the case, then this does not show that the number of twin primes increases indefinitely.

So the next question is, what is the lower bound on $\Delta\pi_2(p_i)$. The cases where $\Delta\pi_2(p_i)$ is minimal is when $p_{i+1} = p_i + 2$. This is because the difference between $[p_{i+1}(p_{i+1} - 2)]$ and (p_i^2) increases dramatically as the difference between p_{i+1} and p_i increases. So substituting $p_i + 2$ for p_{i+1} into the term $[p_{i+1}(p_{i+1} - 2)] - (p_i^2)$ will give us the equation for the lower bound.

$$\begin{aligned} p_{i+1}(p_{i+1} - 2) - p_i^2 &= (p_i + 2)(p_i + 2 - 2) - p_i^2 \\ &= (p_i + 2)p_i - p_i^2 \\ &= p_i^2 + 2p_i - p_i^2 \\ &= 2p_i \end{aligned}$$

Substituting $2p_i$ for $(p_{i+1}(p_{i+1} - 2) - p_i^2)$ into equation 3 for $\Delta\pi_2(p_i)$ gives us a new equation for the lower bound on $\Delta\pi_2(p_i)$.

$$\textbf{Equation 4: } \Delta\pi_2(p_i)* = p_i(1 - 2W(p_i))$$

where $\Delta\pi_2(p_i)*$ represents the lower bound on $\Delta\pi_2(p_i)$.

To validate that no errors were made, I graphed $\Delta\pi_2(p_i)$ versus p (blue line) and $\Delta\pi_2(p_i)*$ versus p (orange line) in Figure 2. Notice that the lower bound $\Delta\pi_2(p_i)*$ coincides with $\Delta\pi_2(p_i)$ exactly at the points where $p_{i+1} = p_i + 2$.

If we show the lower bound $\Delta\pi_2(p_i)*$ is always greater than 1, then we know that $\Delta\pi_2(p_i)$ will always be greater than 1. We will show this by mathematical induction.

Base case for $\Delta\pi_2(p_0)*$:

Using $p_0 = 3$, we get the following

$$\Delta\pi_2(p_0)* = 3(1 - 2W(3)) = 3(1 - 2(1/3)) = 1$$

Next, we assume that $\Delta\pi_2(p_i)* \geq 1$, and prove that $\Delta\pi_2(p_{i+1})* \geq 1$
 $\Delta\pi_2(p_i)* = p_i(1 - 2W(p_i)) \geq 1$

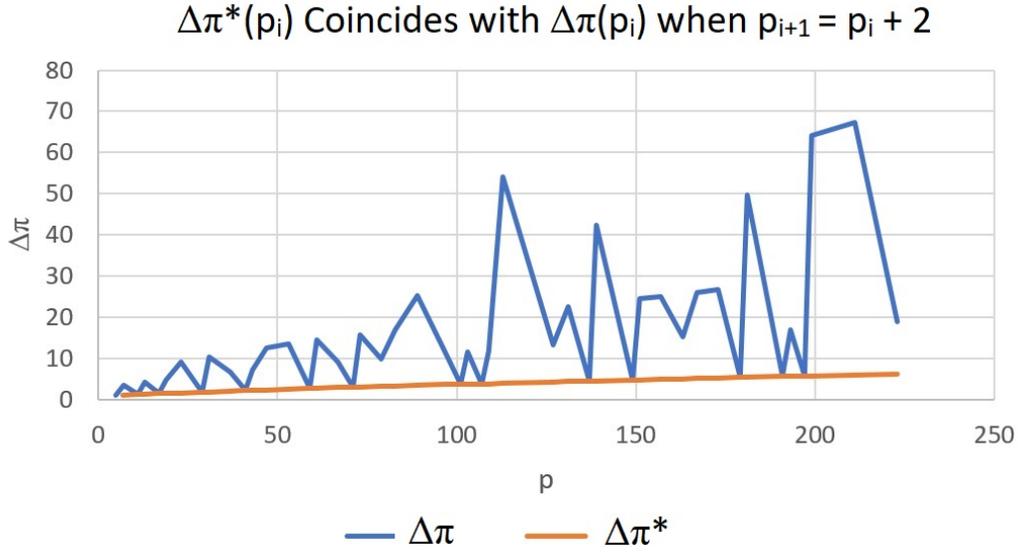


Figure 2: Graph of $\Delta\pi_2(p_i)$ and the lower bound $\Delta\pi_2(p_i)^*$ versus p . Notice that when $p_{i+1} = p_i + 2$, $\Delta\pi_2(p_i)$ coincides with $\Delta\pi_2(p_i)^*$.

Substituting p_{i+1} into the above equation gives:

$$\Delta\pi_2(p_{i+1})^* = p_{i+1}(1 - 2W(p_{i+1}))$$

$$\Delta\pi_2(p_{i+1})^* = p_{i+1}[(p_{i+1} - 2)/p_{i+1}](1 - 2W(p_i)) \quad \text{Using equation 2}$$

$$\Delta\pi_2(p_{i+1})^* = (p_{i+1} - 2)(1 - 2W(p_i))$$

Taking the ratio of $\Delta\pi_2(p_{i+1})^* / \Delta\pi_2(p_i)^*$ gives us the following:

$$\Delta\pi_2(p_{i+1})^* / \Delta\pi_2(p_i)^* = (p_{i+1} - 2)(1 - 2W(p_i)) / (p_i(1 - 2W(p_i)))$$

$$\Delta\pi_2(p_{i+1})^* / \Delta\pi_2(p_i)^* = (p_{i+1} - 2) / p_i$$

Since p_{i+1} is at least equal to $p_i + 2$, the ratio $\Delta\pi_2(p_{i+1})^* / \Delta\pi_2(p_i)^*$ must be greater than or equal to 1. Therefore, the number of twin primes always increases by at least 1 with increasing p_i , and since there are an infinite number of prime numbers p_i , there are an infinite number of twin primes. QED

Note: This also provides evidence for the conjecture that for any p_i there is at least 1 twin prime pair between p_i^2 and $(p_i + 2)^2$. In fact, it may be the case that for any n , there is at least 1 twin prime pair between n^2 and $(n + 2)^2$.

