Prime Sums

Nir Elber, Raymond Feng, Henry Xie
Counselor: Anupam Datta

PROMYS 2021

Abstract

We study the properties of certain graphs involving the sums of primes. Their structure largely turns out to relate to the distribution of prime gaps and can be roughly seen in Cramér's model as well. We also discuss generalizations to the Gaussian integers.

Contents

1 Introduction

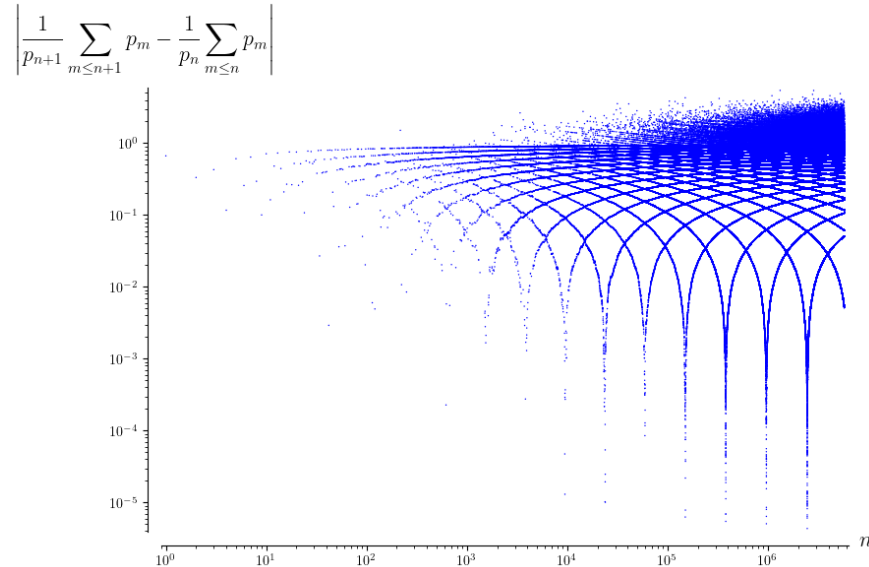The main motivation is why the following figure has so much structure.

Figure 1: Log-log plot of absolute consecutive differences of $a_n := \frac{1}{p_n} \sum_{m \leq n} p_m$

There are a number of features of the above figure which stand out here: the relatively constant horizontal line at the top which gives way to a more chaotic "cloud" of points, the "troughs" which appear to occur at somewhat regular intervals, the various other curves which seem to repeat themselves, and so on. Each of these aspects of the graph will be studied in their own right.

With this in mind, the layout of this paper is as follows. In section 2, we study various aspects of Figure 1. The highlights are that most of the repeating structure comes from prime gaps over various sizes, where the key lemma is as follows.

Proposition 1.1 (Key lemma). *We have that*

$$\lim_{n \to \infty} \frac{\log n}{p_n - p_{n+1}} \left(\frac{1}{p_{n+1}} \sum_{m \leq n} p_m - \frac{1}{p_n} \sum_{m \leq n} p_m \right) = \frac{1}{2}.$$

Relating Proposition 1.1 with Figure 1, lets us more finely discuss its structure. For example, we are able to prove the following.

Proposition 1.2. *As $n \to \infty$, the location of the troughs (provided the troughs exist) in Figure 7 occur at $n \approx e^k$ where k varies over positive integers.*

Remark 1.3. *It should be noted that this is simply stating where the troughs occur only when they exist; it is unknown whether there is a trough for every k (this requires guaranteeing the existence of a specific prime gap value around a certain region of primes), or whether any of the troughs extend infinitely to the right (de Polignac's conjecture [Pol49], if true, would imply this statement because every even integer would appear infinitely often as a prime gap).*

Then in section 3, we quickly compare Figure 1 with a variant of Cramér's model and find that, even though what we can prove about the error term is stronger with Cramér's model (in Proposition 3.1), in practice, the primes appear more structured.

Lastly, in section 4, we extend some of these ideas to the Gaussian integers. Even though studying even the main term of prime sums over $\mathbb{Z}[i]$ appears quite difficult when attempting to keep track of the angle, we are able to achieve something. Namely, we are able to show the following.

Theorem 1.4. *Fix n a nonzero integer. Then, for any real number x,*

$$\sum_{\mathrm{N}(\pi) \leq x} \frac{\pi^n}{\mathrm{N}(\pi)^{n/2}} = O\left(\frac{x}{(\log x)^A}\right),$$

where A is an arbitrarily large (but fixed) constant. Here, the sum is taken over Gaussian primes $\pi \in \mathbb{Z}[i]$ with $\mathrm{N}(\pi) \leq x$; the sum is a real number.

The ideas which go into the proof of Theorem 1.4 motivate a random model for the Gaussian primes as well, which we discuss at the end of the section.

2 Graphing Large Values

As promised, the main cause of the structure of Figure 1 is the following asymptotic.

Proposition 1.1 (Key lemma)**.** *We have that*

$$\lim_{n \to \infty} \frac{\log n}{p_n - p_{n+1}} \left(\frac{1}{p_{n+1}} \sum_{m \leq n} p_m - \frac{1}{p_n} \sum_{m \leq n} p_m\right) = \frac{1}{2}.$$

Proof. This is a culmination of some known results. After factoring out the sum and combining the fractions, we are computing

$$\lim_{n \to \infty} \frac{\log n}{p_n - p_{n+1}} \cdot \frac{p_n - p_{n+1}}{p_{n+1} p_n} \sum_{m \leq n} p_m.$$

Here, the (negative) prime gap $p_n - p_{n+1}$ will cancel out. It is also known that $\sum_{m \leq n} p_m \sim \frac{1}{2} n^2 \log n$ and that $p_n \sim p_{n+1} \sim n \log n$; plugging these all in gives

$$\lim_{n \to \infty} \frac{\log n}{1} \cdot \frac{1}{(n \log n)(n \log n)} \cdot \frac{1}{2} n^2 \log n = \frac{1}{2}.$$

This finishes the proof. $\blacksquare$

Remark 2.1. *The convergence of this function is actually very slow (see Figure 2), but the function itself appears quite smooth. We will discuss this below.*

To analyze the convergence of the above quantity, we prove the following proposition:

Proposition 2.2. *We actually have*

$$\frac{\log n}{p_n p_{n+1}} \sum_{m \leq n} p_m = \frac{1}{2} - \frac{\log \log n}{2 \log n} + \frac{1}{4 \log n} + o\left(\frac{1}{\log n}\right).$$

