# Definitive Proof of the Twin-Prime Conjecture

Kenneth A. Watanabe, PhD

August 6, 2025

## 1 Abstract

A twin prime is defined as a pair of prime numbers $(p_x, p_y)$ such that $p_x + 2 = p_y$. The Twin Prime Conjecture states that there are an infinite number of twin primes. A more general conjecture by de Polignac states 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. In this document, the function $\Pi_2^*(n)$ is derived that closely approximates $\Pi_2(n)$, the actual number of twin primes less than $n$, for large values of $n$. Then by proof by induction on $\Pi_2^*(n)$ , it is shown that for any prime number $p_i$, there is at least one twin prime $(p_x, p_y)$ such that $p_i^2 < p_y < (p_{i+1})^2$. Since there are an infinite number of prime number $p_i$, this proves that there are an infinite number of twin primes, thus proving the Twin Prime Conjecture. Error analysis shows that the relative difference between $\Pi_2^*(n)$ and $\Pi_2(n)$ decreases as $n$ increases. Using this same methodology, the de Polignac Conjecture is also shown to be true.

## 2 Functions

Before we get into the proof, let us define the following functions:

Let the function $\lambda(x)$ represent the largest prime number less than or equal to $x$. For example, $\lambda(10.5) = 7$, $\lambda(20) = 19$ and $\lambda(19) = 19$.

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

Let $|P(n)|$ represent the number of elements in set $P(n)$.

Let $K_p(n)$ represent the number of pairs in $P(n)$ where the $x$ or $y$ coordinate is divisible by prime number $p$ and not divisible by prime numbers less than $p$.

Let $k_p(n)$ represent the number of pairs in $P(n)$ where the $x$ or $y$ coordinate is divisible by prime number $p$.

Let $\pi^*(n)$ represent the number of pairs in $P(n)$ that are not divisible by $p$ or any prime number less than $p$.

Let $W(n)$ represent the fraction of pairs in $P(n)$ that are twin primes.

Let $\Pi_2(n)$ represent the number of twin prime pairs in the set $P(n)$.

Let $\Pi_2^*(n)$ be a close approximation to $\Pi_2(n)$ as defined by the equation 32. As $n$ increases, the value of $\Pi_2^*(n)$ approaches $\Pi_2(n)$. The asterisk (*) is used to denote apporixmation.

# 3    Background

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 [1]. The case where $k = 1$ is the Twin Prime Conjecture. Since its proposition, the de Polignac Conjecture has remained largely unproven until a breakthrough by Chinese mathematician Yitang Zhang in April 2013. Zhang proved that there exists a value N less than 70 million such that there are an infinite number of paired primes separated by N [2]. A year later in 2015, James Maynard [3] has subsequently refined the GPY sieve method [4] to show there is an N less than or equal to 600 such that there are infinitly many primes separated by N.

# 4    Methodology

In this paper, a more straighforward method is used to prove the Twin Prime Conjecture. By pairing odd numbers that differ by 2, then eliminating the pairs that contain a composite number, a function is derived that determines the number of twin primes less than $n$ for large values of $n$. Then by proof by mathematical induction, it is proven that this function increases indefinitely with increasing $n$ thus proving there are an infinite number of twin primes.

To find all the twin primes less than or equal to odd integer $n$, let us first

start with the set of pairs of odd integers and pair them $(x, y)$ such that for each pair $x + 2 = y$ and $y \leq n$. The pair $(1, 3)$ will not be included since 1 is not considered a prime number. For a given odd integer $n$, we see that there are $|P(n)| = (n - 3)/2$ pairs. This give us the following set:

P(n) = {(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) ... (n-4,n-2),(n-2,n)}

Next let us 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 $\lambda(\sqrt{n})$, the largest prime less than or equal to $\sqrt{n}$. There are no prime numbers greater than $\lambda(\sqrt{n})$ 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, excluding the pair (3,5). 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). Note that there are no pairs that have both the $x$ and $y$ coordinate divisible by 3.

P(n) = {(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 $\lfloor(|P(n)| - 1)/3\rfloor$ pairs where the $x$ coordinate is divisible by 3 and $x \neq 3$. We subtract 1 from $P(n)$ since we are not counting pair (3,5). There are $\lfloor|P(n)|/3\rfloor$ pairs where the $y$ coordinate is divisible by 3. Therefore, in total, there are $\lfloor(|P(n)|-1)/3\rfloor + \lfloor|P(n)|/3\rfloor$ pairs where either the $x$ or $y$ coordinates are divisible by 3 but not equal to 3. As $|P(n)|$ gets very large, the value of $|P(n)| - 1$ approaches $|P(n)|$ and the number of pairs divisible by 3 approaches $(2/3)P(n)$. Therefore, the number of pairs $K_3(n)$ that have an $x$ or $y$ coordinate that is divisible by 3 can be represented by the following equation:

$$K_3(n) = \lfloor(|P(n)| - 1)/3\rfloor + \lfloor|P(n)|/3\rfloor \tag{1}$$

$$K_3(n) \lim_{n \to \infty} = (2/3)|P(n)| \tag{2}$$

$$K_3^*(n) = (2/3)|P(n)| \tag{3}$$

So the number of pairs that are not divisible by 3, $\pi_3^*(n)$ , can be repre-

sented by the following equation:

$$\pi_3^*(n) = |P(n)| - K^*(3) \tag{4}$$

$$\pi_3^*(n) = |P(n)| - (2/3)|P(n)| \tag{5}$$

$$\pi_3^*(n) = (1/3)|P(n)| \tag{6}$$

Next, we eliminate the pairs where the $x$ or $y$ coordinate is evenly divisible by 5, excluding the pairs (3,5) and (5,7). 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).

P(n) = {(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 $\lfloor(|P(n)| - 2)/5\rfloor$ pairs where $x$ coordinate is divisible by 5 and $x \neq 5$. We have to subtract 2 from $P(n)$ since we are not counting pairs (3,5) or (5,7). There are $\lfloor(|P(n)| - 1)/5\rfloor$ pairs where $y$ is divisible by 5 and $y \neq 5$. So there are $\lfloor(|P(n)| - 2)/5\rfloor + \lfloor(|P(n)| - 1)/5\rfloor$ pairs where either the $x$ or $y$ coordinates are divisible by 5 but not equal to 5. As $|P(n)|$ gets very large, the values of $|P(n)| - 2$ and $|P(n)| - 1$ approach $|P(n)|$ and the number of pairs divisible by 5 approaches $(2/5)|P(n)|$.

