Unified Finite Fourier Series of Distribution of Prime, Twin Prime, Goldbach Sum and Prime Pairs of Distance $2n$, and their Chebyshev Polynomial Series

Fred Yang
San Jose, California 95133 USA
March 30, 2016

Contents

1 Introduction 2

2 Used Symbol Presentations and Identities 3
  2.1 Used symbol presentations 3
  2.2 Used identities and equations 3

3 Basic Fourier Series $f(p, x)$ 4

4 Finite Fourier Series of Prime, Twin Prime, Goldbach Sum Distribution 7
  4.1 Fourier Series of Co-Prime and Prime Distribution $F(P_i, x)$ 7
  4.2 Normalization of $F(P, x)$ 9
  4.3 Fourier Series of Twin Prime Distribution $H(P_i, x)$ 10
  4.4 Fourier Series of Goldbach Sum Distribution $G(n, x)$ 15
  4.5 Unified Fourier Series $P(p_i, n, x)$ of Distribution of Prime, Twin Prime, Goldbach Sum and Prime Pair of Distance $2n$ 21

5 Transform to Finite Chebyshev Polynomial Series 21
  5.1 Finite Chebyshev Series of $f(p, x)$ 21
  5.2 Finite Chebyshev Polynomial Series of $F(P_i, x)$ 22
  5.3 Finite Chebyshev Polynomial Series of $H(P_i, x)$ 24
  5.4 Finite Chebyshev Polynomial Series of $G(n, x)$ 25
  5.5 Finite Chebyshev Polynomial Series of $P(p_i, n, x)$ 26
  5.6 Map the Properties of Fourier Series to Chebyshev Series 27

6 Further research 27
Abstract

This paper introduces an unified finite Fourier series which zeroes show distribution of relative prime, twin prime, Goldbach sums of $2n$, and prime pairs of distance $2n$. It can visualize the distribution patterns of these objects, and may also provide a new method to analyze and solve prime problems using Fourier series and polynomial theory, and shows the inter-connection between these prime problems.

I also observed that the gap of the zeros of both of the twin prime and Goldbach prime series is less than $\frac{p^2}{2}$. If this bound of gap could be proven, it can lead to prove Twin Prime Conjecture and Goldbach Conjecture all together.

Another observed conjecture is that the density of twin primes over twin co-primes $\rho_2$ is approximately square of density of primes over co-primes $\rho$, i.e. $\rho_2 \approx \rho^2$. The formulas are derived from this conjecture to estimate number of twin primes and Goldbach sums.

For the conjectural results, extensive computer experiment has been done to validate them and data is provided herein for reference.

1 Introduction

Let $p_i$ be the $i^{th}$ prime, define

$$P(p_i, n, x) = \sum_{p \leq p_i} \left(\frac{p}{2} + \sum_{k=1}^{p-1} (p - k) \cos \frac{2kn\pi}{p} \cos \frac{2k\pi}{p} x\right)$$

which zeros show

- When $n = 0$: prime distribution
- When $n = 1$: twin prime distribution as $(x - 1, x + 1)$
- When $n > 1$ and $0 \leq x < n$: distribution of Goldbach sums as $(n - x, n + x)$
- When $n \geq 1$ and $x > n$: distribution of prime pairs of distance of $2n$ as $(x - n, x + n)$

Additionally, for each integer $x$, the normalized form could show the number of primes $p \leq p_i$ that $x \equiv n \pmod{p}$. When $n = 0$, it means the number of prime divisors of $x$ that $\leq p_i$.

$$P_*(p_i, n, x) = \sum_{p \leq p_i} \frac{c_p}{p} \left(1 + 2 \sum_{k=1}^{p-1} (1 - k \frac{1}{p}) \cos \frac{2kn\pi}{p} \cos \frac{2k\pi}{p} x\right), c_p = \begin{cases} 1, & \text{when } p \mid 2n \\ 2, & \text{when } p \nmid 2n \end{cases}$$

Another useful conjecture observed is that in the range $(0, p_i \#)$, the density of twin primes over twin co-primes $\rho_2$ is approximately square of density of primes over co-primes $\rho$, i.e. $\rho_2 \approx \rho^2$, and the same is true for the density of Goldbach sums and prime pairs of distance $2n$ when $n = 2^m$. This demonstrates the similarity between these problems. From this conjecture, the formulas are derived to estimate number of twin primes, Goldbach sums and prime pairs of distance $2n$, and the computed data is provided to validate that these estimates are good and error approaching to 0 when $N$ getting bigger.
It also explained why the even numbers in the form of \( n = 2^m \) has least density of Goldbach sums and how density increases when \( n \) is dividable by more primes.

Finally, these Fourier series are transformed to Chevyshev polynomial series, so that the polynomial theories could be applied to analyze prime number distribution problems.

## 2 Used Symbol Presentations and Identities

### 2.1 Used symbol presentations

\( \mathbb{P} \): Prime numbers  
\( n \): represents any nature number \( n \in \mathbb{N} \)  
\( p \): represents any prime number.

### 2.2 Used identities and equations

Here are some identities that I used in the proof.

\[
\sum_{k=1}^{p-1} \cos \left( \frac{2k\pi}{p} \right) = -1 \quad (1)
\]

\[
\frac{1}{2}(p-1) \sum_{k=1}^{\frac{1}{2}(p-1)} \cos \left( \frac{2k\pi}{p} \right) = -\frac{1}{2} \quad (2)
\]

\[
\left( \sum_{i=1}^{n} a_i \right)^2 = \sum_{i=1}^{n} a_i^2 + 2 \sum_{i<j} a_i a_j \quad (3)
\]

**Proof of (1).** Consider \( p^{th} \) root of unity

\[
r_k = e^{i \frac{2k\pi}{p}} = \cos \frac{2k\pi}{p} + i \sin \frac{2k\pi}{p}, k = 0, 1, \ldots, p-1,
\]

because sum of the root of unity is 0

\[
\sum_{k=0}^{p-1} \cos \left( \frac{2k\pi}{p} \right) = Re \left( \sum_{k=0}^{p-1} r_k \right) = 0 \implies \\
\sum_{k=1}^{p-1} \cos \left( \frac{2k\pi}{p} \right) = -\cos 0 = -1
\]

**Proof of (2).** Because \( \cos \left( \frac{2k\pi}{p} \right) = \cos \left( \frac{2(p-k)\pi}{p} \right) \),

\[
\frac{1}{2}(p-1) \sum_{k=1}^{\frac{1}{2}(p-1)} \cos \left( \frac{2k\pi}{p} \right) = \frac{1}{2} \sum_{k=1}^{p-1} \cos \left( \frac{2k\pi}{p} \right) = -\frac{1}{2}
\]
3 Basic Fourier Series $f(p, x)$

First, let’s create the basic Fourier series that has zeros on the integers not dividable by given prime $p$, and positive otherwise. These basic series are building blocks to construct the more complicated Fourier series to be discussed later.

Define function $f(p, x)$ as following

$$f(p, x) = \frac{p}{2} + \sum_{k=1}^{p-1} (p - k) \cos\left(\frac{2k\pi}{p} x\right)$$

Function $f(p, x)$ has the following properties

$$f(p, x) \begin{cases} = 0 & \text{when } x \text{ is integer and } p \nmid x \\ = \frac{p^2}{4} & \text{when } x \text{ is integer and } p \mid x \\ > 0 & \text{otherwise.} \end{cases}$$

I show some examples first followed by the proof.

Figure 1: $f(2, x)$

Figure 2: $f(3, x)$
Figure 3: $f(5, x)$

**Proof.** These properties are very obvious for $f(2, x) = 1 + \cos(\pi x)$. For odd prime $p$, let’s first prove the following identity

$$f(p, x) = 2 \left( \sum_{k=1}^{\frac{1}{2}(p-1)} \cos \frac{2k\pi}{p} x + \frac{1}{2} \right)^2$$

(4)

Proof. Let $q = \frac{1}{2}(p-1)$ to simply writing. By applying the identity (3),

$$2 \left( \sum_{k=1}^{q} \cos \frac{2k\pi}{p} x + \frac{1}{2} \right)^2$$

$$= 2 \sum_{k=1}^{q} \left( \cos \frac{2k\pi}{p} x \right)^2 + \frac{1}{2} + 2 \sum_{k=1}^{q} \cos \frac{2k\pi}{p} x + 4 \sum_{i<j \leq q} \cos \frac{2i\pi}{p} x \cos \frac{2j\pi}{p} x$$

$$= \sum_{k=1}^{q} \left( 1 + \cos \frac{4k\pi}{p} x \right) + \frac{1}{2} + 2 \sum_{k=1}^{q} \cos \frac{2k\pi}{p} x + 4 \sum_{i<j \leq q} \cos \frac{2i\pi}{p} x \cos \frac{2j\pi}{p} x$$

$$= \frac{p}{2} + \sum_{k=1}^{q} \cos \frac{4k\pi}{p} x + 2 \sum_{k=1}^{q} \cos \frac{2k\pi}{p} x + 2 \sum_{i<j \leq q} \left( \cos \frac{2(j+i)\pi}{p} x + \cos \frac{2(j-i)\pi}{p} x \right)$$