Proof. We use the bounds of

$$p_n = n(\log n + \log \log n - 1 + o(1))$$

and

$$\sum_{m \leq n} p_m = \frac{n^2}{2}\left(\log n + \log \log n - \frac{3}{2} + o(1)\right),$$

which can be found in [Sin15]. Note that

$$\log(n+1) + \log \log(n+1) - 1 + o(1) = \log n + \log\left(1 + \frac{1}{n}\right) + \log \log n + \log\left(\frac{\log(n+1)}{\log n}\right) - 1 + o(1)$$

$$= \log n + \log \log n - 1 + o(1).$$

3

Also, we have

$$\frac{\log n}{\log n + \log\log n - 1 + o(1)} = 1 - \frac{\log\log n - 1 + o(1)}{\log n + \log\log n - 1 + o(1)}$$

$$= 1 - \frac{\log\log n - 1}{\log n} + (\log\log n - 1) \cdot \left[\frac{1}{\log n} - \frac{1}{\log n + \log\log n - 1}\right]$$

$$+ o\left(\frac{1}{\log n}\right)$$

$$= 1 - \frac{\log\log n - 1}{\log n} + o\left(\frac{1}{\log n}\right).$$

Then we have that

$$\frac{\log n}{p_n p_{n+1}} \sum_{m \leq n} p_m = \frac{\log n}{(n+1)(\log n + \log\log n - 1 + o(1))} \cdot \frac{\frac{n^2}{2}\left(\log n + \log\log n - \frac{3}{2} + o(1)\right)}{n(\log n + \log\log n - 1 + o(1))}$$

$$= \frac{\log n}{(n+1)(\log n + \log\log n - 1 + o(1))} \cdot \frac{n}{2}$$

$$\cdot \left[1 - \frac{1}{2(\log n + \log\log n - 1 + o(1))} + o\left(\frac{1}{\log n}\right)\right]$$

$$= \left(1 - \frac{\log\log n - 1}{\log n} + o\left(\frac{1}{\log n}\right)\right)\left(\frac{1}{2} - \frac{1}{2n+2}\right) \cdot \left[1 - \frac{1}{2\log n} + o\left(\frac{1}{\log n}\right)\right]$$

$$= \frac{1}{2} - \frac{1}{4\log n} - \frac{\log\log n - 1}{2\log n} + o\left(\frac{1}{\log n}\right)$$

$$= \frac{1}{2} - \frac{\log\log n}{2\log n} + \frac{1}{4\log n} + o\left(\frac{1}{\log n}\right),$$
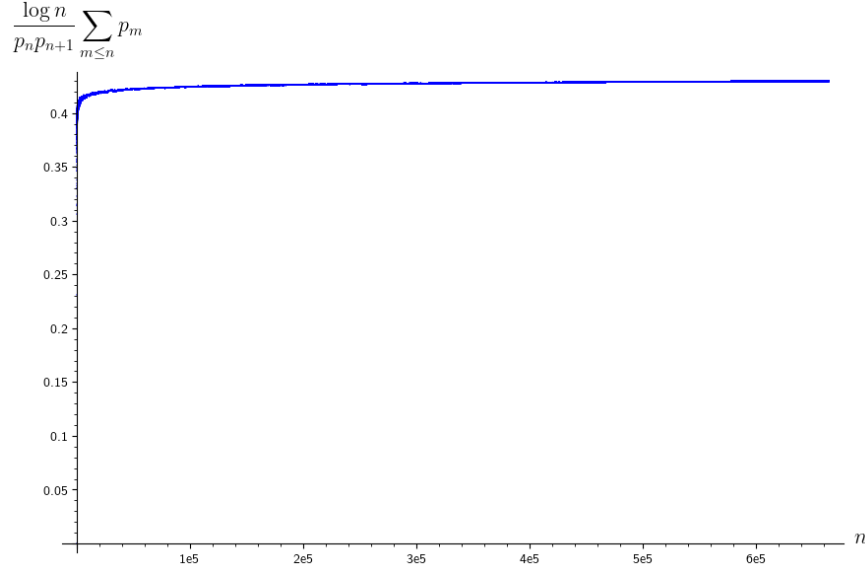
as claimed. ∎

Figure 2: Convergence of Proposition 1.1

We now begin to unravel Proposition 1.1 to talk about the structure of Figure 1. To begin, we see that

$$\log n \left(\frac{1}{p_{n+1}} \sum_{m \leq n} p_m - \frac{1}{p_n} \sum_{m \leq n} p_m\right) \sim \frac{p_n - p_{n+1}}{2},$$
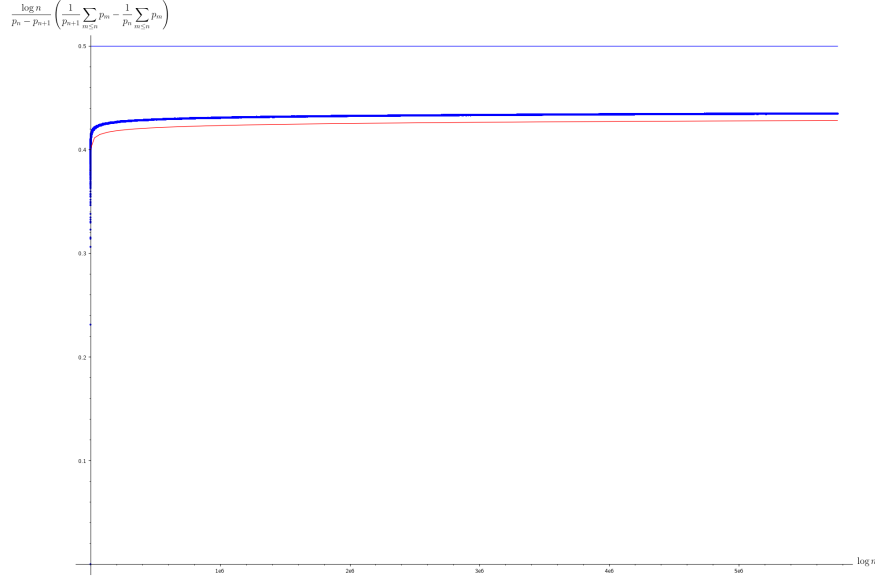
4

$$\frac{\log n}{p_n - p_{n+1}}\left(\frac{1}{p_{n+1}}\sum_{m\leq n}p_m - \frac{1}{p_n}\sum_{m\leq n}p_m\right)$$

Figure 3: Convergence of Proposition 1.1 with asymptotic estimate in red

so graphing $y = \log n \left(\frac{1}{p_{n+1}}\sum_{m\leq n}p_m - \frac{1}{p_n}\sum_{m\leq n}p_m\right)$ with respect to $x = n$ will look like descending horizontal lines, one corresponding to each prime gap. This won't change if we graph with respect to $x = \log n$; compare Figure 4.
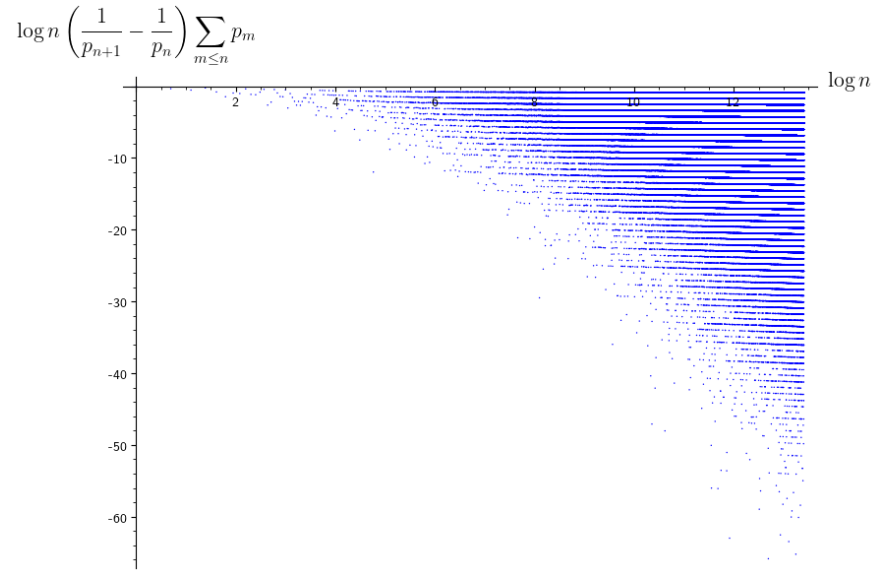