5 Proof of Polignac's Conjecture

The Twin Prime Conjecture is a special case for de Polignac's conjecture where $k = 1$. To prove there are an infinite number of quad primes, i.e. $k = 2$, the odd pairs can be partitioned as follows:

$(3,7), (5,9), (7,11), (9,13), (11,15), (13,17), \dots (n-8,n-4),(n-6,n-2),(n-4,n)$.

Notice that as n gets large, the number of pairs approaches $n/2$ just like for the twin primes.

Eliminating the pairs where the x or y coordinates are divisible by a prime number will yield the quad primes. As it turns out, the equation for the number of quad primes is the exactly same as equation 1.

$$\pi_4(n) = P[1 - 2W(l(\sqrt{n}))]$$

where P is the number of pairs.

In fact, for all values of $k = 2^i$, it can be shown that the number of primes separated by 2^i is the same as the number of twin primes for very large values of n . This is because for any pair (x, y) , the x coordinate is relatively prime to the y coordinate. Thus, by proving the Twin Prime conjecture, we have also proven Polignac's Conjecture for all values of $k = 2^i$ where i is an integer greater than or equal to 0.

For values of $k \neq 2^i$, when partitioning out the odd pairs, when we eliminate the non-prime pairs, there is overlap. For example, if we take the case where $k = 3$, the set of sext primes, we get the following set:

$(3, 9), (5,11), (7,13), (9, 15), (11,17), (13,19), (15, 21) \dots (n-10,n-4),(n-8,n-2),(n-6,n)$.

Now when we eliminate the pairs divisible by 3, we only eliminate only about 1/3rd of the pairs rather than 2/3rds since every pair where the x coordinate is divisible by 3 (yellow), the y coordinate is also divisible by 3 (orange). Thus, the first term of the W function changes from 2/3 to 1/3. This results in a larger number of sext primes relative to number of twin primes. A similar situation holds true for dec primes (primes separated by 10). When eliminating the pairs divisible by 5, we only eliminate about 1/5th of the pairs rather than 2/5ths since every pair where the x coordinate is divisible by 5, the y coordinate is also divisible by 5. Thus the second term of the W function will change from $(1/3)(2/5)$ to $(1/3)(1/5)$. Since the number of sext

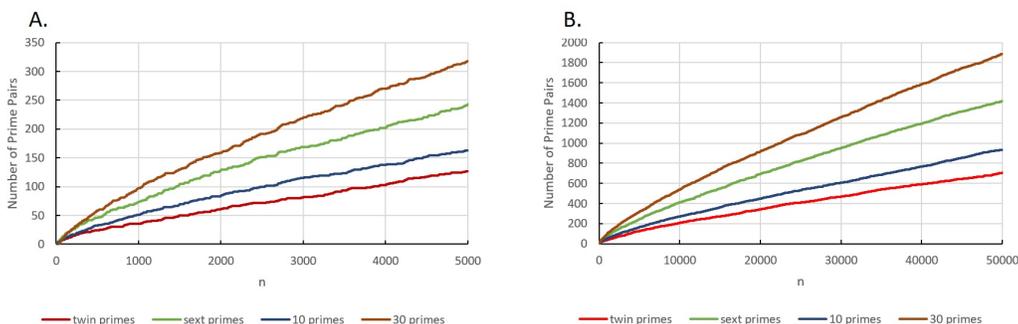


Figure 3: The more factors there are between primes, the more prime pairs exist. There are fewer twin primes (red line) than sext primes (green line), dec primes (blue line) and 30-primes (brown line)

primes, dec primes, 30-primes (primes pairs differing by 30) are larger than the number of twin primes, then Polignac's Conjecture is true for all values of k .

To illustrate this, I graphed the number of prime pairs less than n for twin primes, sext primes, dec primes and 30-primes in Figure 3. Notice that the curve for the twin primes has relatively the fewest number of prime pairs.

6 Summary

I have shown that the number of twin primes less than n approaches the following equation as n gets large:

$$\pi_2(n) = P[1 - 2W(\sqrt{n})]$$

where $l(\sqrt{n})$ is the largest prime number less than \sqrt{n} and $W(x)$ is defined as

$$W(x) = \sum_{(p=3)}^x (1/p) \prod_{(q=3)}^{(l(p))} ((q-2)/q)$$

I have shown by proof by induction, that the above equation for number of twin primes increase indefinitely as n increases the confirming the Twin Prime Conjecture.

References

- [1] Alphonse de Polignac *Recherches nouvelles sur les nombres premiers*
Comptes Rendus des Séances de l'Académie des Sciences (1849)

7 Copyright Notice

This document is protected by U.S. and International copyright laws. Reproduction and distribution of this document or any part thereof without written permission by the author (Kenneth A. Watanabe) is strictly prohibited.

Copyright © 2018 by Kenneth A. Watanabe