We will use lowercase $k_5(n)$ for only the pairs divisible by 5.

$$k_5(n) = \lfloor(|P(n)| - 2)/5\rfloor + \lfloor(|P(n)| - 1)/5\rfloor \tag{7}$$

$$k_5(n)\lim_{n \to \infty} = (2/5)|P(n)| \tag{8}$$

$$k_5^*(n) = (2/5)|P(n)| \tag{9}$$

Notice however, that every third pair (green) where the $x$ coordinate is divisible by 5, the $x$ cooridnate 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, the $y$ coordinate 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 $k_5(n)$ divisible by 5 but not by 3 approaches the following equation as $n$ gets very large.

4

$$K_5(n) \lim_{n\to\infty} = (1/3)k_5^*(n) \tag{10}$$

$$K_5(n) \lim_{n\to\infty} = (1/3)(2/5)|P(n)| \tag{11}$$

$$K_5^*(n) = (1/3)(2/5)|P(n)| \tag{12}$$

By subtracting $K_5^*$ from $\pi_3^*(n)$ gives us the number of pairs that are not divisible by 3 or 5. This can be represented by the following equation:

$$\pi_5^*(n) = \pi_3^*(n) - K_5^*(n) \tag{13}$$

$$\pi_5^*(n) = (1/3)|P(n)| - (1/3)(2/5)|P(n)| \tag{14}$$

$$\pi_5^*(n) = (1/3)(3/5)|P(n)| \tag{15}$$

Next, we eliminate the pairs where the $x$ or $y$ coordinate is divisible by 7, excluding pairs (5,7) and (7,9). 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` )
. . .

There are $\lfloor(|P(n)| - 3)/7\rfloor$ pairs where $x$ coordinate is divisible by 7 and $x \neq 7$. We have to subtract 3 from $P(n)$ since we are not counting pairs (3,5), (5,7) or (7,9). There are $\lfloor(|P(n)| - 2)/7\rfloor$ pairs where $y$ is divisible by 7 and $y \neq 7$. We will use lowercase $k_7(n)$ for only the pairs divisible by 7.

$$k_7(n) = \lfloor(|P(n)| - 3)/7\rfloor + \lfloor(|P(n)| - 2)/7\rfloor \tag{16}$$

$$k_7(n) \lim_{n\to\infty} = (2/7)|P(n)| \tag{17}$$

$$k_7^*(n) = (2/7)|P(n)| \tag{18}$$

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.

$$K_7(n) \lim_{n \to \infty} = (1/3)(3/5)k_7^*(n) \tag{19}$$

$$K_7(n) \lim_{n \to \infty} = (1/3)(3/5)(2/7)P(n) \tag{20}$$

$$K_7^*(n) = (1/3)(3/5)(2/7)|P(n)| \tag{21}$$

By subtracting $K_7$ from $\pi^5$ give us the number of pairs that are not divisible by 3, 5 or 7. This can be represented by the following equation:

$$\pi_7^*(n) = \pi_5^*(n) - K_7^*(n) \tag{22}$$

$$\pi_7^*(n) = (1/3)(3/5)|P(n)| - (1/3)(3/5)(2/7)|P(n)| \tag{23}$$

$$\pi_7^*(n) = (1/3)(3/5)(5/7)|P(n)| \tag{24}$$

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

$$k_p(n) = \left\lfloor \frac{(|P(n)| - \frac{(p-1)}{2})}{p} \right\rfloor + \left\lfloor \frac{(|P(n)| - \frac{(p-1)}{2} + 1)}{p} \right\rfloor \tag{25}$$

$$k_p(n) \lim_{n \to \infty} = (2/p)|P(n)| \tag{26}$$

$$k_p^*(n) = (2/p)|P(n)| \tag{27}$$

The general formula for number of pairs divisible by prime number $p$ but not divisible by a prime number less than $p$ is as follows

$$K_p(n) \lim_{n \to \infty} = (1/3)(3/5)(5/7)...(\lambda(p) - 2)/\lambda(p)(2/p)P(n)$$
or

$$K_p(n) \lim_{n \to \infty} = P(n)(2/p) \prod_{q=3}^{\lambda(p)} ((q - 2)/q) \tag{28}$$

where the product is over prime numbers only.

The number of pairs that are not divisible by any prime number less than or equal to $p$ can be represented by the following equation:

$$\pi_p^*(n) = (1/3)(3/5)(5/7)...\frac{(p - 2)}{p}|P(n)| \tag{29}$$

For any integer $n$, the number of twin prime pairs can be defined as follows:

$$\Pi_2^*(n) = |P(n)| \prod_{p=3}^{\lambda(\sqrt{n})} \frac{p-2}{p} \tag{30}$$

Let us define the function $W(n)$ to be the fraction of pairs that are twin primes as follows:

$$W(n) = \prod_{p=3}^{\lambda(\sqrt{n})} \frac{p-2}{p} \tag{31}$$

Then the equation for the number of twin primes less than $n$ simplifies to the following:

$$\Pi_2^*(n) = |P(n)|W(n) \tag{32}$$

Since $|P(n)| = (n-3)/2$, the equation in terms of $n$ is the following:

$$\Pi_2^*(n) = \left( \frac{n-3}{2} \right) W(n) \tag{33}$$

To verify that the derivation of the equation for $\Pi_2^*(n)$ was correct and to determine at what point this equation begins to accurately determine the number of twin primes less than $n$, I plotted the actual value of $\pi(n)$ (blue line) and the aproximation $\Pi_2^*(n)$ (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 $\Pi_2^*(n)$ for values of $n < 5,000$. This is not a problem since this errs on the side of caution. But as $n$ increases, equation for $\Pi_2^*(n)$ very closely estimates $\Pi_2(n)$. 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 for $\Pi_2^*(n)$ for large values of $n$.