$$\log n \left(\frac{1}{p_{n+1}} - \frac{1}{p_n}\right)\sum_{m\leq n}p_m$$

Figure 4: Prime gaps

Continuing, we next take out the $\log n$ factor. Because $y = \log n \left(\frac{1}{p_{n+1}}\sum_{m\leq n}p_m - \frac{1}{p_n}\sum_{m\leq n}p_m\right)$ as a function of $x = \log n$ looks like horizontal lines, graphing $y = \frac{1}{p_{n+1}}\sum_{m\leq n}p_m - \frac{1}{p_n}\sum_{m\leq n}p_m$ with respect to $x = \log n$ will look like

$$xy \sim \frac{p_n - p_{n+1}}{2}.$$

5

Equivalently, this looks like

$$y \sim \frac{p_n - p_{n+1}}{2} \cdot \frac{1}{x}.$$

Namely, these are descending hyperbolas, one hyperbola for each prime gap. See Figure 5.
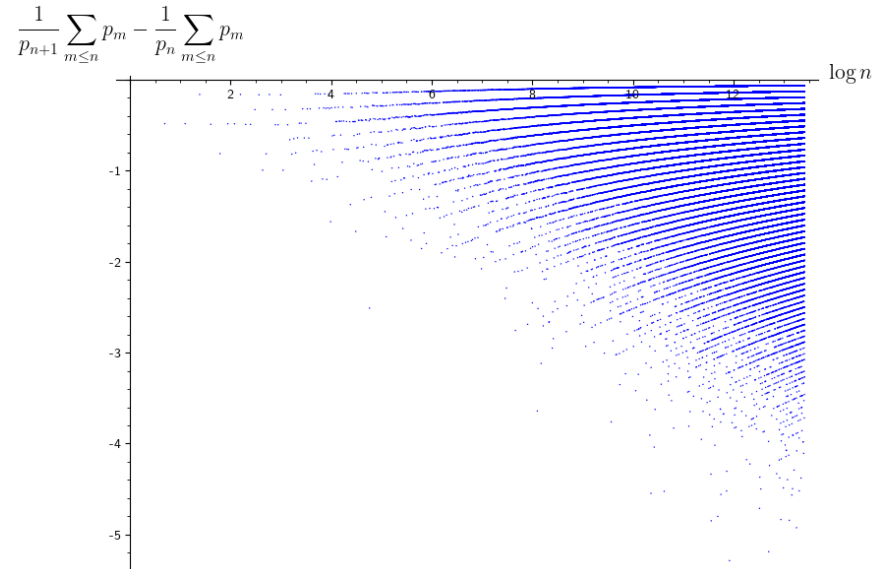
Figure 5: Hyperbolas

We now correct the index of $\frac{1}{p_{n+1}} \sum_{m \leq n} p_m$ by simply adding $1 = p_{n+1}/p_{n+1}$ to y. This means we want to graph $y = \frac{1}{p_{n+1}} \sum_{m \leq n+1} p_m - \frac{1}{p_n} \sum_{m \leq n} p_m$ as a function of $x = \log n$, which will look like

$$y = 1 + \frac{p_n - p_{n+1}}{2} \cdot \frac{1}{x}.$$

This shifts all hyperbolas up by 1, giving Figure 6.

Figure 6: Hyperbolas shifted up by 1, with negatives colored red

6

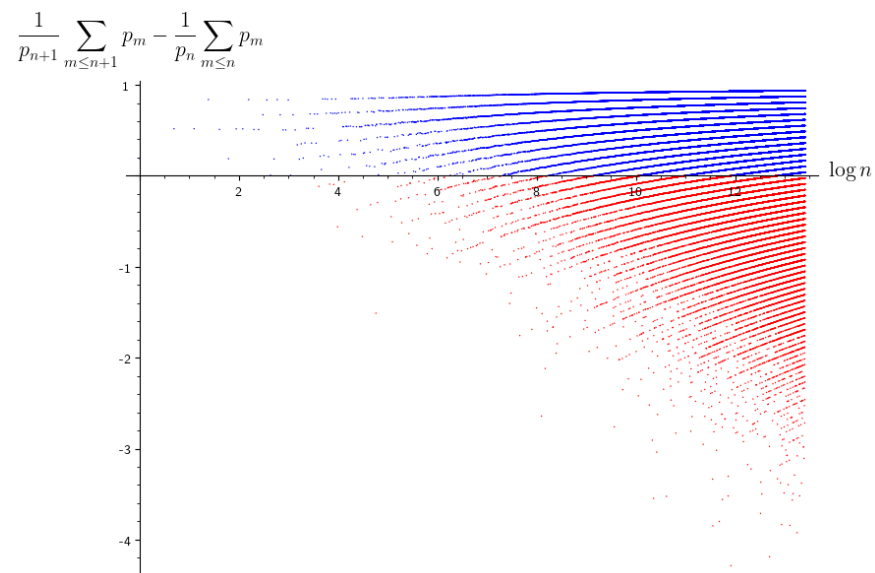Now, we see that a log-log plot of $\frac{1}{p_{n+1}} \sum_{m \leq n+1} p_m - \frac{1}{p_n} \sum_{m \leq n} p_m$ roughly corresponds to to graphing $y = \log \left| \frac{1}{p_{n+1}} \sum_{m \leq n+1} p_m - \frac{1}{p_n} \sum_{m \leq n} p_m \right|$ with respect to $x = \log n$. This will look like