Now consider given $k \leq q$. There are $q-k$ combinations for $j-i = k$, and $\lfloor \frac{k-1}{2} \rfloor$ combinations for $i+j = k$, so the coefficient of $\cos \frac{2k\pi}{p} x$ is

$$a_k = \begin{cases} 2(1 + (q - k) + \frac{k-1}{2}) & \text{when } k \text{ is odd} \\ 2(1 + (q - k) + \frac{k-2}{2}) + 1 & \text{when } k \text{ is even number} \end{cases}$$

$$= 2q + 1 - k = p - k$$

For $q < k \leq p - 1$, there’s no combination for $j-i = k$ because $1 \leq i < j \leq q$, and $\lfloor \frac{p-k}{2} \rfloor$ combination for $i+j = k$, so the coefficient of $\cos \frac{2k\pi}{p} x$ is

$$a_k = \begin{cases} 2(\frac{p-k}{2}) & \text{when } k \text{ is odd} \\ 2(\frac{p-k-1}{2}) + 1 & \text{when } k \text{ is even number} \end{cases}$$

$$= p - k$$

5
Therefore for all \(1 \leq k \leq p - 1\), \(a_k = p - k\), and this gives the identity

\[
f(p, x) = \frac{p}{2} + \sum_{k=1}^{p-1} (p - k) \cos \left(\frac{2k\pi}{p} x\right) = 2\left(\sum_{k=1}^{\frac{1}{2}(p-1)} \cos \frac{2k\pi}{p} x + \frac{1}{2}\right)^2
\]

\[\blacksquare\]

**Note:** I also found the following identity but it’s more complicated to prove. Proof of this identity is not necessary for this article so it’s not included.

\[
\frac{p}{2} + \sum_{k=1}^{p-1} (p - k) \cos \left(\frac{2k\pi}{p} x\right) = 2^{p-2} \prod_{k=1}^{p-1} (1 - \cos \frac{2\pi}{p} (x - k))
\]

Now let’s use the identity (4) to prove the properties of \(f(p, x)\).

**Case 1:** When \(x \in \mathbb{N}\) and \(p \nmid x\), \(f(p, x) = 0\).

*Subproof.* Due to the identity (2),

\[
f(p, 1) = 2\left(\sum_{k=1}^{\frac{1}{2}(p-1)} \cos \frac{2k\pi}{p} + \frac{1}{2}\right)^2
\]

\[
= 2\left(-\frac{1}{2} + \frac{1}{2}\right)^2 = 0
\]

For \(n \in \mathbb{N}\) that \(p \nmid n\), set \(\{\cos \frac{2nk\pi}{p}, k = 1, \ldots, \frac{1}{2}(p - 1)\}\) is just permutation of set \(\{\cos \frac{2k\pi}{p}, k = 1, \ldots, \frac{1}{2}(p - 1)\}\), so that \(f(p, n) = f(p, 1) = 0\)

\[\blacksquare\]

**Case 2:** When \(p \mid x\), \(f(p, x) = \frac{p^2}{2}\).

*Subproof.* Let \(x = np\)

\[
f(p, x) = f(p, np) = \frac{p}{2} + \sum_{k=1}^{p-1} (p - k) \cos(2nk\pi)
\]

\[
= \frac{p}{2} + \sum_{k=1}^{p-1} (p - k)
\]

\[
= \frac{p}{2} + \frac{p(p - 1)}{2} = \frac{p^2}{2}
\]

\[\blacksquare\]

**Case 3:** Otherwise \(f(p, x) > 0\).
Subproof. Due to identity (4), \( f(p, x) \geq 0 \) and it is cosine function of period \( p \). 
\[
\cos \left( \frac{2(p-1)\pi}{p} \right)
\] is sinusoid of highest frequency \( p - 1 \) and it has \( p - 1 \) minima during period \( p \), hence \( f(p, x) \) at most has \( p - 1 \) minima during period \( p \). Because \( k = 1, \ldots, p - 1 \) are all the \( p - 1 \) 0 minima on \([0, p]\), therefore \( f(p, x) > 0 \) on all other points.  ■

4 Finite Fourier Series of Prime, Twin Prime, Goldbach Sum Distribution

Now we can use the above basic Fourier series \( f(p, x) \) to construct the series that show relative prime, twin prime and Goldbach sum distributions as zeros.

4.1 Fourier Series of Co-Prime and Prime Distribution \( F(P_i, x) \)

Given a set of primes \( P \), define

\[
F(P, x) = \sum_{p \in P} f(p, x)
\]

then \( F(P, x) \) is a cosine Fourier series of period \( \text{prod}(P) \) that has the following properties

\[
F(P, x) \begin{cases} 
= 0, & \forall x \in \mathbb{Z}, x \text{ co-prime to } \forall p \in P \\
= \frac{x^2}{2}, & \forall x \in P \\
> 0, & \text{elsewhere}
\end{cases}
\]

Particularly, let \( p_i \) be the \( i \)th prime and \( P_i = \{ p \in \mathbb{P} : p \leq p_i \} \), then \( F(P_i, x) \) has period \( p_i \# \) and is symmetric at half period point \( \frac{1}{2}p_i\# \), and it has the following properties

\[
F(P_i, x) = \sum_{p \leq p_i} \left( \frac{p}{2} + \sum_{k=1}^{p-1} (p - k) \cos \frac{2k\pi}{p} x \right) \begin{cases} 
= 0, & \forall x \in \mathbb{Z}, x \text{ co-prime to } \forall p \in P_i \\
= \frac{x^2}{2}, & \forall x \in P_i \\
> 0 & \text{elsewhere}
\end{cases}
\]

Let’s see example of \( F(P_3, x) = f(2, x) + f(3, x) + f(5, x) \), which has period 30 and half period 15. In this example, except trivial zero 1, all zeros of \( F(P_3, x) \) are primes when

\[
n < p_1^2 = 49 \text{ and co-primes of } P_3 \text{ when } n \geq 49, \text{ and it has reflection symmetry at half period point } 15. \text{ It’s also worth note that the first non-prime zero (except 1) is always } p_{i+1}^2 \text{ which is 49 in this case.}
\]

![Figure 4: F(P_3, x) shows all primes 5 < p ≤ 7^2](image)
Theorem 4.1. The number of zeros of $F(P_i, x)$ on domain $[0, p_i \#)$ is exactly

$$\prod_{p \leq p_i} (p - 1)$$

In the $F(P_3, x)$ example above, it has period 30 and exactly $(5 - 1) \cdot (3 - 1) \cdot (2 - 1) = 8$ zeros on $[0, 30)$.

Figure 5: $F(P_4, x)$ shows all primes $p \leq 11^2$

Figure 6: $F(P_4, x)$ shows all co-primes when $n \geq 11^2$

Let’s see another example of $F(P_4, x)$. Because $7\# = 210$, $F(P_4, x)$ has period 210 and mirror symmetry at half period point 105, and it has exactly $(7 - 1) \cdot (5 - 1) \cdot (3 - 1) \cdot (2 - 1) = 48$ zeros on $[0, 210)$ and 24 zeros on $[0, 105)$. 
The proof is due to Chinese Remainder Theorem cited here.

**Theorem.** [1] **Chinese Remainder Theorem (CRT)** Let \( m_1, m_2, \ldots, m_n \) be pairwise coprime, (i.e. \( \forall i \neq j, \gcd(m_i, m_j) = 1 \)). Then the system of \( n \) congruence equations

\[
\begin{align*}
x &\equiv a_1 \pmod{m_1} \\
x &\equiv a_2 \pmod{m_2} \\
&\vdots \\
x &\equiv a_n \pmod{m_n}
\end{align*}
\]

has unique solution for \( x \) modulo \( M \), where \( M = m_1 m_2 \cdots m_n \). Let \( b_i = M/m_i \) and \( b'_i = b_i^{-1} \pmod{m_i} \), then

\[
x = \sum_{i=1}^{n} a_i b_i b'_i \pmod{M}
\]

**Proof of Theorem 3.1.** Let \( p_i \) be \( i^{th} \) prime. Consider the congruence equation system

\[
x \equiv a_k \pmod{p_k}, \quad k \leq i.
\]

Let \( M = p_i \# \),

\[
p_k \nmid x \iff a_k \neq 0, a_k \in \mathbb{Z}_{p_k}
\]

Therefore for each \( p_k \), \( a_k \) has \( p_k - 1 \) choices. So that for all \( p \leq p_i \), the total combinations of \( a_k \) are

\[
Y = \prod_{p \leq p_i} (p - 1)
\]

This means the congruence system has \( Y \) solutions on \([0, M]\), so that \( F(p_i, x) \) has \( Y \) zeros on \([0, p_i \#)\). \( \square \)