# 5   The Proof of the Twin Prime Conjecture

To prove the Twin Prime conjecture, that there is an infinite number of twin primes, we will use proof by induction. To do this, we will show that there is at least 1 twin prime between $\Pi_2^*(p_i^2)$ and $\Pi_2^*(p_{i+1}^2)$, and since it has been proven that there is an infinite number of prime numbers, this proves there
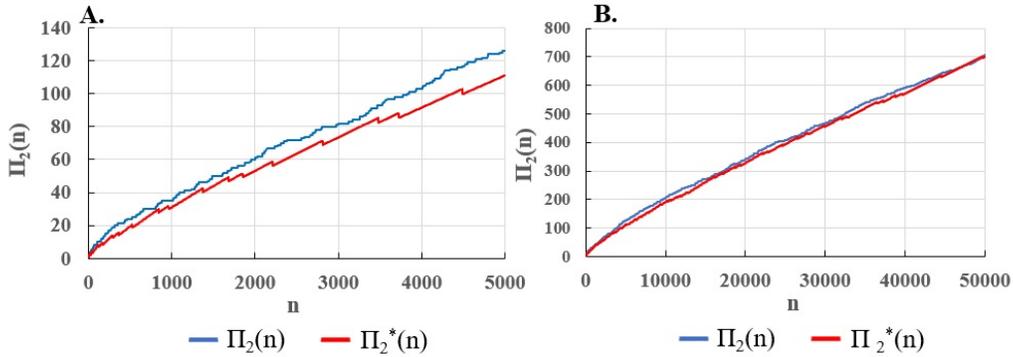
Figure 1: The actual number of twin primes $\Pi_2(n)$ (blue line) is underestimated by $\Pi_2^*(n) = |P(n)|W(n)$ (red line) for values of $n < 5{,}000$. But as $n$ gets larger, the value of $\Pi_2^*(n)$ approaches the $\Pi_2(n)$.

is an infinite number of twin primes. Let $\Delta\Pi_2(p_i)$ be the number of prime numbers between $\Pi_2^*(p_i^2)$ and $\Pi_2^*(p_{i+1}^2)$.

We will examine vaules of $\Pi_2^*(p_i^2)$ rather than values of $\Pi_2^*(p_i)$ because if you look at the values of $W(n)$, the vaule of $W(n) = W(n+1)$ unless $n+1 = p^2$. For example $W(9) = W(10) = W(11) = \ldots = W(24)$. But $W(9)$ is not equal to $W(25)$. Because of this, we will only look at values of $n$ that are squared primes.

To perform proof by induction, we must first get $W(p_{i+1}^2)$ in terms of $W(p_i^2)$. Upon examining $W(n)$ we find the following equation to hold true:

$$W(p_{i+1}^2) = \frac{(p_{i+1} - 2)}{p_{i+1}} W(p_i^2) \tag{34}$$

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:

8

$$\Delta\Pi_2(p_i) = \frac{(p_{i+1}^2 - 3)}{2}W(p_{i+1}^2) - \frac{(p_i^2 - 3)}{2}W(p_i^2) \tag{35}$$

$$\Delta\Pi_2(p_i) = \frac{(p_{i+1}^2 - 3)}{2}\frac{(p_{i+1} - 2)}{p_{i+1}}W(p_i^2) - \frac{(p_i^2 - 3)}{2}W(p_i^2) \tag{36}$$

$$\Delta\Pi_2(p_i) = \frac{W(p_i^2)}{2}\left(\frac{(p_{i+1}^2 - 3)(p_{i+1} - 2)}{p_{i+1}} - (p_i^2 - 3)\right) \tag{37}$$

$$\Delta\Pi_2(p_i) = \frac{W(p_i^2)}{2}\left(\frac{(p_{i+1}^3 - 2p_{i+1}^2 - 3p_{i+1} + 6)}{p_{i+1}} - (p_i^2 - 3)\right) \tag{38}$$

$$\Delta\Pi_2(p_i) = \frac{W(p_i^2)}{2}\left(p_{i+1}^2 - 2p_{i+1} - 3 + \frac{6}{p_{i+1}} - p_i^2 + 3)\right) \tag{39}$$

$$\Delta\Pi_2(p_i) = \frac{W(p_i^2)}{2}\left(p_{i+1}^2 - 2p_{i+1} + \frac{6}{p_{i+1}} - p_i^2)\right) \tag{40}$$

We know that $W(p_i^2)$ is greater than 0 and $p_{i+1}^2 - 2p_{i+1} + \frac{6}{p_{i+1}} - p_i^2$ is greater than 0 so the product must be greater than 0. This proves that $\Delta\Pi_2(n)$ is always increasing and thus the Twin Prime conjecture that there is an infinite number of twin primes holds true. To be more precise, let us examine what is the minimum value of $\Delta\Pi_2(p_i)$.

We can define $p_{i+1} = p_i + x$ where $x$ is a postive even integer. So substituting $p_i + x$ for $p_{i+1}$ into the term $p_{i+1}^2 - 2p_{i+1} + \frac{6}{p_{i+1}} - p_i^2$ gives us the following:

$$p_{i+1}^2 - 2p_{i+1} + \frac{6}{p_{i+1}} - p_i^2 = (p_i + x)^2 - 2(p_i + x) + \frac{6}{p_i + x} - p_i^2 \tag{41}$$

$$= (p_i^2 + 2xp_i + x^2) - 2p_i - 2x + \frac{6}{p_i + x} - p_i^2 \tag{42}$$

$$= p_i(2x - 2) + x^2 - 2x + \frac{6}{p_i + x} \tag{43}$$

Substituting this into equation 40 give us the following:

$$\Delta\Pi_2(p_i) = \frac{W(p_i^2)}{2} \times \left(p_i(2x - 2) + x^2 - 2x + \frac{6}{p_i + x}\right) \tag{44}$$

So the next question is, what is the lower bound on $\Delta\Pi_2(p_i)$. The case where $\Delta\Pi_2(p_i)$ is minimal is when $x = 2$. In other words, when $p_{i+1} = p_i + 2$.

9

Substituting 2 for $x$ into equation 44 gives us the new equation for the lower bound for $\Delta\Pi_2(p_i)$.

$$\Delta\Pi_2(p_i) \geq \frac{W(p_i^2)}{2} \times \left(p_i(4-2) + 2^2 - 4 + \frac{6}{p_i+2}\right) \tag{45}$$

$$\Delta\Pi_2(p_i) \geq \frac{W(p_i^2)}{2} \times \left(2p_i + \frac{6}{p_i+2}\right) \tag{46}$$

$$\Delta\Pi_2(p_i) \geq p_i(W(p_i^2)) \tag{47}$$

$$\Delta\Pi_2^*(p_i) = p_i(W(p_i^2)) \tag{48}$$

where $\Delta\Pi_2^*(p_i)$ represents the lower bound on $\Delta\Pi_2(p_i)$.