$$y = \log \left| 1 + \frac{p_n - p_{n+1}}{2} \cdot \frac{1}{x} \right|.$$

So we get a "curtain" of $\log \left| 1 - \frac{k}{x} \right|$ for each positive integer k. (How clearly this current appears corresponds to how densely a particular prime gap appears.) This is actually what Figure 1. Figure 7 is Figure 1 with the dots colored by sign.
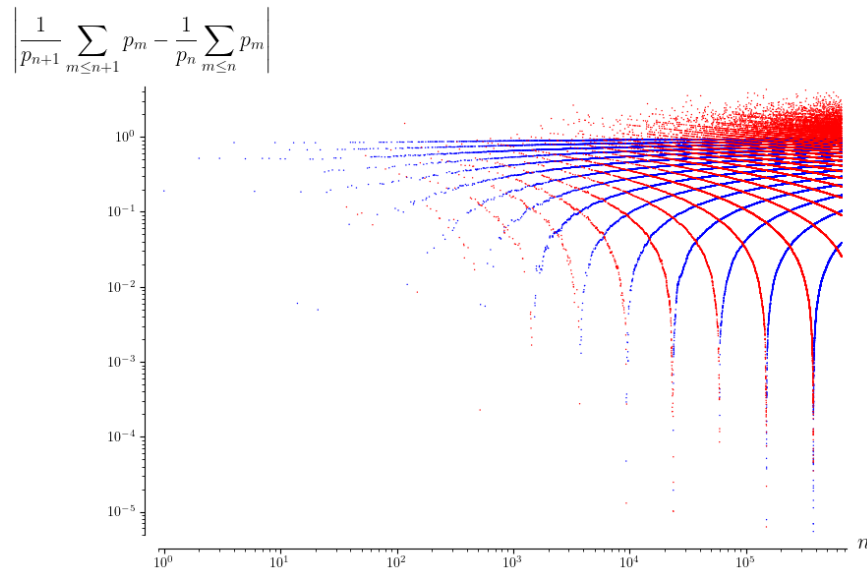

Figure 7: Log-log plot of $\left| \frac{1}{p_{n+1}} \sum_{m \leq n+1} p_m - \frac{1}{p_n} \sum_{m \leq n} p_m \right|$, with negatives red.

At this point, we can begin to tease out structure from the graph of Figure 7. For example, all of the positive differences (in blue), are converging to $10^0 = 1$ because the log-log plot should look like $\log \left| 1 - \frac{k}{x} \right|$ for various positive integers k, and as $x \to \infty$ in such a plot, the value of this goes to 0, which is $10^0 = 1$.

More interestingly, we can predict where the "troughs" should be as n grows large. This should occur, roughly speaking, when the graph of

$$y = \log \left| 1 - \frac{p_{n+1} - p_n}{2} \cdot \frac{1}{x} \right|$$

approaches $-\infty$. Namely, fixing a particular prime gap $g := p_{n+1} - p_n$, we want $1 - \frac{g/2}{x}$ to be roughly 0, which corresponds to $\log n = x \approx g/2$. Thus, we have the following proposition.

Proposition 1.2. *As $n \to \infty$, the location of the troughs (provided the troughs exist) in Figure 7 occur at $n \approx e^k$ where k varies over positive integers.*

Proof. This follows from the above discussion. ∎

We can also discuss the red points above Figure 7.

Proposition 2.3. *As $n \to \infty$, the red cluster of points above the y-axis in Figure 7 continues to grow without bound.*

Proof. Notice that

$$\left| \frac{1}{p_{n+1}} \sum_{m \leq n+1} p_m - \frac{1}{p_n} \sum_{m \leq n} p_m \right| = 1 + \frac{\sum_{m \leq n} p_m}{p_n p_{n+1}} \cdot (p_{n+1} - p_n) \sim \frac{p_{n+1} - p_n}{2} \cdot \frac{1}{\log n},$$

7

but it is a known result [Wes31] that there are arbitrarily large primes p such that

$$p_{n+1} - p_n > c \log n$$

for any constant c. In other words,

$$\limsup_{n \to \infty} \frac{p_{n+1} - p_n}{\log n} = \infty.$$

This implies that the difference attains arbitrarily large values, so the red cluster of points in Figure 7 has points of arbitrarily large y-value.

On a somewhat related note, this also means that there are arbitrarily low/negative values in Figure 6. ∎

There is also a related claim we can make about the largest values in Figure 6, which correspond to the smallest prime gaps. The largest values are always less than one, as

$$\frac{1}{p_{n+1}} \sum_{m \leq n+1} p_m - \frac{1}{p_n} \sum_{m \leq n} p_m = 1 - \frac{p_{n+1} - p_n}{p_{n+1} p_n} \cdot \sum_{m \leq n} p_n < 1.$$

On the other hand, we can show that there are values in Figure 6 that are arbitrarily close to 1, since it is known [GPY07] that

$$\liminf_{n \to \infty} \frac{p_{n+1} - p_n}{\log n} = 0.$$

In fact, much more is known; it is known now that $\liminf_{n \to \infty} |p_{n+1} - p_n|$ is finite [Zha14].

3   Comparing with Cramér's Model

In this section we study whether the structure in the graphs of the prime sums from earlier can be attributed to the relatively random behavior of the prime numbers, or if the structure is special to the prime numbers.

To this end, our investigation focuses on the following modification of Cramér's random model: let $p_1 = 3$, and let $i = 1$. For each odd number $k \geq 5$, starting at $k = 5$ and incrementing by two each step, we do both of the following with probability $\frac{2}{\log k}$ (the factor of two comes from the fact that we are only choosing odd numbers, as we want to preserve the fact that the difference between consecutive terms is even):

- Set $p_{i+1} = k$, and

- increment i by 1.

In this manner, we generate a sequence of odd numbers which have the same asymptotic distribution as the prime numbers (this can be proven by the prime number theorem because we know that $\pi(n) \sim \frac{n}{\log n}$); in other words, for any interval of length N, the expected number of p_i chosen by the procedure described above grows asymptotically at the same rate as the number of primes in the interval, as N grows arbitrarily large.

As an example to compare, Figure 8 is the corresponding version of Figure 7 for the random model. What is remarkable here is that Figure 8 is somehow less stable than the case for primes; this is not a fluke. For example, the error terms which we can theoretically prove from Cramér's model in Theorem A.2 also turn out to have incorrect constant factors in comparison to the actual primes.

The main point of this section is to establish the following result, in analogy with Proposition 2.2. We refer the reader to Appendix A for the proof of the relatively known computation Theorem A.2.

Proposition 3.1. *Fix $\varepsilon > 0$. Over all sequences $\mathcal{P} \subseteq \mathbb{Z}$ with probability weighted by Cramér's model, we order $\mathcal{P} = \{p_1, p_2, \ldots\}$ with $p_1 < p_2 < \cdots$. Then we have that*

$$\frac{\log n}{p_n p_{n+1}} \sum_{m \leq n} p_m = \frac{\log n}{\text{Li}^{-1}(n)^2} \int_2^{\text{Li}^{-1}(n)} \frac{t}{\log t} \, dt + O_\varepsilon \left(x^{-1/2+\varepsilon} \right)$$

holds almost surely. Here, Li refers to the logarithmic integral.