### 4.2 Normalization of \( F(P, x) \)

One observation of \( F(P, x) \) is that it puts all points of \( p \in P_i \) on parabola \( y = \frac{x^2}{2} \), which grows quickly and may make the graphing unpractical when \( p_i \) get very big. We can normalize the \( f(p, x) \) so that \( 0 \) shows \( p \nmid x \) and \( 1 \) shows \( p \mid x \) as following

\[
f_*(p, x) = \frac{2}{p^2} f(p, x) \begin{cases} 
0 & \text{when } x \text{ is integer and } p \nmid x \\
1 & \text{when } x \text{ is integer and } p \mid x \\
> 0 & \text{otherwise.}
\end{cases}
\]

Let’s denote the normalized format of \( F(P, x) \) as \( F_*(P, x) \),

\[
F_*(P, x) = \sum_{p \leq p_i} f_*(p, x) = \sum_{p \leq p_i} \frac{1}{p} \left( 1 + 2 \sum_{k=1}^{p-1} (1 - \frac{k}{p}) \cos \frac{2k\pi x}{p} \right)
\]

\[
\begin{cases} 
0, \forall x \in \mathbb{Z}, \forall p \in P_i, (x, p) = 1 \\
\text{number of } p|x, p \in P_i \\
> 0 \text{ otherwise}
\end{cases}
\]

9
Figure 7: $F_*(P_4, x)$ shows number of divisors of $x$ in $P_4$

A good feature of the normalized function $F_*(P_i, x)$ is that it nicely shows how many prime numbers $p \leq p_i$ that divide $x$. So it keeps all co-primes as $y = 0$ and put all power of primes on $y = 1$, and put all numbers with 2 unique prime factors on the $y = 2$, and so on.

In the example of $F_*(P_4, x)$, $P_4 = 2, 3, 5, 7$, the graph shows that prime numbers like 11 and 17 have no divisor; 6, 24, 100 have 2 divisors; 30, 42, 60, 70, 84, 90 have 3 divisors; and 0 has 4 divisors.

4.3 Fourier Series of Twin Prime Distribution $H(P_i, x)$

First let’s consider twin co-primes of given prime $p$, which equals to that $p$ cannot divide both $x$ and $x + 2$. Due to the properties of $f(p, x)$,

$$(p \nmid x) \land (p \nmid x + 2) \iff f(p, x) + f(p, x + 2) = 0$$

$$f(p, x) + f(p, x + 2) = p + \sum_{k=1}^{p-1} (p - k) \left( \cos \frac{2k\pi}{p} x + \cos \frac{2k\pi}{p} (x + 2) \right)$$

$$= p + 2 \sum_{k=1}^{p-1} (p - k) \cos \frac{2k\pi}{p} \cos \frac{2k\pi}{p} (x + 1)$$

Let $t = x + 1$ and define

$$h(p, t) = \frac{p}{2} + \sum_{k=1}^{p-1} (p - k) \cos \frac{2k\pi}{p} \cos \frac{2k\pi}{p} t$$

Then $\forall t_0 \in \mathbb{R}$,

$$h(p, t_0) = 0 \iff (t_0 - 1, t_0 + 1) \text{ is a pair of twin co-primes of } p$$

Let’s see some examples.

Figure 8: $h(3, t)$
Theorem 4.2.

\[ h(p, t) = 0 \iff t \neq \pm 1 \pmod{p}, t \in \mathbb{Z} \]

Proof. By definition of \( t=x+1 \) and \( h(p, t) \) above,

\[
h(p, t) = \frac{1}{2}(f(p, t-1) + f(p, t+1)) \implies \]
\[
h(p, t) = 0 \iff f(p, t-1) = 0 \text{ and } f(p, t+1) = 0 \iff p \nmid (t-1) \text{ and } p \nmid (t+1), t \in \mathbb{Z} \iff t - 1 \neq 0 \pmod{p} \text{ and } t + 1 \neq 0 \pmod{p}, t \in \mathbb{Z} \iff t \neq \pm 1 \pmod{p}, t \in \mathbb{Z}
\]

In summary, \( h(p, t) \) has following properties

\[
h(p, t) = \begin{cases} 
0, & \text{when } t \neq \pm 1 \pmod{p}, t \in \mathbb{Z} \\
\frac{p^2}{2c_p}, & \text{when } t \equiv \pm 1 \pmod{p}, t \in \mathbb{Z} \\
> 0, & \text{elsewhere}
\end{cases}
\]

where

\[
c_p = \begin{cases} 
1, & \text{when } p = 2 \\
2, & \text{when } p \geq 3
\end{cases}
\]
Therefore the normalized form that scales the max value to 1 is

\[ h_*(p,t) = \frac{2c_p}{p^2} h(p,t) \]

\[ = \frac{c_p}{p} \left( 1 + 2 \sum_{k=1}^{p-1} \left( 1 - \frac{k}{p} \right) \cos \frac{2k\pi}{p} \cos \frac{2k\pi}{p} t \right) \]

**Corollary 4.3.** \( h(p,t) \) has \( p - 2 \) zeros during period \( p \).

Given a set of primes \( P \), define

\[ H(P,t) = \sum_{p \in P} h(p,t) = \sum_{p \in P} \left( \frac{p}{2} + \sum_{k=1}^{p-1} (p - k) \cos \frac{2k\pi}{p} \cos \frac{2k\pi}{p} t \right) \]

\[ H_*(P,t) = \sum_{p \in P} h_*(p,t) = \sum_{p \in P} \frac{c_p}{p} \left( 1 + 2 \sum_{k=1}^{p-1} \left( 1 - \frac{k}{p} \right) \cos \frac{2k\pi}{p} \cos \frac{2k\pi}{p} t \right) \]

Then set \( H_P = \{ x : x = t \pm 1, H(P,t) = 0 \} \) is the set of all twin co-primes of set \( P \). Particularly, let \( p_i \) be the \( i^{th} \) prime and \( P_i = \{ p \in \mathbb{P} : p \leq p_i \} \), then \( H_{P_i} \) is set of all twin co-primes of \( p \leq p_i \), and set \( \{ x : x \in H_P, x < p_i^2 + 1 \} \) is set of twin primes less than \( p_i^2 + 1 \).

---

**Figure 11:** \( H(P_4,t) = 0 \) shows all twin primes \((t-1, t+1)\) for \( t < 11^2 - 1 \)

**Figure 12:** \( H_*(P_4,t) = 0 \) shows number of \( p \in P_i \) that \( t \equiv \pm 1 \) (mod \( p \))
Figure of $H_s(P_4, t)$ shows that there are 3 primes $p = [2, 3, 7]$ for 13 that $13 \equiv \pm 1 \pmod{p}$, 4 primes $[2, 3, 5, 7]$ for 29, 1 prime 5 for 54, 4 primes $[2, 3, 5, 7]$ for 71, and 2 primes $[3, 5]$ for 94, and so on. It would be an interesting study of distribution of the integers on these horizontal lines.

**Theorem 4.4.** The number of zeros of $H(P_i, x)$ on domain $[0, p_i \#)$ is exactly

$$\prod_{3 \leq p \leq p_i} (p - 2)$$

**Proof.** Consider the congruence equation system $x \equiv a_k (\mod{p_k}), k = 1, 2, \cdots, i$,

$$h(p_k, x) = 0 \iff a_k \neq \pm 1, a_k \in \mathbb{Z}_{p_k}$$

Therefore for $p_1 = 2, x = 0 \pmod{2}$ and hence $a_1 = 0$, and for each $p_k \geq 3$, $a_k$ has $p_k - 2$ choices. So that for all $p \leq p_i$, the total combinations are

$$Y = \prod_{3 \leq p \leq p_i} (p - 2).$$

According to the CRT, this congruence equation system has $Y$ unique solutions on $[0, p_i \#)$, and each solution is a zero of $H(P_i, x)$.

**Corollary 4.5.** Zeros of $h(2, x)$ are even numbers and zeros of $h(3, x)$ are multiplications of 3, so that zeros of $H(P_i, x), i \geq 2$ must be multiplication of 6. Therefore twin primes must be in format of $6n \pm 1$.

The properties of $H(P_i, x)$ provides a new approach to prove twin prime theorem, by approving the following conjecture.

**Conjecture 4.6.** For any $p_i, \exists t_0 < p_{i+1}^2 - 1, H(P_i, t_0) = 0$.

If this conjecture is true, because co-prime of $P_i$ less then $p_{i+1}^2$ must be prime, $(t_0 - 1, t_0 + 1)$ are prime pair greater than $p_i$. Therefore for any arbitrarily big $p_i$, there always exists twin prime pair greater than $p_i$, therefore there are infinitely many twin primes. This is stronger statement of the Twin Prime Conjecture.

I also found the following conjecture about the relationship between the density of prime and twin prime pairs. Proof of this conjecture could also lead to proof of twin prime conjecture, and additionally provide a very good twin prime estimate function.

**Conjecture 4.7.** Given domain $[0, p_i \#)$, let $\rho_2$ be the density of twin prime pairs over twin co-prime pairs, and $\rho$ be the density of prime numbers over co-prime numbers, then

$$\rho_2 \approx \rho^2.$$ 

Using the notation $\pi_2(x)$ for number of twin prime pairs less than $x$, and the theorem 3.1 and 3.4, the conjecture could be written as following

$$\frac{\pi_2(p_i \#)}{\prod_{3 \leq p \leq p_i} (p - 2)} \approx \left(\frac{\pi(p_i \#)}{\prod_{p \leq p_i} (p - 1)}\right)^2.$$
Actually, \( \rho_2 \) could be the density of any prime pairs over co-prime pairs, including twin prime pairs, Goldbach sum pairs of \( 2n \), and prime pairs with distance \( 2n \), and the same formula \( \rho_2 \approx \rho^2 \) is observed to be true.