A graph of $\Delta\Pi_2(p_i)$ versus $p$ (blue line) and $\Delta\Pi_2^*(p_i)$ versus $p$ (orange line) in Figure 2 confirms that $\Delta\Pi_2^*(p_i)$ is always less than or equal to $\Delta\Pi_2(p_i)$. Notice that $\Delta\Pi_2^*(p_i)$ coincides with $\Delta\Pi_2(p_i)$ only at the points where $p_{i+1} = p_i + 2$.

Now that we know that $\Delta\Pi_2^*(p_i)$ is always less than or equal to $\Delta\Pi_2(p_i)$, if we show that $\Delta\Pi_2^*(p_i)$ is always greater than or equal to 1, then we know that $\Delta\Pi_2(p_i)$ will always be greater than 1. We can prove this by mathematical induction.

Base case for $\Delta\Pi_2^*(p_0)$:
Using $p_0 = 3$, we get the following

$$\Delta\Pi_2^*(p_0) = 3W(3^2) = 3(1/3) = 1 \tag{49}$$

Inductive step:
Next, we assume that $\Delta\Pi_2^*(p_i) \geq 1$, and prove that $\Delta\Pi_2^*(p_{i+1}) \geq 1$
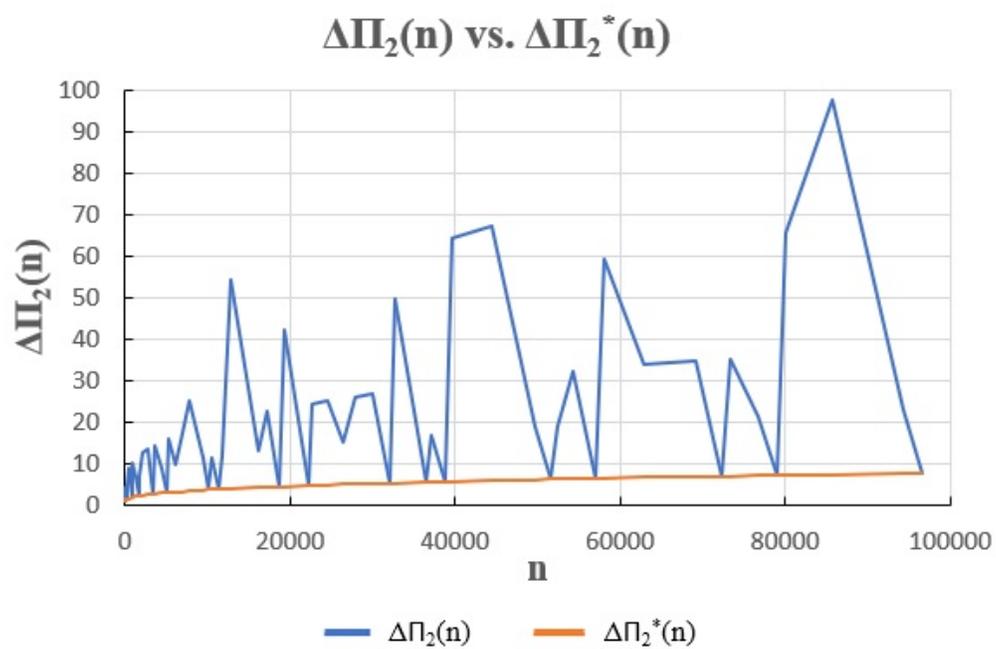Substituting $p_{i+1}$ into $\Delta\Pi_2^*(p_i) = p_i(1 - 2W(p_i^2)) \geq 1$, and using equation 34 gives:

Figure 2: Graph of $\Delta\Pi_2(n)$ and the lower bound $\Delta\Pi_2^*(n)$ versus $n$ where $n = p_i^2$. $\Delta\Pi_2^*(n)$ is always less than or equal to $\Delta\Pi_2(n)$ and they coincide only at the points where $n = p_i^2$ and $p_{i+1} = p_i + 2$.

$$\Delta\Pi_2^*(p_{i+1}) = p_{i+1}(W(p_{i+1}^2)) \tag{50}$$

$$\Delta\Pi_2^*(p_{i+1}) = p_{i+1}\frac{(p_{i+1}-2)}{p_{i+1}}W(p_i^2) \tag{51}$$

$$\Delta\Pi_2^*(p_{i+1}) = (p_{i+1}-2)W(p_i^2) \tag{52}$$

$$\Delta\Pi_2^*(p_{i+1}) = \frac{(p_{i+1}-2)}{p_i}p_iW(p_i^2) \tag{53}$$

$$\Delta\Pi_2^*(p_{i+1}) = \frac{(p_{i+1}-2)}{p_i}\Delta\Pi_2^*(p_i) \tag{54}$$

Since we assumed that $\Delta\Pi_2^*(p_i)$ is greater than or equal to 1, and $\frac{(p_{i+1}-2)}{p_i}$ is greater than or equal to 1, then the product must be greater than or equal to 1. Therefore $\Delta\Pi_2^*(p_{i+1})$ must be greater than or equal to 1 for all $p_i$ and $p_{i+1}$. Therefore, the number of twin primes always increases by at least 1 between $n = p_i^2$ and $n = p_{i+1}^2$, and since it has been proven by Euclid that there are an infinite number of prime numbers $p_i$, there are an infinite number of twin primes.

Since $p_{i+1}$ is at least $p_i + 2$, this proves 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 integer $n$, there is at least 1 twin prime pair between $n^2$ and $(n + 2)^2$. Let this be known as the Watanabe Conjecture.

# 6    Error Analysis

We have shown that $\Pi_2^*(n) = \frac{(n-3)}{2}W(n)$ is an estimate of the number of twin primes less than $n$ and as $n$ increases, this approximation approaches the actual number of twin primes less than $n$ or $\Pi_2(n)$. Using this approximation, we have proven that as $n$ increases indefinitely, so does the value of $\Pi_2^*(n)$, thus proving the Twin Prime conjecture as true. However, since $\Pi_2^*(n)$ is an approximation of $\Pi_2(n)$, the difference between $\Pi_2^*(n)$ and $\Pi_2(n)$ (i.e. the error) may increase as $n$ increases to the point where any increase in $\Pi_2^*(n)$ can be offset by the error.