8

$$\left| \frac{1}{p_{n+1}} \sum_{m \le n+1} p_m - \frac{1}{p_n} \sum_{m \le n} p_m \right|$$
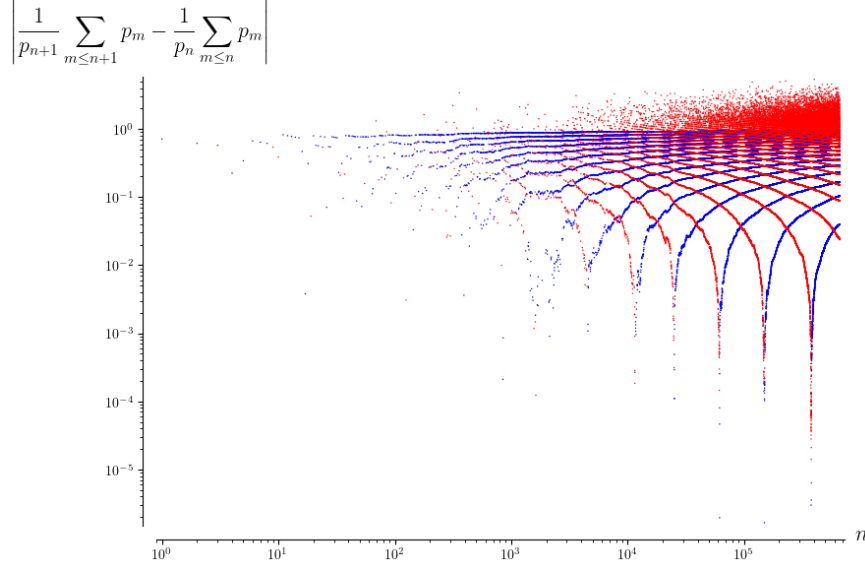
Figure 8: Randomized version of Figure 7

Proof. This proof is mostly done by force, converting each term in the sum to its asymptotic according to Theorem A.2. As such, we will split this proof into many parts.

Lemma 3.2. *Fix everything as above. Then* $p_n = \mathrm{Li}^{-1}(n) + O_\varepsilon\left(n^{1/2+\varepsilon}\right)$ *almost surely. In particular,* $p_n = O_\varepsilon\left(n^{1+\varepsilon}\right).$

Proof. Most of the work here is done by Theorem A.2, which after plugging in $x = p_n$ yields

$$\mathrm{Li}(p_n) = \int_2^{p_n} \frac{1}{\log t}\, dt = n + O_\varepsilon\left(p_n^{1/2+\varepsilon}\right) \tag{3.1}$$

with probability $1 - p_n^{-2}$. We would like to take Li^{-1} everywhere. This function is strictly increasing and, in fact, for real numbers $e < x < y$, we have

$$\mathrm{Li}(x) - \mathrm{Li}(y) = \int_y^x \frac{1}{\log t}\, dt \le x - y.$$

In particular, we see that $|\mathrm{Li}(x) - \mathrm{Li}(y)| < |x - y|$ for any real numbers $x, y > e$. Thus, we may use Equation 3.1 to write

$$\left| p_n - \mathrm{Li}^{-1}(n) \right| < |\mathrm{Li}(p_n) - n| = O_\varepsilon\left(p_n^{1/2+\varepsilon}\right).$$

So we see that $p_n = \mathrm{Li}^{-1}(n) + O_\varepsilon\left(p_n^{1/2+\varepsilon}\right).$

Now, $n \sim \mathrm{Li}(p_n)$ lets us get $p_n \sim n \log n$ in the same way that the Prime number theorem implies $p_n \sim n \log n$, so we get $p_n = O\left(n^{1+\varepsilon}\right)$, implying $p_n = \mathrm{Li}^{-1}(n) + O_\varepsilon\left(p_n^{1/2+\varepsilon}\right)$, where perhaps we have to switch out our constant factors. $\blacksquare$

The above lemma, among other things, gives us relatively easy access to the fraction $\frac{1}{p_n p_{n+1}}$.

Lemma 3.3. *Fix everything as above. We have that* $\frac{1}{p_{n+1}p_n} = \frac{1}{\mathrm{Li}^{-1}(n)^2} + O_\varepsilon\left(n^{-5/2+\varepsilon}\right)$ *almost surely.*

Proof. We begin by converting $1/(p_{n+1}p_n)$ into $1/p_n^2$. We see

$$\frac{1}{p_{n+1}p_n} = \frac{1}{p_n^2} - \frac{1}{p_n}\left(\frac{1}{p_n} - \frac{1}{p_{n+1}}\right) = \frac{1}{p_n^2} - \frac{p_{n+1} - p_n}{p_n^2 p_{n+1}}.$$

9

It is known [Cra36] that Cramér's model gives an estimate of $p_{n+1} - p_n = O\left((\log p_n)^2\right)$, so the error term here comes out to $O\left(p_n^{-3}(\log p_n)^2\right) = O_\varepsilon\left(n^{-3+\varepsilon}\right)$ by readjusting ε's constant factors as necessary.

We now convert $1/p_n^2$ to $1/\operatorname{Li}^{-1}(n)^2$. For this, we write

$$\left|\frac{1}{p_n^2} - \frac{1}{\operatorname{Li}^{-1}(n)^2}\right| = \left|\frac{1}{p_n} + \frac{1}{\operatorname{Li}^{-1}(n)}\right| \cdot \left|\frac{1}{p_n} - \frac{1}{\operatorname{Li}^{-1}(n)}\right|.$$

The first term is the sum of two fractions, both of whose denominators are bigger than n with probability $1 - n^{-2}$, so we may bounds this as $O(1/n)$. We can estimate the second term as

$$\frac{\left|\operatorname{Li}^{-1}(n) - p_n\right|}{p_n \operatorname{Li}^{-1}(n)}.$$

The numerator is $O_\varepsilon\left(n^{1/2+\varepsilon}\right)$ while the denominator is $O\left(1/n^2\right)$, so we accumulate $O_\varepsilon\left(n^{-3/2+\varepsilon}\right)$ here. Combining the two terms, we get $O_\varepsilon\left(n^{-1} \cdot n^{-3/2+\varepsilon}\right)$, which is what we wanted. ∎

Lastly, we turn the sum into an integral.

Lemma 3.4. *Fix everything as above. We have that*

$$\sum_{m \le n} p_m = \int_2^{\operatorname{Li}^{-1}(n)} \frac{t}{\log t}\, dt + O_\varepsilon\left(x^{3/2+\varepsilon}\right)$$

almost surely.

Proof. Plugging in p_n into Theorem A.2, we get that

$$\sum_{m \le n} p_m = \int_2^{p_n} \frac{t}{\log t}\, dt + O_\varepsilon\left(p_n^{3/2+\varepsilon}\right)$$

with probability $1 - p_n^{-2} > 1 - n^{-2}$. Plugging in our asymptotic $p_n = \operatorname{Li}^{-1}(n) + O_\varepsilon\left(n^{1/2+\varepsilon}\right)$ (which holds in the same situation by partial summation), we get

$$\int_2^{p_n} \frac{t}{\log t}\, dt = \int_2^{\operatorname{Li}^{-1}(n)} \frac{t}{\log t}\, dt + O_\varepsilon\left(n^{1/2+\varepsilon} \cdot \frac{p_n}{\log p_n}\right).$$

Now, the error term here will collapse into $O_\varepsilon\left(n^{3/2+\varepsilon}\right)$ once we use $p_n = O\left(n^{1+\varepsilon}\right)$ and readjust the constant factors. Combining our error terms finishes. ∎