As observed, \( \rho_2 \approx \rho^2 \) is also true for arbitrary big domain \((0, N)\).

I’m still in progress to prove this conjecture, but the computed data tabulated below validates that \( \rho_2 \approx \rho^2 \) when \( p > 11 \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>( p_n # )</th>
<th>( \rho )</th>
<th>( \rho^2 )</th>
<th>( \rho_2)</th>
<th>Diff%</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2,310</td>
<td>0.71458333</td>
<td>0.51062934</td>
<td>0.51111111</td>
<td>-0.09%</td>
</tr>
<tr>
<td>6</td>
<td>30,030</td>
<td>0.56388889</td>
<td>0.31797068</td>
<td>0.31515152</td>
<td>0.89%</td>
</tr>
<tr>
<td>7</td>
<td>510,510</td>
<td>0.45932075</td>
<td>0.21097555</td>
<td>0.20812570</td>
<td>1.35%</td>
</tr>
<tr>
<td>8</td>
<td>9,699,690</td>
<td>0.33657788</td>
<td>0.15166106</td>
<td>0.15172113</td>
<td>-0.04%</td>
</tr>
<tr>
<td>9</td>
<td>223,092,870</td>
<td>0.38943685</td>
<td>0.15166106</td>
<td>0.15172113</td>
<td>0.53%</td>
</tr>
<tr>
<td>10</td>
<td>6,469,693,230</td>
<td>0.29394128</td>
<td>0.08640147</td>
<td>0.08599424</td>
<td>0.47%</td>
</tr>
</tbody>
</table>

Corollary 4.8. Let \( N = p_i\# \), then

\[
\pi_2(N) \approx \left( \prod_{3 \leq p \leq p_i} \frac{(p-2)}{(p-1)^2} \right) \pi^2(N) = 2 \left( \prod_{3 \leq p \leq p_i} \frac{p(p-2)}{(p-1)^2} \right) \frac{\pi^2(N)}{N} \tag{5}
\]

For big \( p_i \), use \( \pi(N) \approx \frac{N}{\ln N} \), the formula matches the First Hardy–Littlewood conjecture [3].

\[
\prod_{3 \leq p \leq p_i} \frac{p(p-2)}{(p-1)^2} \approx C_2 = \prod_{3 \leq p} \frac{p(p-2)}{(p-1)^2} \sim 0.6601618158\ldots
\]

and this corollary could be written as

\[
\pi_2(N) \approx 2C_2 \frac{\pi^2(N)}{N} \tag{6}
\]

\[
\approx 2C_2 \frac{N^2}{N \ln^2 N} = 2C_2 \frac{N}{\ln^2 N} \tag{7}
\]

The table shows \( \pi_2(N) \) estimation using formula (5) and (6) vs. actual numbers. Note that estimation has been rounded but the error rate is calculated more preciously with decimals before round.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( N )</th>
<th>( \pi(N) )</th>
<th>( \pi_2(N) )</th>
<th>Est(5)</th>
<th>Error(5)</th>
<th>Est(6)</th>
<th>Error(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2,310</td>
<td>343</td>
<td>140</td>
<td>138</td>
<td>-1.43%</td>
<td>134</td>
<td>-4.29%</td>
</tr>
<tr>
<td>6</td>
<td>30,030</td>
<td>3,248</td>
<td>936</td>
<td>944</td>
<td>0.85%</td>
<td>928</td>
<td>-0.85%</td>
</tr>
<tr>
<td>7</td>
<td>510,510</td>
<td>42,331</td>
<td>9,272</td>
<td>9,398</td>
<td>1.36%</td>
<td>9,269</td>
<td>-0.03%</td>
</tr>
<tr>
<td>8</td>
<td>9,699,690</td>
<td>646,029</td>
<td>114,906</td>
<td>114,860</td>
<td>-0.04%</td>
<td>113,620</td>
<td>-1.12%</td>
</tr>
<tr>
<td>9</td>
<td>223,092,870</td>
<td>12,283,531</td>
<td>1,792,124</td>
<td>1801718</td>
<td>0.54%</td>
<td>1,785,958</td>
<td>-0.34%</td>
</tr>
<tr>
<td>10</td>
<td>6,469,693,230</td>
<td>300,369,796</td>
<td>36,927,426</td>
<td>37,102,300</td>
<td>0.47%</td>
<td>36,824,700</td>
<td>-0.28%</td>
</tr>
</tbody>
</table>
4.4 Fourier Series of Goldbach Sum Distribution $G(n, x)$

**Theorem.** Goldbach Conjecture Every even integer greater than 2 can be expressed as the sum of two prime numbers.

Given a big even number $2n$, the number pair $(x, 2n - x)$ being co-prime to $p$ equals to $f(p, x) + f(p, 2n - x) = 0$.

$$f(p, x) + f(p, 2n - x) = p + \sum_{k=1}^{p-1} (p - k) \left( \cos \frac{2k\pi}{p} x + \cos \frac{2k\pi}{p} (2n - x) \right)$$

$$= p + 2 \sum_{k=1}^{p-1} (p - k) \cos \frac{2kn\pi}{p} \cos \frac{2k\pi}{p} (x - n)$$

Let $t = x - n$ and $g(p, n, t) = \frac{1}{2} (f(p, x) + f(p, 2n - x))$, then

$$g(p, n, t) = \frac{p}{2} + \sum_{k=1}^{p-1} (p - k) \cos \frac{2kn\pi}{p} \cos \frac{2k\pi}{p} t$$

Then $\forall t_0 \in [0, n]$,

$$g(p, n, t_0) = 0 \iff (n - t_0, n + t_0) \text{ is a pair of co-prime of } p \text{ which sum is } 2n$$

And we can also see that $\forall t_0 > n$,

$$g(p, n, t_0) = 0 \iff (t_0 - n, t_0 + n) \text{ is a pair of co-prime of } p \text{ which distance is } 2n.$$  

When $p \mid n$, because $\cos \frac{2kn\pi}{p} = 1$, $g(p, n, x) = f(p, x)$. Take $2n = 100$ as example, Let’s see some graphs for $p \nmid n$ and $p \mid n$,

![Figure 13: g(3, 50, t)](image)

**Theorem 4.9.**

$$g(p, n, t) = 0 \iff t \neq \pm n(mod p), t \in \mathbb{Z}$$

**Proof.** By definition of $t = x - n$ and $g(p, n, t)$ above, and $f(p, x) = f(p, -x)$

$$g(p, n, t) = \frac{1}{2} (f(p, t + n) + f(p, t - n)) \implies$$

$$g(p, n, t) = 0 \iff f(p, t - n) = 0 \text{ and } f(p, t + n) = 0$$

$$\iff p \nmid (t - n) \text{ and } p \nmid (t + n), t \in \mathbb{Z}$$

$$\iff t - n \neq 0(mod p) \text{ and } t + n \neq 0(mod p), t \in \mathbb{Z}$$

$$\iff t \neq \pm n(mod p), t \in \mathbb{Z}$$
Because $g(p, n, t) = f(p, t)$ when $p \mid n$ and the above theorem, it’s easy to see the following corollary.

**Corollary 4.10.** $g(p, n, t)$ has $p - 1$ zeros during period $p$ when $p \mid n$ and $p - 2$ zeros when $p \nmid n$.

For example, because $(50 \mod 7) = 1$, $g(7, 50, 1)$ and $g(7, 50, 6)$ equal to $\frac{7^2}{4}$ and all other integers are zeros during period $[0,7)$.

Now define

$$G(n, t) = \sum_{p < \sqrt{2n}} g(p, n, t) = \sum_{p < \sqrt{2n}} \left( \frac{p}{2} + \sum_{k=1}^{p-1} (p - k) \cos \frac{2kn\pi}{p} \cos \frac{2k\pi}{p} t \right)$$

And let set $G_{2n} = \{ x | x = (n - t, n + t), G(n, t) = 0, t \in [0,n] \}$, then $G_{2n}$ is the set of all prime pairs sum to $2n$, excluding the pairs $(p, 2n - p)$ because $(n - p) \equiv n \pmod{p}$. 