We will first calculate the error in $\pi_3^*(n) = (1/3)|P(n)$, the number of pairs not divisible by 3. From equation 1, the exact number of pairs of $n$

that are divisible by 3 is defined by the following equation:

$$k_3(n) = \lfloor (|P(n)| - 1)/3 \rfloor + \lfloor |P(n)|/3 \rfloor \tag{55}$$

Therefore the exact number of pairs of $n$ that are not divisible by 3 , $\bar{k}_3(n)$, is defined by the following equation:

$$\bar{k}_3(n) = |P(n)| - k_3(n) \tag{56}$$

As $n$ get large, $k_3(n)$ was approximated by the following equation:

$$k_3^*(n) = (2/3)|P(n)| \tag{57}$$

Therefore, the approximation of the number of pairs of $n$ that are not divisible by 3, $\bar{k}_3^*(n)$, is defined by the following equation:

$$\bar{k}_3^*(n) = |P(n)| - k_3^*(n) \tag{58}$$

The difference between $\bar{k}_3(n)$ and $\bar{k}_3^*(n)$ is the error in $\pi_3^*(n)$, $e_3(n)$, and can be defined as follows:

$$e_3(n) = \bar{k}_3(n) - \bar{k}_3^*(n) \tag{59}$$
$$e_3(n) = |P(n)| - k_3(n) - |P(n)| + k_3^*(n) \tag{60}$$
$$e_3(n) = k_3^*(n) - k_3(n) \tag{61}$$
$$e_3(n) = (2/3)|P(n)| - \lfloor (|P(n)| - 1)/3 \rfloor - \lfloor |P(n)|/3 \rfloor \tag{62}$$

Note that this is the same as the difference between $k_3(n)$ and $k_3^*(n)$.

The maximum values of $e_3(n)$ occurs when $|P(n)| = 2 + 3i$ where $i$ is an integer greater than 0. This is because $|P(n)| = 2$ is the largest integer where the floor functions of $e_3(n)$ will round down the fractions to 0. Plugging values of $|P(n)|$ into equation 62 confirms that the maximum value of $e_3(n)$ is 1 1/3 as shown in figure 3A.

$$e_3 = 1\frac{1}{3} \tag{63}$$

Next, in order to calculate the maximum error in $\pi_5^*(n)$, we must first calculate the maximum error in $k_5^*(n)$. From equation 7, the exact number of pairs of $n$ that are divisible by 5 is defined by the following equation:

$$k_5(n) = \lfloor (|P(n)| - 2)/5 \rfloor + \lfloor (|P(n)| - 1)/5 \rfloor \tag{64}$$

13

**A.**

| P(n) | x | y | div 3 | (2/3)P(n) | diff |
|---|---|---|---|---|---|
| 1 | 3 | 5 | 0 | 0.666667 | 0.666667 |
| 2 | 5 | 7 | 0 | 1.333333 | 1.333333 |
| 3 | 7 | 9 | 1 | 2 | 1 |
| 4 | 9 | 11 | 2 | 2.666667 | 0.666667 |
| 5 | 11 | 13 | 2 | 3.333333 | 1.333333 |
| 6 | 13 | 15 | 3 | 4 | 1 |
| 7 | 15 | 17 | 4 | 4.666667 | 0.666667 |
| 8 | 17 | 19 | 4 | 5.333333 | 1.333333 |
| 9 | 19 | 21 | 5 | 6 | 1 |
| 10 | 21 | 23 | 6 | 6.666667 | 0.666667 |
| 11 | 23 | 25 | 6 | 7.333333 | 1.333333 |
| 12 | 25 | 27 | 7 | 8 | 1 |
| 13 | 27 | 29 | 8 | 8.666667 | 0.666667 |
| 14 | 29 | 31 | 8 | 9.333333 | 1.333333 |
| 15 | 31 | 33 | 9 | 10 | 1 |
| 16 | 33 | 35 | 10 | 10.66667 | 0.666667 |
| 17 | 35 | 37 | 10 | 11.33333 | 1.333333 |
| 18 | 37 | 39 | 11 | 12 | 1 |
| 19 | 39 | 41 | 12 | 12.66667 | 0.666667 |
| 20 | 41 | 43 | 12 | 13.33333 | 1.333333 |
| 21 | 43 | 45 | 13 | 14 | 1 |
| 22 | 45 | 47 | 14 | 14.66667 | 0.666667 |
| 23 | 47 | 49 | 14 | 15.33333 | 1.333333 |
| 24 | 49 | 51 | 15 | 16 | 1 |

**B.**

| P(n) | x | y | div 5 | (2/5)P(n) | diff |
|---|---|---|---|---|---|
| 1 | 3 | 5 | 0 | 0.4 | 0.4 |
| 2 | 5 | 7 | 0 | 0.8 | 0.8 |
| 3 | 7 | 9 | 0 | 1.2 | 1.2 |
| 4 | 9 | 11 | 0 | 1.6 | 1.6 |
| 5 | 11 | 13 | 0 | 2 | 2 |
| 6 | 13 | 15 | 1 | 2.4 | 1.4 |
| 7 | 15 | 17 | 2 | 2.8 | 0.8 |
| 8 | 17 | 19 | 2 | 3.2 | 1.2 |
| 9 | 19 | 21 | 2 | 3.6 | 1.6 |
| 10 | 21 | 23 | 2 | 4 | 2 |
| 11 | 23 | 25 | 3 | 4.4 | 1.4 |
| 12 | 25 | 27 | 4 | 4.8 | 0.8 |
| 13 | 27 | 29 | 4 | 5.2 | 1.2 |
| 14 | 29 | 31 | 4 | 5.6 | 1.6 |
| 15 | 31 | 33 | 4 | 6 | 2 |
| 16 | 33 | 35 | 5 | 6.4 | 1.4 |
| 17 | 35 | 37 | 6 | 6.8 | 0.8 |
| 18 | 37 | 39 | 6 | 7.2 | 1.2 |
| 19 | 39 | 41 | 6 | 7.6 | 1.6 |
| 20 | 41 | 43 | 6 | 8 | 2 |
| 21 | 43 | 45 | 7 | 8.4 | 1.4 |
| 22 | 45 | 47 | 8 | 8.8 | 0.8 |
| 23 | 47 | 49 | 8 | 9.2 | 1.2 |
| 24 | 49 | 51 | 8 | 9.6 | 1.6 |