We are now ready to prove the proposition. To begin, we write

$$\sum_{m \le n} p_m = \int_2^{\operatorname{Li}^{-1}(n)} \frac{t}{\log t}\, dt + O_\varepsilon\left(x^{3/2+\varepsilon}\right)$$

and multiply both sides by $1/(p_n p_{n+1}) = 1/\operatorname{Li}^{-1}(n)^2 + O_\varepsilon\left(n^{-5/2+\varepsilon}\right)$ to get

$$\frac{1}{p_n p_{n+1}} \sum_{m \le n} p_m = \frac{1}{\operatorname{Li}^{-1}(n)^2} \int_2^{\operatorname{Li}^{-1}(n)} \frac{t}{\log t}\, dt + O_\varepsilon\left(n^{-5/2+\varepsilon} \cdot \operatorname{Li}^{-1}(n)^2\right).$$

The error term here becomes $O_\varepsilon\left(n^{-1/2}\right)$ after adjusting constant factors, and we get the result after multiplying both sides by $\log n$. ∎

Remark 3.5. *One of the costs of the small error term $O_\varepsilon\left(x^{-1/2+\varepsilon}\right)$ is that the actual main term here is relatively unreadable. For example, it is no longer immediately obvious that the limit here is $1/2$.*

The point of the above result is that we are able to describe the behavior of the sum from Proposition 1.1 with very high accuracy—on the order of $n^{-1/2}$—using Cramér's model. This is more accuracy than we could previously achieve, yet its constant factors still look surprisingly off, as in Figure 8.

4   Generalization to Gaussian Integers

4.1   A Motivating Prime Sum

We now move the study of prime sums to primes in $\mathbb{Z}[i]$. One problem here is that there is no truly standard way to choose one prime while discarding its associates in the same way that we choose the positive primes in $\mathbb{Z}$. This is a real problem because it prevents us from, for example, usefully summing

$$\sum_{N(\pi) \leq x} \pi = 0,$$

where the sum is over irreducible elements $\pi \in \mathbb{Z}[i]$. Indeed, for every prime π with $N(\pi) \leq x$, we have that $-\pi$ is a distinct element which is also a prime of norm $\leq x$, so we in total sum to 0 after constructing pairs $(\pi, -\pi)$. In fact, we still can't even usefully sum

$$\sum_{N(\pi) \leq x} \pi^2 = 0$$

because, for each prime $\pi \in \mathbb{Z}[i]$ with $N(\pi)$, we see that $i\pi$ is another distinct prime of norm $\leq x$. In the same way as before, we can create pairs of primes $(\pi, i\pi)$ (even though $\pi \mapsto i\pi$ is not involutive, it is invertible), in total still summing to $\pi^2 + (i\pi)^2 = 0$.

   However, once we take fourth powers, then the sum becomes more interesting: the associates of a prime $\pi \in \mathbb{Z}[i]$ look like $\{\pi, i\pi, -\pi, -i\pi\}$, which all give the same fourth power, so there is no "trivial" cancellation. However, it turns out that there is still some (nontrivial) global cancellation, no matter what our exponent is. The main result of this section is to show the following.

Theorem 1.4. *Fix n a nonzero integer. Then, for any real number x,*

$$\sum_{N(\pi) \leq x} \frac{\pi^n}{N(\pi)^{n/2}} = O\left(\frac{x}{(\log x)^A}\right),$$

where A is an arbitrarily large (but fixed) constant. Here, the sum is taken over Gaussian primes $\pi \in \mathbb{Z}[i]$ with $N(\pi) \leq x$; the sum is a real number.

This statement essentially says that there is nontrivial cancellation among the $\pi^n / N(\pi)^{n/2}$. In particular, the prime number theorem for $\mathbb{Z}[i]$ implies that

$$\sum_{N(\pi) \leq x} 1 = \Theta\left(\frac{x}{\log x}\right),$$

so to get $o(x/\log x)$, there must be some cancellation.

   We have to build some machinery to prove Theorem 1.4. We take the following definition.

Definition 4.1 (Angle of a Gaussian integer)**.** *Let $\alpha = a + bi \in \mathbb{Z}[i] \setminus \{0\}$ be a nonzero Gaussian integer. Then we define the angle of α to be the angle $\theta_\alpha \in [0, 2\pi)$ such that $\alpha = \sqrt{N(\alpha)} e^{i\theta_\alpha}$.*

   The main intuition for Theorem 1.4 is that we are summing

$$\frac{\pi^n}{N(\pi)^{n/2}} = e^{ni\theta_\pi},$$

and because we expect θ_π to roughly be a random angle, we expect there to be a fairly large amount of cancellation in the "random walk." This intuition plus a technical result is convertible into the result.

Proof of Theorem 1.4. We begin by showing that the sum is actually real. Essentially, the idea is that, if π is a prime of bounded norm, then $\bar{\pi}$ also has the same bounded norm and will cancel out any imaginary contribution. To make this rigorous, we divide the sum up into the pieces

$$\sum_{N(\pi) \leq x} \frac{\pi^n}{N(\pi)^{n/2}} = \sum_{\substack{N(\pi) \leq x \\ \operatorname{Im} \pi > 0}} \frac{\pi^n}{N(\pi)^{n/2}} + \sum_{\substack{N(\pi) \leq x \\ \operatorname{Im} \pi = 0}} \frac{\pi^n}{N(\pi)^{n/2}} + \sum_{\substack{N(\pi) \leq x \\ \operatorname{Im} \pi < 0}} \frac{\pi^n}{N(\pi)^{n/2}}.$$

11

The middle sum is entirely real, so we won't pay more attention to it. Further, notice that for every prime π with $\operatorname{Im}\pi > 0$ and $N(\overline{\pi}) = N(\pi) \leq x$, we have $\overline{\pi}$ has $\operatorname{Im}\overline{\pi} < 0$ and still $N(\pi) \leq x$. This map lets us pair each π in the first sum with $\overline{\pi}$ living in the second. Thus,

$$\sum_{N(\pi)\leq x} \frac{\pi^n}{N(\pi)^{n/2}} = \sum_{\substack{N(\pi)\leq x \\ \operatorname{Im}\pi>0}} \left(\frac{\pi^n + \overline{\pi}^n}{N(\pi)^{n/2}}\right) + \sum_{\substack{N(\pi)\leq x \\ \operatorname{Im}\pi=0}} \frac{\pi^n}{N(\pi)^{n/2}}.$$

Now, $\pi^n + \overline{\pi}^n$ is self-conjugate and hence real. Thus, the entire sum is real, finishing this part of the proof.

We now turn to bounding the sum. As suggested by our intuition earlier, we write $\pi = \sqrt{N(\pi)}e^{i\theta_\pi}$ for each π, meaning that we want to bound

$$S := \sum_{N(\pi)\leq x} \frac{\pi^n}{N(\pi)^{n/2}} = \sum_{N(\pi)\leq x} e^{ni\theta_\pi}.$$

For peace of mind, we note that we already know this sum is real, so it suffices to take real parts and estimate

$$S = \sum_{N(\pi)\leq x} \cos(n\theta_\pi). \tag{4.1}$$

The idea, now, is to split up the various θ_π into small sectors and pretend that cos is constant over each sector. To make this work however, we need the technical Theorem 4 from [Kov76].