16
**Theorem 4.11.** Let $p_i$ be the largest prime less than $\sqrt{2n}$ and $p \leq p_i$, then the number of zeros of $G(n,t)$ on domain $[0,p_i\#)$ is exactly

$$\prod_{p|n}(p-1) \prod_{p \nmid n}(p-2)$$

**Proof.** Consider the congruence equation system $x \equiv a_k(\mod p_k), k = 1, 2, \ldots, i$.

$$g(p_k, n, t) = 0 \iff a_k \neq \pm n(\mod p_k), a_k \in \mathbb{Z}_{p_k}$$

Therefore for $p_1 = 2, a_1 \neq n(\mod 2)$ has only one choice either 0 or 1, and for each $p_k \geq 3$, $a_k$ has $p_k - 2$ choices if $p_k \nmid n$ and $p_k - 1$ choices if $p_k | n$. So that for all $p_k \leq p_i$, the total number of combinations of $a_k$ is

$$\prod_{p|n}(p-1) \prod_{p \nmid n}(p-2)$$

Based on CRT, each set of $a_k$ for this congruence system has an unique solution on $[0, p\#)$, and each solution is one zero of $G(n, t)$, so that $G(n, t)$ has same number of zeros on $[0, p\#)$.

\[\square\]

![Figure 16: $G(50, t)$](image1)

![Figure 17: $G_{100}$ shows prime pairs sum to 100, excluding (3,97)](image2)
**Corollary 4.12.** If \( n \) is power of 2, let \( p_i \) be the largest prime less than \( \sqrt{2n} \), then the number of zeros of \( G(n,t) \) on domain \([0, p\#)\) is exactly

\[
\prod_{3 \leq p \leq p_i} (p - 2)
\]

which is same as number of zeros of \( H(P_i, t) \) on the same domain \([0, p\#)\).

This is a hint that the distribution of Goldbach sums of \( 2n \) and twin primes less than \( n \) are similar, and could be estimated by the same formula. The difference is that now the Goldbach pairs are distributed on \([0, 2n]\) vs. \([0, n]\) for twin prime. This leads to the following proposition.

**Proposition 4.13.** Given \( N = 2n \) be a big number of power of 2, let \( \pi_g(N) \) denote number of ways that \( N \) could be written as sum of two primes, and \( \pi_n(N) \) denote number of prime pairs of distance \( n \) less than \( 2n \), then

\[
\pi_g(N) \approx \pi_n^N(N) \approx C_2 \frac{\pi^2(N)}{N} \approx C_2 \frac{\pi^2(N)}{\ln^2 N}
\]

The following is the table of computed data of actual vs. estimation with formula (8). Note that estimation has been rounded but the error rate is calculated more preciously with decimals before round.
Some numbers show more offs but overall the formula (8) is a good estimation, especially when $N$ is big. It’s also interesting to see that number of Goldbach prime sums of $N$ is approximately same as the number of prime pairs of distance $\frac{N}{2}$ less than $N$.

Now let’s consider the big even number $N = 2^n$ that has other prime divisors. Let $p_i$ be the largest prime less than $\sqrt{2^n}$ and $3 \leq p \leq p_i$, due to theorem 3.11, the number of zeros of $G(n, t)$ on domain $[0, p_i \#)$ is exactly

$$\prod_{p|n} (p - 1) \prod_{p|n} (p - 2) = \prod_{p|n} \frac{p - 1}{p - 2} \prod_{3 \leq p \leq p_i} (p - 2)$$

Applying the corollary 3.12 and formula (8) for $N$ of power of 2 that has no any other prime divisor, this gives us the following general estimation formula of $\pi_g(N)$ and $\pi_n(N)$,

**Proposition 4.14.** Let $N = 2^n,$

$$\pi_g(N) \approx \pi_n(N) \approx C_2 \frac{\pi^2(N)}{N} \prod_{p|n} \frac{p - 1}{p - 2}$$

(10)

$$\approx C_2 \frac{N}{\ln^2 N} \prod_{p|n} \frac{p - 1}{p - 2}$$

(11)

<table>
<thead>
<tr>
<th>$n$</th>
<th>$N = 2^n$</th>
<th>Est(8)</th>
<th>$\pi_g(N)$</th>
<th>$\pi_g(N)$</th>
<th>$\pi_g$ Error</th>
<th>$\pi_g$ Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1,024</td>
<td>19</td>
<td>21</td>
<td>18</td>
<td>-9.52%</td>
<td>6.11%</td>
</tr>
<tr>
<td>11</td>
<td>2,048</td>
<td>31</td>
<td>25</td>
<td>32</td>
<td>24.00%</td>
<td>-2.50%</td>
</tr>
<tr>
<td>12</td>
<td>4,096</td>
<td>51</td>
<td>52</td>
<td>51</td>
<td>-1.92%</td>
<td>-0.20%</td>
</tr>
<tr>
<td>13</td>
<td>8,192</td>
<td>85</td>
<td>76</td>
<td>84</td>
<td>11.84%</td>
<td>1.19%</td>
</tr>
<tr>
<td>14</td>
<td>16,384</td>
<td>145</td>
<td>150</td>
<td>151</td>
<td>-3.33%</td>
<td>-3.31%</td>
</tr>
<tr>
<td>15</td>
<td>32,768</td>
<td>248</td>
<td>243</td>
<td>254</td>
<td>2.06%</td>
<td>-1.97%</td>
</tr>
<tr>
<td>16</td>
<td>65,536</td>
<td>431</td>
<td>435</td>
<td>429</td>
<td>-0.92%</td>
<td>0.47%</td>
</tr>
<tr>
<td>17</td>
<td>131,072</td>
<td>756</td>
<td>749</td>
<td>754</td>
<td>0.93%</td>
<td>0.27%</td>
</tr>
<tr>
<td>18</td>
<td>262,144</td>
<td>1,332</td>
<td>1,313</td>
<td>1,311</td>
<td>1.45%</td>
<td>1.60%</td>
</tr>
<tr>
<td>19</td>
<td>524,288</td>
<td>2,370</td>
<td>2,367</td>
<td>2,377</td>
<td>0.13%</td>
<td>-0.29%</td>
</tr>
<tr>
<td>20</td>
<td>1,048,576</td>
<td>4,236</td>
<td>4,238</td>
<td>4,199</td>
<td>-0.05%</td>
<td>0.88%</td>
</tr>
<tr>
<td>21</td>
<td>2,097,152</td>
<td>7,622</td>
<td>7,584</td>
<td>7,542</td>
<td>0.20%</td>
<td>0.51%</td>
</tr>
<tr>
<td>22</td>
<td>4,194,304</td>
<td>13,785</td>
<td>13,704</td>
<td>13,785</td>
<td>0.59%</td>
<td>0.00%</td>
</tr>
<tr>
<td>23</td>
<td>8,388,608</td>
<td>25,048</td>
<td>24,955</td>
<td>24,955</td>
<td>0.48%</td>
<td>0.37%</td>
</tr>
<tr>
<td>24</td>
<td>16,777,216</td>
<td>45,715</td>
<td>45,840</td>
<td>45,840</td>
<td>-0.07%</td>
<td>-0.27%</td>
</tr>
<tr>
<td>25</td>
<td>33,554,432</td>
<td>83,789</td>
<td>84,002</td>
<td>84,002</td>
<td>0.39%</td>
<td>-0.25%</td>
</tr>
<tr>
<td>26</td>
<td>67,108,864</td>
<td>154,092</td>
<td>154,236</td>
<td>154,236</td>
<td>0.16%</td>
<td>0.38%</td>
</tr>
<tr>
<td>27</td>
<td>134,217,728</td>
<td>284,363</td>
<td>284,236</td>
<td>284,236</td>
<td>0.22%</td>
<td>0.25%</td>
</tr>
<tr>
<td>28</td>
<td>268,435,456</td>
<td>526,440</td>
<td>526,440</td>
<td>526,440</td>
<td>0.23%</td>
<td>0.34%</td>
</tr>
<tr>
<td>29</td>
<td>536,870,912</td>
<td>977,362</td>
<td>976,849</td>
<td>976,849</td>
<td>0.17%</td>
<td>0.12%</td>
</tr>
<tr>
<td>30</td>
<td>1,073,741,824</td>
<td>1,589,609</td>
<td>1,542,054</td>
<td>1,542,054</td>
<td>3.08%</td>
<td>0.10%</td>
</tr>
<tr>
<td>31</td>
<td>2,147,483,648</td>
<td>3,395,516</td>
<td>3,390,036</td>
<td>3,390,036</td>
<td>0.16%</td>
<td>0.08%</td>
</tr>
<tr>
<td>32</td>
<td>4,294,967,296</td>
<td>6,351,566</td>
<td>6,341,423</td>
<td>6,341,423</td>
<td>0.16%</td>
<td>0.05%</td>
</tr>
<tr>
<td>33</td>
<td>8,589,934,592</td>
<td>11,907,089</td>
<td>11,891,654</td>
<td>11,891,654</td>
<td>0.13%</td>
<td>0.07%</td>
</tr>
</tbody>
</table>
The following is the table of computed data of actual vs. estimation with formula (10). Note that estimation has been rounded but the error rate is calculated more precisely with decimals before round.