**C.**

| P(n) | x | y | div 7 | (2/7)P(n) | diff |
|---|---|---|---|---|---|
| 1 | 3 | 5 | 0 | 0.285714 | 0.285714 |
| 2 | 5 | 7 | 0 | 0.571429 | 0.571429 |
| 3 | 7 | 9 | 0 | 0.857143 | 0.857143 |
| 4 | 9 | 11 | 0 | 1.142857 | 1.142857 |
| 5 | 11 | 13 | 0 | 1.428571 | 1.428571 |
| 6 | 13 | 15 | 0 | 1.714286 | 1.714286 |
| 7 | 15 | 17 | 0 | 2 | 2 |
| 8 | 17 | 19 | 0 | 2.285714 | 2.285714 |
| 9 | 19 | 21 | 1 | 2.571429 | 1.571429 |
| 10 | 21 | 23 | 2 | 2.857143 | 0.857143 |
| 11 | 23 | 25 | 2 | 3.142857 | 1.142857 |
| 12 | 25 | 27 | 2 | 3.428571 | 1.428571 |
| 13 | 27 | 29 | 2 | 3.714286 | 1.714286 |
| 14 | 29 | 31 | 2 | 4 | 2 |
| 15 | 31 | 33 | 2 | 4.285714 | 2.285714 |
| 16 | 33 | 35 | 3 | 4.571429 | 1.571429 |
| 17 | 35 | 37 | 4 | 4.857143 | 0.857143 |
| 18 | 37 | 39 | 4 | 5.142857 | 1.142857 |
| 19 | 39 | 41 | 4 | 5.428571 | 1.428571 |
| 20 | 41 | 43 | 4 | 5.714286 | 1.714286 |
| 21 | 43 | 45 | 4 | 6 | 2 |
| 22 | 45 | 47 | 4 | 6.285714 | 2.285714 |
| 23 | 47 | 49 | 5 | 6.571429 | 1.571429 |
| 24 | 49 | 51 | 6 | 6.857143 | 0.857143 |

Figure 3: The difference between the actual number of pairs divisible by 3, 5 and 7 and $(2/3)|P(n)|$, $(2/5)|P(n)|$ and $(2/7)|P(n)|$ respectively. The maximum errors are highlighted in yellow.

As $n$ get large, $k_5(n)$ was approximated by the following equation:

$$k_5^*(n) = (2/5)|P(n)| \qquad (65)$$

The difference between $\bar{k}_5(n)$ and $\bar{k}_5^*(n)$ is the error, $e_5(n)$ but since this is the same as the difference between $k_5(n)$ and $k_5^*(n)$, $e_5(n)$ can be defined as follows:

$$e_5(n) = k_5^*(n) - k_5(n) \qquad (66)$$
$$e_5(n) = (2/5)|P(n)| - \lfloor(|P(n)|-2)/5\rfloor - \lfloor(|P(n)|-1)/5\rfloor \qquad (67)$$

The maximum values of $e_5(n)$ occurs when $|P(n)| = 5i$ where $i$ is an integer greater than 0. This is because $|P(n)| = 5$ is the largest integer where both floor functions of $e_5(n)$ will round down the fractions to 0. Plugging values of $|P(n)|$ into equation 67 confirms that the maximum value of $e_5(n)$ is 2 as shown in figure 3B.

$$e_5 = 2 \qquad (68)$$

Next we need ot determine the error in $\pi_5^*(n)$, the number of pairs not divisible by 5 or 3. To do this, let us define the following equations:

$a(n)$ = number of pairs divisible by 3 excluding (3,n-3) and (n-3,3).
$b(n)$ = number of pairs divisible by 5 excluding (5,n-5) and (n-5,5).
$c(n)$ = number of pairs divisible by both 3 and 5

In this case $a(n) = k_3(n)$ and $b(n) = k_5(n)$. The function $c(n)$ would include all pairs divisible by 15. The function c(n) can be defined as follows:

$$c(n) = \lfloor(P(n)+9)/15\rfloor + \lfloor(P(n)+8)/15\rfloor \qquad (69)$$

The inclusion/exclusion principle states that given two set A and B, the number of elements in A union B is $|A \cup B| = |A| + |B| - |A \cap B|$. Also note that $|A \cup B| \leq |A| + |B|$. Using the inclusion/exclusion principle, we can define number of pairs of $n$ that are evenly divisible by 3 or 5, $K_5(n)$, as follows:

$$K_5(n) = a(n) + b(n) - c(n) \qquad (70)$$

15

Subtracting $K_5(n)$ from $|P(n)|$ gives us $\pi_5(n)$:

$$\pi_5(n) = |P(n)| - K_5(n) \tag{71}$$
$$\pi_5(n) = |P(n)| - a(n) - b(n) + c(n) \tag{72}$$
$$\pi_5(n) = |P(n)| - k_3(n) - k_5(n) + c(n) \tag{73}$$
$$\begin{aligned}\pi_5(n) = |P(n)| &- (\lfloor(|P(n)| - 1)/3\rfloor + \lfloor|P(n)|/3\rfloor) \\ &- (\lfloor(|P(n)| - 2)/5\rfloor + \lfloor(|P(n)| - 1)/5\rfloor) \\ &+ (\lfloor(P(n) + 9)/15\rfloor + \lfloor(P(n) + 8)/15\rfloor)\end{aligned} \tag{74}$$

We also know from equation 15, that

$$\pi_5^*(n) = (1/3)(1/5)|P(n)| \tag{75}$$
$$\pi_5^*(n) = P(n) - (2/3)P(n) - (2/5)P(n) + (4/15)P(n) \tag{76}$$

Now we can calculate the error in $\Pi_5^*(n)$. Let us call that error $e_{max5}(n)$.

$$e_{max5}(n) = \pi_5(n) - \pi_5^*(n) \tag{77}$$
$$\begin{aligned}e_{max5}(n) = P(n) &- a(n) - b(n) + c(n) \\ &- (P(n) - (2/3)P(n) - (2/5)P(n) + (4/15)P(n))\end{aligned} \tag{78}$$
$$e_{max5}(n) = (2/3)P(n) - a(n) + (2/5)P(n) - b(n) - ((4/15)P(n) - c(n)) \tag{79}$$