Lemma 4.2 (Koval'chik)**.** *Let x be a real number going to infinity, and let S be a sector of angle at least $\Theta^{-1/4+\varepsilon}$ for some $\varepsilon > 0$. Then*

$$\sum_{\substack{N(\pi)\leq x \\ \theta_\pi\in S}} 1 = \frac{\Theta}{2\pi} \cdot 8\operatorname{Li}(x)\left(1 + O\left(\frac{1}{(\log x)^B}\right)\right)$$

for any arbitrary large constant B.

We now set $N = \lfloor x^{1/3}\rfloor > 1$ and define the $|n|N$ sectors

$$S_k := \left[\frac{2\pi k}{|n|N}, \frac{2\pi(k+1)}{|n|N}\right),$$

where k varies from 1 to $|n|N$. We note that each sector has angle exceeding $x^{-1/4+\varepsilon_0}$, so these sectors are fair play for Lemma 4.2.

Anyways, we take Equation 4.1 and divide up by our sectors so that we are bounding

$$S = \sum_{k=1}^{|n|N} \sum_{\substack{N(\pi)\leq x \\ \theta_\pi\in S_k}} \cos(|n|\theta_\pi).$$

Now, fixing one sector S_k, we see $n\theta_\pi \in [2\pi k/N, 2\pi(k+1)/N)$, so we have that

$$\cos(n\theta_\pi) \in \left[\cos\frac{2\pi k}{N}, \cos\frac{2\pi(k+1)}{N}\right).$$

This means that our sum turns into

$$S = \sum_{k=1}^{|n|N} \sum_{\substack{N(\pi)\leq x \\ \theta_\pi\in \widetilde{S}_k}} \left[\cos\frac{2\pi k}{N} + O\left(\left|\cos\frac{2\pi(k+1)}{N} - \cos\frac{2\pi k}{N}\right|\right)\right]. \tag{4.2}$$

The error term here comes out to $-2\sin\left(\frac{\pi}{N}\right)\sin\left(\frac{\pi}{N}(2x+1)\right)$ using a sum-to-product formula; the estimate $\sin(\theta) < \theta$ for $\theta > 0$ is enough to conclude that the error is $O(1/N)$, which is good enough for our purposes.

Now, the terms in Equation 4.2 do not actually depend on the individual prime, so we may just count the number of primes by using Lemma 4.2 to get

$$S = \sum_{k=1}^{nN} \frac{8 \operatorname{Li}(x)}{nN} \left(1 + O\left(\frac{1}{(\log x)^A}\right)\right) \left(\cos \frac{2\pi k}{N} + O\left(\frac{1}{N}\right)\right),$$

where A is an arbitrarily large constant. Expanding this directly, we get that

$$S = \sum_{k=1}^{nN} \left[\frac{8 \operatorname{Li}(x)}{nN} \cos \frac{2\pi k}{N} + O\left(\frac{8 \operatorname{Li}(x)}{nN(\log x)^A} \cos \frac{2\pi k}{N}\right) + O\left(\frac{8 \operatorname{Li}(x)}{nN^2}\right) + O\left(\frac{8 \operatorname{Li}(x)}{nN^2(\log x)^A}\right)\right].$$

The first term will vanish when summed. As for the remaining error terms, we sum over k and are left with

$$S = O\left(\frac{\operatorname{Li}(x)}{(\log x)^A}\right) + O\left(\frac{\operatorname{Li}(x)}{N}\right) + O\left(\frac{\operatorname{Li}(x)}{N(\log x)^A}\right).$$

The new first error term is the main term to worry about because $N \sim x^{1/3}$ is large. This leave us with $O\left(x/(\log x)^{A+1}\right)$ because $\operatorname{Li}(x) \sim x/\log x$. We set $B = A + 1$ to finish the proof. ∎

Looking at the above proof, we note that there is really no need to restrict n to be a nonzero integer. Using the reframing with angles of Gaussian integers, the core of the above proof lies in the following result, which is proven in the same way as above.

Proposition 4.3. *Fix y a nonzero real number. Then, for any real number x,*

$$\sum_{N(\pi) \leq x} e^{iy\theta_\pi} = O\left(\frac{x}{(\log x)^A}\right),$$

where A is an arbitrarily large (but fixed) constant. Here, the sum is taken over Gaussian primes $\pi \in \mathbb{Z}[i]$ with $N(\pi) \leq x$; the sum is a real number.

Remark 4.4. *It is probably possible to generalize Proposition 4.3 further, allowing y to vary with x, though the authors have not worked this out explicitly.*

Remark 4.5. *Comparing Proposition 4.3 with Weyl's equidistribution criterion, it is unsurprising that we were forced to appeal to an equidistribution result like Lemma 4.2 in order to achieve Proposition 4.3.*

We close this section by actually computing the case of sums of fourth powers, which follows quickly from Theorem 1.4.

Corollary 4.6. *For real numbers x, we have that*

$$\sum_{N(\pi) \leq x} \pi^4 = O\left(\frac{x^3}{(\log x)^A}\right),$$

where A is an arbitrarily large constant.

Proof. By Theorem 1.4, we have that

$$\sum_{N(\pi) \leq x} \frac{\pi^4}{N(\pi)^2} = O\left(\frac{x}{(\log x)^A}\right).$$

In order to use partial summation, we rewrite the desired sum as

$$\sum_{2 \leq k \leq x} \left(\sum_{N(\pi) = k} \pi^4\right).$$

Partial summation turns this into

$$\sum_{2 \leq k \leq x} \left(\sum_{N(\pi) = k} \frac{\pi^4}{N(\pi)^2} \cdot k^2\right) = x^2 O\left(\frac{x}{(\log x)^A}\right) + O\left(\int_2^x \frac{t \cdot 2t}{(\log t)^A} \, dt\right),$$

for any arbitrarily large constant A. Both terms are $O\left(x^3/(\log x)^A\right)$, which finishes. (The second term is safe because the maximum value of the integral is $x^2/\log x$, for x sufficiently large.) ∎

4.2 A Random Model for Gaussian Integers

Motivated by the previous section, we now describe a random model for the Gaussian primes, akin to Cramér's model. We define our probability space to be over all subsets Π of $\mathbb{C}$ satisfying the following properties.

- We will fix $\pm 1 \pm i$ always live in Π.

- The set of norms $\mathrm{N}(\Pi) := \left\{ |z|^2 : z \in \Pi \right\}$ consists exclusively of integers, where an integer $n > 1$ has probability $1/(2 \log n)$ of appearing as a norm.

- For each norm $n \in \mathrm{N}(\Pi)$, there are eight elements of Π. An angle θ is chosen randomly from $[0, \pi/4)$, and then our eight elements are $\sqrt{n} e^{\pm i\theta}$ and their associates.

Note that we are choosing norms with probability $1/(2 \log n)$ because all but a density-zero set of the norms in $\mathbb{Z}[i]$ are 1 (mod 4) primes, which consist of roughly half (the density) of the rational primes; in other words, we have rigged the Prime number theorem for $\mathbb{Z}[i]$ into our norm distribution. Also, it is worth noting that the symmetry of associates and conjugates is built into the model.