| $N$          | $p|n$ | $\prod_{p>3, p^\frac{1}{2}}$ | $\text{Est}(10)$ | $\pi_g(N)$ | $\pi_n(N)$ | $\pi_g \text{ Err}$ | $\pi_n \text{ Err}$ |
|--------------|-------|-----------------------------|------------------|------------|------------|---------------------|---------------------|
| 2,044        | 2,7,73| 1.21690                     | 37               | 40         | 37         | -6.80%              | 0.76%               |
| 2,048        | 2^{11}| 1.00000                     | 31               | 25         | 32         | 24.00%              | -2.50%              |
| 2,052        | 2,3,19| 2.11765                     | 65               | 69         | 64         | -6.33%              | 0.98%               |
| 32,764       | 2,8191| 1.00012                     | 248              | 248        | 245        | 0.17%               | 1.39%               |
| 32,768       | 2^{15}| 1.00000                     | 248              | 243        | 254        | 0.06%               | -1.97%              |
| 32,772       | 2,3,2731| 2.00073                    | 497              | 499        | 497        | -0.38%              | 0.02%               |
| 262,140      | 2,3,5,17,257| 2.85560| 3,804 | 3,751 | 3,792 | 1.14% | 0.31% |
| 262,144      | 2^{18}| 1.00000                     | 1,332            | 1,331      | 1,331      | 1.60%               | 1.45%               |
| 262,148      | 2,65537| 1.00000| 1,332          | 1,331      | 1,320      | 0.09%               | 0.23%               |
| 2,097,148    | 2,524287| 1.00000| 7,623         | 7,516      | 7,619      | 1.42%               | 0.05%               |
| 2,097,152    | 2^{21}| 1.00000                     | 7,622            | 7,471      | 7,584      | 2.02%               | 0.51%               |
| 2,097,156    | 2,3,174763| 2.00001| 15,245      | 15,125      | 15,158     | 0.79%               | 0.57%               |
| 33,554,428   | 2,47,178481| 1.02223| 85,652       | 85,358      | 85,358     | 0.05%               | 0.34%               |
| 33,554,432   | 2^{25}| 1.00000                     | 83,789           | 83,467     | 84,002     | 0.39%               | -0.25%              |
| 33,554,436   | 2,3,2796203| 2.00000| 167,579      | 167,024     | 167,024    | 0.16%               | 0.33%               |
| 268,435,452  | 2,3,2731,8191| 2.00098| 1,053,395   | 1,052,224  | 1,051,397  | 0.11%               | 0.19%               |
| 268,435,456  | 2^{28}| 1.00000                     | 526,440          | 525,236    | 524,648    | 0.23%               | 0.34%               |
| 268,435,460  | 2,5,53,157,1613| 1.36910| 720,748      | 718,617     | 718,617    | 0.08%               | 0.30%               |
| 2,147,483,644| 2,233,1103,2089| 1.00572| 3,414,948   | 3,408,771  | 3,408,771  | 0.13%               | 0.18%               |
| 2,147,483,648| 2^{31}| 1.00000                     | 3,395,516        | 3,390,038  | 3,392,835  | 0.16%               | 0.08%               |
| 2,147,483,652| 2,3,59,3033,169| 2.03509| 6,910,176   | 6,897,845  | 6,897,845  | 0.07%               | 0.18%               |
| 8,589,934,588| 2,2147483647| 1.00000| 11,907,088  | 11,890,625 | 11,890,625 | 0.09%               | 0.14%               |
| 8,589,934,592| 2^{33}| 1.00000                     | 11,907,089       | 11,898,310 | 11,898,310 | 0.13%               | 0.07%               |
| 8,589,934,596| 2,3,715827883| 2.00000| 23,814,177  | 23,781,722 | 23,781,722 | 0.07%               | 0.14%               |

From the formula and the data, we can see that the even number in the form of $N = 2^k$ has least density of Goldbach sums, so if Goldbach conjecture could be proven with $2^k$ of any big $k$, then it could be proven for all big even numbers.

Similar to $H(P, t)$ for twin prime, the properties of $G(n, t)$ gives a new possible approach to prove Goldbach conjecture by approving the following conjecture.

**Conjecture 4.15.** For any big even number $2n$, $\exists t_0 < n, G(n, t_0) = 0$.

Let $p_i$ be the largest prime less than $\sqrt{2n}$. If this conjecture is true, because $2n < p_i^2 + 1 \Rightarrow n \pm t_0 < p_{i+1}^2$, $n - t_0$ and $n + t_0$ are a pair of prime numbers sum to $2n$. 

20
4.5 Unified Fourier Series $P(p_i, n, x)$ of Distribution of Prime, Twin Prime, Goldbach Sum and Prime Pair of Distance $2n$

Now let $p_i$ be the $i^{th}$ prime, and define

$$P(p_i, n, x) = \sum_{p \leq p_i} \left( \frac{p}{2} + \sum_{k=1}^{p-1} (p - k) \cos \frac{2kn\pi}{p} \cos \frac{2k\pi}{p} x \right)$$

then because $\cos 0 = 1$,

$$P(p_i, n, x) = \begin{cases} 
F(P_i, x), & \text{when } n = 0, \text{ prime distribution} \\
H(P_i, x), & \text{when } n = 1, \text{ twin prime distribution} \\
G(n, x), & \text{when } n \geq 1, \text{ Goldbach sums and prime pair of distance } 2n 
\end{cases}$$

The unified formula to calculate number of zeros $L$ on $[0, p_i)$ is, for all $3 \leq p \leq p_i$,

$$L = \prod_{p \mid n} (p - 1) \prod_{p \notmid n} (p - 2) \tag{12}$$

where $n = 0$ for prime series $F(P_i, x)$, so $L = \text{prod}(p - 1)$; $n = 1$ for twin prime series $H(P_i, x)$, so $L = \text{prod}(p - 2)$.

Actually $H(P_i, x)$ of twin prime distribution is a special case of $G(n, x)$ of prime pairs of distance $2n$ when $n = 1$, so they share the similar distribution properties with Goldbach sum distribution. The unified Fourier series $P(p_i, n, x)$ shows the internal relationship between prime, twin prime conjecture and Goldbach conjecture, and I hope this can inspire some ideas toward final proof of these conjectures.

Corresponding to the normalization of $F(P_i, x)$, the normalized form of $P(p_i, n, x)$ is given by

$$P^*(p_i, n, x) = \sum_{p \leq p_i} \frac{c_p}{p} \left( 1 + 2 \sum_{k=1}^{p-1} (1 - \frac{k}{p}) \cos \frac{2k\pi}{p} \cos \frac{2\pi}{p} x \right), \quad c_p = \begin{cases} 
1, & \text{when } p \mid 2n \\
2, & \text{when } p \notmid 2n 
\end{cases}$$

5 Transform to Finite Chebyshev Polynomial Series

These finite Fourier Series could be transformed to finite Chebyshev Polynomial Series, and their zeroes are 1-1 mapped to roots of the polynomials, so that the polynomial theory could be applied to study the distribution of primes, twin primes and Goldbach conjectures.

5.1 Finite Chebyshev Series of $f(p, x)$

The Chebyshev polynomials of the first kind are defined by the recurrence relation [2]

$$T_0(x) = 1$$
$$T_1(x) = x$$
$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$
where \( x \in \mathbb{R}[−1, 1] \). Let \( \theta = \arccos(x) \), its equivalent trigonometric definition is

\[
T_n(\cos \theta) = \cos(n\theta)
\]

\( T_n(x) \) is degree \( n \) polynomial of \( x \). Now consider

\[
f(p, x) = \frac{p}{2} + \sum_{k=1}^{p-1} (p - k) \cos \left( \frac{2k\pi}{p} x \right)
\]

Let \( t = \cos \left( \frac{2\pi}{p} x \right) \), \( x \in \mathbb{R}[0, \frac{p}{2}] \), then \( t \in \mathbb{R}[−1, 1] \) and

\[
\cos \left( \frac{2k\pi}{p} x \right) = T_k(t)
\]

and the \( f(p, x) \) is transformed to the following Chebyshev polynomial series of degree \( p - 1 \).

\[
f_T(p, t) = \frac{p}{2} + \sum_{k=1}^{p-1} (p - k)T_k(t)
\]

\( t = \cos \left( \frac{2\pi}{p} x \right) \) and its reverse function \( x = \frac{p}{2\pi} \arccos(t) \) establishes a bijection between \( \mathbb{R}[−1, 1] \) and \( \mathbb{R}[0, \frac{p}{2}] \), and \( f_T(p, t) = f(p, x) \). Especially, it creates bijection between zeros of \( f(p, x) \) and zeros of \( f_T(p, t) \).

\( f_T(p, t) \) has simple coefficient for calculations, but because \( f_T(p, 1) = f(p, 0) = \frac{p^2}{2} \), it increases quickly when \( p \) getting bigger. So for purpose of graphing, its normalized form becomes handy.

\[
f_{T^*}(p, t) = \frac{2}{p^2} f(p, t) = \frac{1}{p} \left( 1 + 2 \sum_{k=1}^{p-1} (1 - \frac{k}{p})T_k(t) \right)
\]

\( f_{T^*}(p, t) \) reserves all zeros and scales the max \( f_{T^*}(p, 1) \) to 1.

For odd prime \( p \), the roots of \( f_T \) on \([−1, 1]\) are mapped to natural numbers 1 to \( \frac{p-1}{2} \), the zeros of \( f \) on \([0, \frac{p}{2}]\), in reversed order. When \( p = 2 \), the zero is mapped to 1.