Since $e_3(n)$ and $e_5(n)$ were derived by equations 62 and 67, we can substitute $e_3(n)$ and $e_5(n)$ into equation 79 to get:

$$e_{max5}(n) = e3 + e5 - ((4/15)P(n) - c(n)) \tag{80}$$

Since $|A \cup B| \leq |A| + |B|$, we can simplify the equation as follows:

$$e_{max5}(n) \leq e3 + e5 \tag{81}$$

The general formula for error $e_p$ is as follows:

$$e_p(n) = k_p^*(n) - k_p(n) \tag{82}$$

Substituting for $k_p(n)$ and $k_p^*(n)$ using equation 25 gives us the following:

$$e_p(n) = \left(\frac{2}{p}\right)|P(n)| - \left\lfloor\frac{(|P(n)| - \frac{(p-1)}{2})}{p}\right\rfloor + \left\lfloor\frac{(|P(n)| - \frac{(p-1)}{2} + 1)}{p}\right\rfloor \tag{83}$$

16

The maximum error for any $p$ occurs when the numerator of the second floor term $|P(n)| - \frac{(p-1)}{2} + 1$ equals $p - 1$. This is the maximum value of $|P(n)|$ where both floor functions will round down to 0.

$$|P(n)|_{max} - \frac{(p-1)}{2} + 1 = p - 1 \tag{84}$$

$$|P(n)|_{max} = p - 1 + \frac{(p-1)}{2} - 1 \tag{85}$$

$$|P(n)|_{max} = \frac{3p - 5}{2} \tag{86}$$

Plugging the value of $|P(n)|_{max}$ into equation 83 will give us the maximum error contribution for $p$.

$$e_p = \left(\frac{2}{p}\right)\frac{3p - 5}{2} - 0 - 0 \tag{87}$$

$$e_p = \frac{3p - 5}{p} \tag{88}$$

So the maximum error contribution for any prime number $p$ can be calculated from equation 88. Below are a few values:

$$e_7 = 16/7 = 2.285714 \tag{89}$$
$$e_{11} = 28/11 = 2.545455 \tag{90}$$
$$e_{13} = 34/13 = 2.615385 \tag{91}$$
$$e_{17} = 46/17 = 2.705882 \tag{92}$$
$$etc.$$

As we add errors for $e_7$, $e_{11}$, $e_{13}$, etc., the inclusion/exclusion principle starts getting exponentially complex since we have to subtract overlapping values and add back in other overlapping values. Since $|A \cup B| \leq |A| + |B|$, we can simplify the equation by adding the individual errors without subtracting the overlap since this errs on the side of caution. If we sum up all these errors,

we get the maximum error for equation 32.

$$e_{max}(n) = \sum_{p=3}^{\lambda(\sqrt{n})} e_p \tag{93}$$

$$e_{max}(n) = \sum_{p=3}^{\lambda(\sqrt{n})} \frac{3p-5}{p} \tag{94}$$

$$e_{max}(n) < \sum_{p=3}^{\lambda(\sqrt{n})} 3 \tag{95}$$

$$e_{max}^*(n) = 3\pi(\sqrt{n}) \tag{96}$$

where $\pi(x)$ is the prime counting function.

To confirm no errors were made in the error calculation, a graph of $\Pi_2^*(n) \pm e_{max}^*(n)$ vs. $n$ was compared to $\Pi_2(n)$. As can be seen in figure 4, $\Pi_2(n)$ lies comfortably within $\Pi_2^*(n) \pm e_{max}^*(n)$. Since $e_{max}^*(n)$ increases at a much slower rate than $\Pi_2^*(n)$, this proves that there are an infinite number of twin primes.

# 7 Proof of de 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), ... (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 32.

$$\pi_4(n) = |P(n)|W(n)$$

where $\pi_4(n)$ is the number of quad primes less than $n$.

In fact, for all values of $k = 2^i$, it can be shown that the number of primes separted by $2^i$ is the same as the number of twin primes for very large values
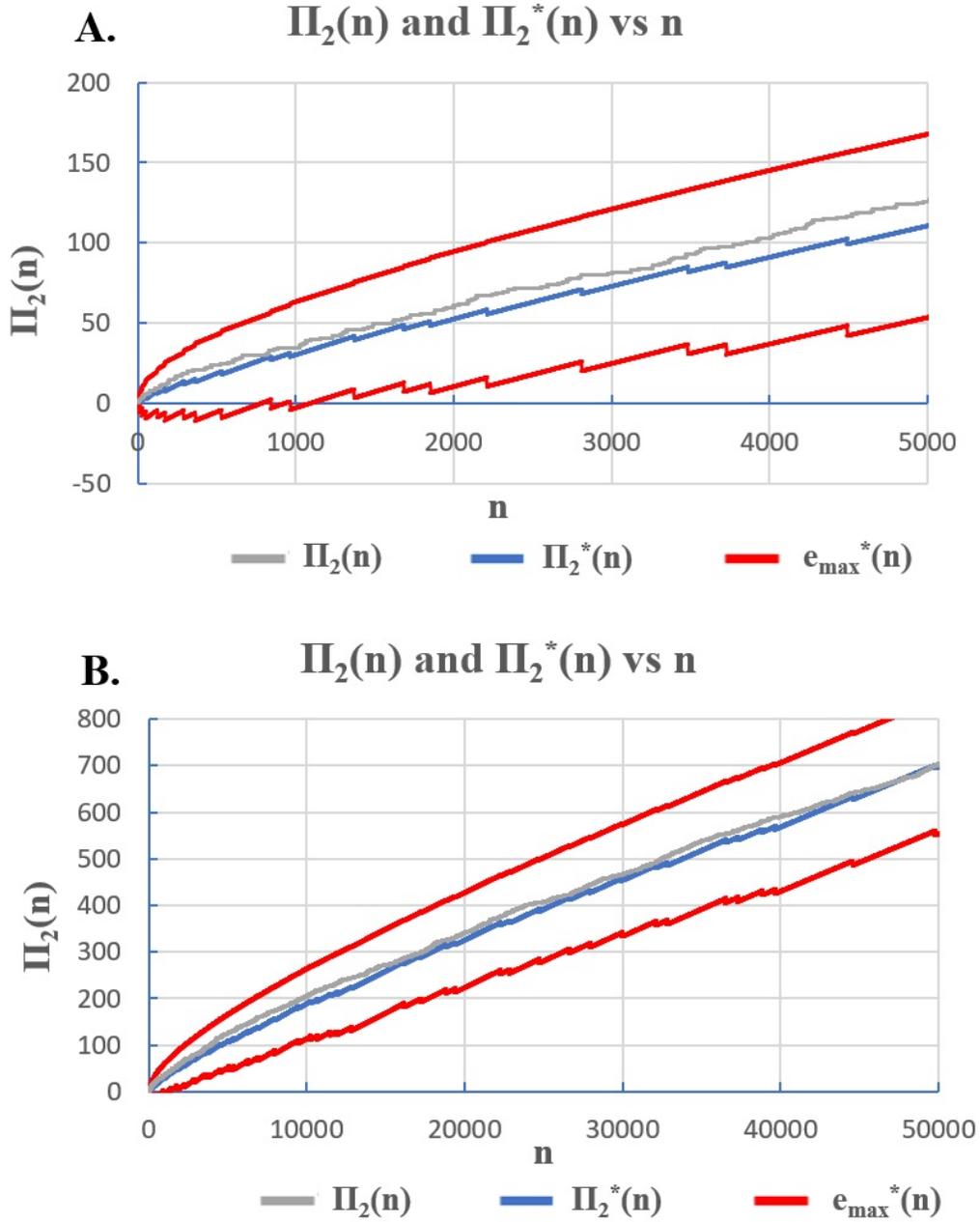
Figure 4: Graph of $\Pi_2^*(n)$ blue line versus $\Pi_2(n)$ gray line. The red lines represent $\Pi_2^*(n) \pm e_{max}(n)$. Notice that $\Pi_2(n)$ lies comfortably between $\Pi_2^*(n) + e_{max}(n)$ and $\Pi_2^*(n) - e_{max}(n)$.