<table>
<thead>
<tr>
<th>( f_T(p, t) )</th>
<th>( f_T(2, t) )</th>
<th>( f_T(3, t) )</th>
<th>( f_T(5, t) )</th>
<th>( f_T(7, t) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t )</td>
<td>-1</td>
<td>-0.5</td>
<td>0.309</td>
<td>-0.809</td>
</tr>
<tr>
<td>( x = \frac{p}{2\pi} \arccos(t) )</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

### 5.2 Finite Chebyshev Polynomial Series of \( F(P_i, x) \)

\( F(P_i, x) \) has period \( p_i\# \). Let \( M = p_i\# \) and \( m_p = \frac{M}{p} \).

Let \( t = \cos \left( \frac{2\pi}{M} x \right) \), then \( t \) maps \( x \in [0, \frac{M}{2}] \) to \([-1, 1]\) in reversed order, and

\[
\cos \left( \frac{2k\pi}{p} x \right) = \cos \left( \frac{2\pi}{M} km_p x \right) = T_{km_p}(t)
\]
Figure 18: $f_T(2, t), f_T(3, t), f_T(5, t), f_T(7, t)$

Figure 19: $f_{T^*}(2, t), f_{T^*}(3, t), f_{T^*}(5, t), f_{T^*}(7, t)$

Therefore

$$F(P, x) = \sum_{p \leq p_i} \left( \frac{p}{2} + \sum_{k=1}^{p-1} (p - k) \cos \frac{2k\pi}{p} x \right)$$

$$= \sum_{p \leq p_i} \left( \frac{p}{2} + \sum_{k=1}^{p-1} (p - k) T_{km_p}(t) \right)$$

Denote this as $F_T(P, t)$. Because $km_p$ is max when $p = p_i$ and $k = p - 1$, $F_T(P, x)$ is polynomial of degree $(p_i - 1)m_{p_i - 1}$. e.g. for $p_4 = 7$, $F(4, x)$ is polynomial of degree $(7 - 1) \times 30 = 180$. The normalized form of $F_T(P, t)$ is

$$F_{T^*}(P, t) = \sum_{p \leq p_i} \frac{1}{p} \left( 1 + 2 \sum_{k=1}^{p-1} (1 - \frac{k}{p}) T_{km_p}(t) \right)$$

Figure 20: $F_T(P_4, t)$
The following table shows the 1-1 mapping between roots of $F_T$ and zeros of $F$, $x = \frac{105}{\pi} \arccos(t)$

<table>
<thead>
<tr>
<th>$t$</th>
<th>0.9995</th>
<th>0.9463</th>
<th>0.9253</th>
<th>0.8734</th>
<th>0.7724</th>
<th>0.6466</th>
<th>0.5998</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>1</td>
<td>11</td>
<td>13</td>
<td>17</td>
<td>19</td>
<td>23</td>
<td>29</td>
</tr>
<tr>
<td>$t$</td>
<td>0.4473</td>
<td>0.3374</td>
<td>0.28043</td>
<td>0.16382</td>
<td>-0.01496</td>
<td>-0.1932</td>
<td>-0.2515</td>
</tr>
<tr>
<td>$x$</td>
<td>37</td>
<td>41</td>
<td>43</td>
<td>47</td>
<td>53</td>
<td>59</td>
<td>61</td>
</tr>
<tr>
<td>$t$</td>
<td>-0.5256</td>
<td>-0.5756</td>
<td>-0.7125</td>
<td>-0.7911</td>
<td>-0.8876</td>
<td>-0.9715</td>
<td>-0.9928</td>
</tr>
<tr>
<td>$x$</td>
<td>71</td>
<td>73</td>
<td>79</td>
<td>83</td>
<td>89</td>
<td>97</td>
<td>101</td>
</tr>
</tbody>
</table>

The number of roots is

$$\frac{1}{2} \prod_{p \leq 7} (p - 1) = 24$$

### 5.3 Finite Chebyshev Polynomial Series of $H(P_i, x)$

Using the same definition of $M$, $m_p$ and $t$,

$$h_T(p, t) = \frac{p}{2} + \sum_{k=1}^{p-1} (p - k)(\cos \frac{2k\pi}{p})T_k(t)$$

$$h_{T^*}(p, t) = \frac{c_p}{p} \left( 1 + 2 \sum_{k=1}^{p-1} (1 - \frac{k}{p})(\cos \frac{2k\pi}{p})T_k(t) \right)$$

$$c_p = \begin{cases} 1, & \text{when } p = 2 \\ 2, & \text{when } p \geq 3 \end{cases}$$

$$H_T(P_i, t) = \sum_{p \leq p_i} h_T(p, T_{m_p}(t)) = \sum_{p \leq p_i} \left( \frac{p}{2} + \sum_{k=1}^{p-1} (p - k)(\cos \frac{2k\pi}{p})T_{km_p}(t) \right)$$

$$H_{T^*}(P_i, t) = \sum_{p \leq p_i} h_{T^*}(p, T_{m_p}(t)) = \sum_{p \leq p_i} \frac{c_p}{p} \left( 1 + 2 \sum_{k=1}^{p-1} (1 - \frac{k}{p})(\cos \frac{2k\pi}{p})T_{km_p}(t) \right)$$
Figure 22: $H_{T^*}(P_4, t)$ zeroes show twin prime distribution

The following table shows the 1-1 mapping between roots of $H_{T^*}$ and zeros of $H_s$, by $x = \frac{105}{\pi} \arccos(t)$, and $(x - 1, x + 1)$ is pair of twin primes greater than $p_4 = 7$.

<table>
<thead>
<tr>
<th>$(t, x)$</th>
<th>(0.9362,12)</th>
<th>(0.8585,18)</th>
<th>(0.6233,30)</th>
<th>(0.309,42)</th>
<th>(-0.2226,60)</th>
<th>(-0.551,72)</th>
<th>(-0.996,102)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(x \pm 1)$</td>
<td>(11,13)</td>
<td>(17,19)</td>
<td>(29,31)</td>
<td>(41,43)</td>
<td>(59,61)</td>
<td>(71,73)</td>
<td>(101,103)</td>
</tr>
</tbody>
</table>

The number of roots is

$$\frac{1}{2} \left( \prod_{3 \leq p \leq 7} (p - 2) - 1 \right) = 7$$

5.4 Finite Chebyshev Polynomial Series of $G(n, x)$

Let $p_i$ be the largest prime less than $\sqrt{2n}$, using the same definition of $M, m_p$ and $t$,

$$g_T(p, n, t) = \frac{p}{2} + \sum_{k=1}^{p-1} (p - k)(\cos \frac{2kn\pi}{p})T_k(t)$$

$$g_{T^*}(p, n, t) = \frac{2c_p}{p^2} g_T(p, n, t), c_p = \begin{cases} 
1, & \text{when } p \mid 2n \\
2, & \text{when } p \nmid 2n 
\end{cases}$$

$$= \frac{c_p}{p} \left( 1 + 2 \sum_{k=1}^{p-1} (1 - \frac{k}{p})(\cos \frac{2kn\pi}{p})T_k(t) \right)$$

$$G_T(n, t) = \sum_{p<\sqrt{2n}} g_T(p, n, T_{m_p}(t)) = \sum_{p<\sqrt{2n}} \left( \frac{p}{2} + \sum_{k=1}^{p-1} (p - k)(\cos \frac{2kn\pi}{p})T_{km_p}(t) \right)$$

$$G_{T^*}(n, t) = \sum_{p<\sqrt{2n}} g_{T^*}(p, n, T_{m_p}(t)) = \sum_{p<\sqrt{2n}} \frac{c_p}{p} \left( 1 + 2 \sum_{k=1}^{p-1} (1 - \frac{k}{p})(\cos \frac{2kn\pi}{p})T_{km_p}(t) \right)$$

Let’s see an example for $n = 60$. In this example, $p_i = p_4 = 7$ because the next prime
11 > √2n. 2,3 and 5 divide n, and (60 mod 7)=4.

\[ G_{T_*}(60, t) = \sum_{p < \sqrt{120}} g_{T_*}(p, n, T_{m_p}(t)) \]
\[ = g_{T_*}(2, 60, T_{105}(t)) + g_{T_*}(3, 60, T_{70}(t)) + g_{T_*}(5, 60, T_{42}(t)) + g_{T_*}(7, 60, T_{30}(t)) \]
\[ = \frac{1}{2} g_T(2, 60, T_{105}(t)) + \frac{2}{9} g_T(3, 60, T_{70}(t)) + \frac{2}{25} g_T(5, 60, T_{42}(t)) + \frac{4}{49} g_T(7, 60, T_{30}(t)) \]
\[ = \frac{1}{2} \left( 1 + T_{105}(t) \right) + \frac{2}{9} \left( \frac{3}{2} + 2T_{70}(t) + T_{140}(t) \right) + \]
\[ + \frac{2}{25} \left( \frac{5}{2} + 4T_{42}(t) + 3T_{84}(t) + 2T_{126}(t) + T_{168}(t) \right) + \frac{4}{49} \left( 1 + 6 \cos \frac{8\pi}{7} T_{30}(t) \right) + \]
\[ + \frac{4}{49} \left( 5 \cos \frac{2\pi}{7} T_{60}(t) + 4 \cos \frac{10\pi}{7} T_{90}(t) + 3 \cos \frac{4\pi}{7} T_{120}(t) + 2 \cos \frac{12\pi}{7} T_{150}(t) + \cos \frac{6\pi}{7} T_{180}(t) \right) \]

Figure 23: \( G_{T_*}(60, t) \) zeroes show Goldbach sum distribution of 120

The following table shows the mapping between zeros of \( G_{T_*} \) to zeros of \( G_* \), by \( x = \frac{105}{\pi} \) \( \arccos(t) \). And \( (n - x, n + x) \) is pair of primes of Goldbach sum of \( 2n = 120 \).

| \((t, x)\) | (0.99955, 1) | (0.9781, 7) | (0.9253, 13) | (0.8428, 19) | (0.7724, 23) |
| \((n - x, n + x)\) | (59, 61) | (53, 67) | (47, 73) | (41, 79) | (37, 83) |
| (0.6464, 29) | (0.4473, 37) | (0.3373, 41) | (0.2804, 43) | (0.1638, 47) | (0.1045, 49) |
| (31, 89) | (23, 97) | (19, 101) | (17, 103) | (13, 107) | (11, 109) |

When \( x > n, (x - n, x + n) \) is a pair of primes or co-primes of distance \( 2n = 120 \).

| \((t, x)\) | (-0.2517, 61) | (-0.5255, 71) | (-0.6691, 77) | (-0.7124, 79) |
| \((x - n, x + n)\) | (1, 121) | (11, 131) | (17, 137) | (19, 139) |
| (-0.7911, 83) | (-0.8876, 89) | (-0.9135, 91) | (-0.9715, 97) | (-0.9982, 103) |
| (23, 143) | (29, 149) | (31, 151) | (37, 157) | (43, 163) |

The total number of roots is by given \( p < \sqrt{120} \)

\[ \frac{1}{2} \prod_{p \mid 60} (p - 1) \prod_{p \notmid 60} (p - 2) = 20 \]

### 5.5 Finite Chebyshev Polynomial Series of \( P(p_i, n, x) \)

Summarizing all the Chebyshev polynomial series discussed above, and using the same definition of \( M, m_p \) and \( t \), we can get the following unified Chebyshev polynomial series that
works for distribution of prime when \( n = 0 \), twin prime when \( n = 1 \), and Goldbach sums or prime pairs of distance \( 2n \) when \( n > 1 \).

\[
P_T(p_i, n, t) = \sum_{p \leq p_i} \left( \frac{p}{2} + \sum_{k=1}^{p-1} (p-k) \cos \frac{2kn\pi}{p} T_{km_p}(t) \right)
\]

And its normalization

\[
P_{T*}(p_i, n, x) = \sum_{p \leq p_i} \frac{c_p}{p} \left( 1 + 2 \sum_{k=1}^{p-1} \left( 1 - \frac{k}{p} \right) \cos \frac{2kn\pi}{p} T_{km_p}(t) \right), c_p = \begin{cases} 1, & \text{when } p \mid 2n \\ 2, & \text{when } p \nmid 2n \end{cases}
\]

5.6 Map the Properties of Fourier Series to Chebyshev Series

\( t = \cos \frac{2\pi}{M} x \) maps \( x \in [0, \frac{M}{2}] \) to \([-1, 1]\) in reversed order, also creates a bijection between zeros of \( \tilde{P} \) and \( P_T \), while retaining the properties of \( P \).

For example, the number of roots of \( H_T \) on \([-1, 1]\) equals to number of zeros of \( H \) on \([0, \frac{M}{2}]\), which is exactly

\[
\frac{1}{2} \left( \prod_{3 \leq p \leq p_i} (p-2) - 1 \right)
\]

And the conjectures 3.15 could be converted to the following conjecture of Chebyshev series.

**Conjecture 5.1.** For any big even number \( 2n \), \( \exists t_0 \in (\cos \frac{2n\pi}{M}, 1] \), \( G_T(n, t_0) = 0 \).

Proof of this conjecture is equivalent to proof of conjecture of 3.15, which will result in proof of the Goldbach conjecture.

6 Further research

- Is polynomial \( F_T(P_i, t) \) solvable? If it’s solvable, the solution of roots may derive some prime generating function.

- Study the structure of zeros of these Fourier series on \((0, p_i\#)\). I found many interesting distribution patterns, and these are several examples

**Conjecture 6.1.** For any big \( p_i > 7 \), the gap between zeros of \( H(P_i, x) \) is less than \( \frac{1}{2}p_i^2 \). This means that for any \( p \), there’s at least one zero, hence at least one twin prime, between \( p \) and \( \frac{p_i^2}{2} \), and another between \( \frac{p^2}{2} \) and \( p^2 \).

e.g. for \( p_5 = 11 \), the biggest gap is \( 42 < 60.5 \); for \( p_6 = 13 \), the biggest gap is \( 66 < 84.5 \); for \( p_8 = 19 \), the biggest gap is \( 150 < 180.5 \).

**Conjecture 6.2.** Given any big even number \( 2n \), let \( p_i \) be the largest prime less than \( \sqrt{2n} \), the gap between zeros of \( G(n, x) \) is less than \( \frac{1}{2}p_i^2 \).
If this conjecture is true, it means that for any given even number $2n$, $\exists x_0 < \frac{1}{2}p_i^2 < n$, and $(n - x_0, n + x_0)$ is a pair of primes of sum $2n$. And it leads to proof of Goldbach conjecture (refer to page 20).

Taking the numbers $N = 2n$ as power of 2 for example, which supposed to have least density of Goldbach sums, for $N = 128$ and $p_5 = 11$, the biggest gap is $42 < 60.5$; for $N = 256$ and $p_6 = 13$, the biggest gap is $78 < 84.5$; for $N = 512$ and $p_8 = 19$, the biggest gap is $132 < 180.5$.

We know that $P(p_i, n, x)$ has exactly half zeros on it’s half period, but is there smaller segment that has even distribution of zeros? To my observation the answer is yes as described in the following conjecture.

**Conjecture 6.3** (Near linear distribution of zeros of $P(p_i, n, x)$). Let $M = p_i\#$, $L$ be the number of zeros of $P(p_i, n, x)$ on domain $(0, p_i\#)$, and $S = \gcd(M, L)$, then the domain $(0, p_i\#)$ can be divided into $S$ segments of length $\frac{M}{S}$, and each segment has approximately $\frac{L}{S}$ zeros. In other words, the zeros of $P(p_i, n, x)$ are almost evenly distributed to each segment.

The conjecture is saying that, even though the distribution of individual primes and co-primes are very irregular, but there exist segments that the distribution on these segments are much more regular, predictable and calculable. Here I show one example for each function. Refer to unified equation 12 on page 21 for calculation of $L$.

- $F(P_4, x)$: $M = 210$, $L = (3 - 1) \cdot (5 - 1) \cdot (7 - 1) = 48$, $S = \gcd(210, 48) = 6$. $[0, 210)$ can be divided into 6 segments of length 35, and each segment has exactly 8 zeros. When $x < 11^2$, they are all prime numbers excluding 1, so there are exactly $4 + 7 = 11$ primes $< 35$, 8 primes between 35 and 70, and 8 primes between 70 and 105. Refer to table on page 24.

- $H(P_5, x)$: $M = 2310$, $L = (3 - 2) \cdot (5 - 2) \cdot (7 - 2) \cdot (11 - 2) = 135$, $S = \gcd(2310, 135) = 15$. $[0, 210)$ can be divided into 15 segments of length 154, and each segment has about 9 zeros. For each segment $[154k, 154(k + 1))$, $k = 0, \cdots , 14$, the sequence of number of zeros is

$$10, 9, 9, 10, 8, 9, 10, 8, 9, 9, 8, 9, 8, 9, 10, 9, 9$$

Because $154 < 13^2$, each zero $x_0$ on $(0, 154)$ represents a pair of twin prime $(x_0 - 1, x_0 + 1)$. Therefore it has 9 twin primes not involving elements of $P_5$. Plus $(3, 5), (5, 7)$ and $(11, 13)$, There are totally 12 twin prime pairs less than 154.

- $G(60, x)$: $M = 210$, $L = (3 - 1) \cdot (5 - 1) \cdot (7 - 2) = 40$, $S = \gcd(210, 40) = 10$. $[0, 210)$ can be divided into 10 segments of length 21, and each segment has exactly 4 zeros. When $x < 60$, each zero $x_0$ represents a pair of primes $(60 - x_0, 60 + x_0)$ of sum 120. Therefore there are exactly 4 pairs on $(39, 81)$, and another 4 pairs on $(18, 39) \cup (81, 102)$. Refer to table on page 26.

Combine this conjecture and the conjecture 4.7 on page 13 that $\rho_2 \approx \rho_2^2$, we can understand more of the structure of prime, twin prime, Goldbach sum distributions and their relationship.
• Convert the Number Theory Theorems into expression of these series, and they may come together to generate some new results.

References