19

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 sexy primes, we get the following set:
(3, <mark>9</mark> ), (5,11), (7,13), ( <mark>9</mark> , <mark>15</mark> ), (11,17), (13,19),( <mark>15</mark> , <mark>21</mark> ) ... (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 $W(n)$ function changes from 1/3 to 2/3. This results in a larger number of sext primes relative to number of twin primes. A similar situtation 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 $W(n)$ function will change from (3/5) to (4/5). Since the number of sext 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, sexy primes, dec primes and 30-primes in Figure 5. Notice that the curve for the twin primes has relatively the fewest number of prime pairs.

# 8    Summary

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

$$\Pi_2^*(n) = |P(n)|W(n)$$

where $W(n)$ is defined as

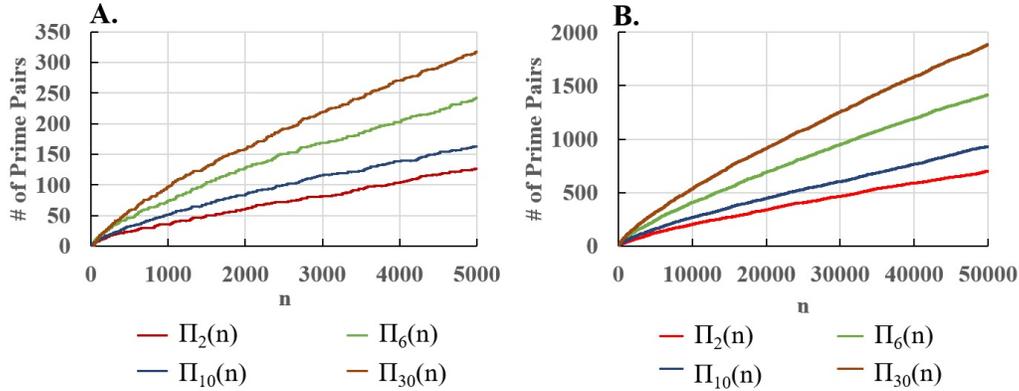$$W(n) = \prod_{(p=3)}^{(\lambda(\sqrt{n}))}((p-2)/p)$$

20

Figure 5: The more factors there are between primes, the more prime pairs exist. There are fewer twin primes (red line) than sexy primes (green line), dec primes (blue line) and 30-primes (brown line).

where $\lambda(\sqrt{n})$ is the largest prime number less than or equal to $\sqrt{n}$ and the product is over prime numbers.

I have shown by proof by induction, that the above equation for number of twin primes increase indefinitely as $n$ increases the proving the Twin Prime Conjecture. I also have shown that the deviation between $\Pi_2^*(n)$ and $\Pi_2(n)$, the actual number of twin primes less than $n$ increases at a slower rate than $\Pi_2^*(n)$, thus proving the Twin Prime Conjecture.

# 9   Future Directions

Future work will involve applying this technique of pairing numbers to prove the Goldbach Conjecture [5]. The Goldbach Conjecture states that every even integer greater than 2 can be expressed as the sum of two primes. To prove the Goldbach Conjecture, we first pair odd numbers $(x, y)$ such that $x + y = n$. For example, (3,n-3),(5,n-5),(7,n-7),(9,n-9)...,(n-5,5),(n-3,3). Then by eliminating pairs that are divisible by 3, 5, 7, 11 etc, the remaining pairs are the prime pairs that sum up to $n$.

I will show that for the set of even integers $n$ not divisible by a prime number less than $sqrtn$, the number of prime pairs that sum to $n$ will approach the following equation as $n$ gets large:

$$\pi(n) = |P(n)|W(n)$$

21

where $\pi(n)$ is the number of prime pairs that add up to $n$ and $W(n)$ is the fraction of prime pairs.

This equation is identical to equation 32. What this means is, that for large values of $n$ not divisible by a prime less than *sqrtn*, the number of prime pairs that sum to $n$ will approach the number ot twin primes less than $n$. Thus, the proof of the Goldbach's Conjecture is reduced to the proof of the Twin Prime Conjecture.

For other cases of the Goldbach Conjecture for $n = 6p, n = 10p$ or $n = 30p$ will reduce to case of Polignac's Conjecture for primes separated by 6, 10 or 30.

Applying this technique to other prime number conjectures will lead to further proofs.

# References

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

[2] Yitang Zhang. Bounded gaps between primes. *Annals of Mathematics*, 179(3):1121–1174, 2014.

[3] James Maynard. Small gaps between primes. *Annals of Mathematics*, 181(1):383–413, 2015.

[4] Cem Y. Yildirim Daniel A. Goldston, Janos Pintz. Primes in tuples. *Annals of Mathematics*, 170(2):819–862, 2009.

[5] Christian Goldbach. *Letter to Euler, Correspondance mathématique et physique de quelques célèbres géomètres du XVIIIème siècle (band 1).* St. Petersbourg, 1843.

# 10   Copyright Notice

written permission by the author (Kenneth A. Watanabe) is strictly prohibited.