In practice, the model behaves as two infinite tuples of random variables: the norms, which are a subset of $\mathbb{Z}$ chosen akin to a "sparse" Cramér's model; and the angles, which are essentially randomly chosen real numbers in $[0, \pi/4)$. As an example of what we can do, we have the following case of Proposition 4.3.

Proposition 4.7. *Over all subsets $\Pi \subseteq \mathbb{C}$ weighted according to the model above, we have that*

$$\sum_{\substack{\mathrm{N}(\pi) \leq x \\ \pi \in \Pi}} e^{4i\theta_\pi} = O\left(x^{1/2+\varepsilon} \right)$$

with probability $1 - x^{-2}$.

Proof. We begin by rewriting without the complex numbers. For each norm $n \in \mathrm{N}(\Pi)$, we let $\theta_n \in [0, \pi/4)$ be the corresponding angle. This makes the sum equal to

$$X := \sum_{\substack{n \in \mathrm{N}(\pi) \\ n \leq x}} 8 \cos(4\theta_n)$$

after accounting for associates and conjugates. At this point, the computation is essentially a random walk, but we will write it out for completeness; for example, there might be worries about the distribution of the norms, but this does not have an effect.

As mentioned above, the θ_ns are essentially random variables on their own, so we are able to immediately compute the expected value here as

$$\mathbb{E}[X] = \sum_{\substack{n \in \mathrm{N}(\pi) \\ n \leq x}} \mathbb{E}\left[8 \cos(4\theta_n) \right] = 0$$

by linearity. We now study higher moments in order to appeal to Chebychev's inequality, as in Theorem A.2; we refer to its proof for a more explicit computation. In particular, we look at the expansion of

$$\left(\sum_{\substack{n \in \mathrm{N}(\pi) \\ n \leq x}} \cos(4\theta_n) \right)^k$$

for an integer k to be fixed later. Because the θ_n are independent variables, the expected value here will vanish for all terms in the expansion, except for those which do not contain exactly one $\cos(4\theta_n)$ term for any particular n.

As in Theorem A.2, there are at most $O_k\left(x^{k/2}\right)$ of these terms to worry about (note that the number of norms is less than or equal to x), and each term is $O(1)$ at most, so we see

$$\mathbb{E}\left[X^k\right] = \mathbb{E}\left[\left|\sum_{\substack{n \in \mathrm{N}(\pi) \\ n \leq x}} \cos(4\theta_n)\right|^k\right] = O\left(x^{k/2}\right).$$

By Chebychev's inequality, we get that

$$\mathbb{P}\left[|X| \geq C_k^{1/k} x^\varepsilon \cdot x^{1/2}\right] \leq x^{-\varepsilon k},$$

for some constant C_k. Setting k large enough (with respect to ε) gets us $X = O\left(x^{1/2+\varepsilon}\right)$ with probability $1 - x^{-2}$. ∎

Remark 4.8. *Technically, we have not used the distribution of norms in the above proof, but this is roughly because the summands are not very sensitive to the size or number of the primes. In contrast, if we were just counting the number of primes, then this distribution of norms will matter.*

Note that the above proposition continues the paradigm that Cramér-like models roughly allow us to reduce the error term by a $1/2$ in the exponent. This error term also fits the data; see Figure 9 to compare the error term for the sum over the actual Gaussian primes.
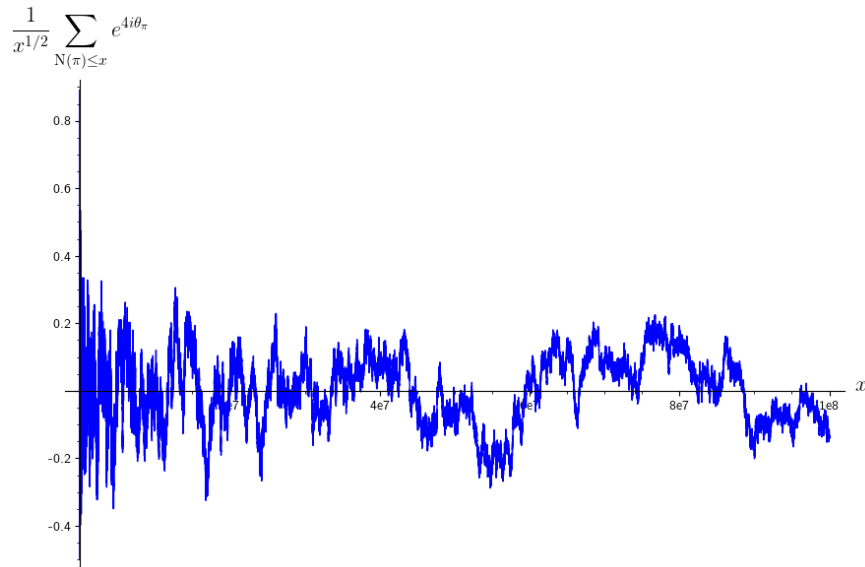

$$\frac{1}{x^{1/2}} \sum_{\mathrm{N}(\pi) \leq x} e^{4i\theta_\pi}$$

Figure 9: Sharpness of the error of Proposition 4.7

With this in mind, we make the following conjecture.

Conjecture 4.9. *Fix y a nonzero real number and $\varepsilon > 0$. Then, for any real number x,*

$$\sum_{\mathrm{N}(\pi) \leq x} e^{iy\theta_\pi} = O_\varepsilon\left(x^{1/2+\varepsilon}\right).$$

5 Acknowledgements

A   Estimates with Cramér's Model

Before doing any estimations with Cramér's model, we pick up the following technical lemma.

Lemma A.1. *Let α be a nonnegative real number. Then*

$$\sum_{3\leq n\leq x} \frac{n^\alpha}{\log n} = \int_2^x \frac{t^\alpha}{\log t}\,dt + O\left(\frac{x^\alpha}{\log x}\right).$$

Proof. The idea is to integrate by parts once to make the error terms easier to control in a continuous setting, and then integrate by parts back to get the statement. Set $f(t) := t^\alpha/\log t$ for brevity; it is continuously differentiable for $t > 1$. Using summation by parts, we see that

$$\sum_{3\leq n\leq x} f(n) = \lfloor x\rfloor f(x) - 2f(2) - \int_2^x \lfloor t\rfloor f'(t)\,dt.$$

Note that $\lfloor x\rfloor f(x) = xf(x) + O(f(x))$. As for the integral, we can express it as

$$\int_2^x \lfloor t\rfloor f'(t)\,dt = \int_2^x tf'(t)\,dt - \int_2^x \{t\}f'(t)\,dt,$$

and the right-hand integral is bounded by $\int_2^x f'(t)\,dt = O(f(x))$. Putting everything together, we see

$$\sum_{3\leq n\leq x} f(n) = xf(x) + O(f(x)) - 2f(2) - \int_2^x tf'(t)\,dt + O(f(x)).$$

The error terms combine to $O(f(x))$. The rest can be collapsed using integration by parts into

$$\sum_{3\leq n\leq x} f(n) = \int_2^x f(t)\,dt + O(f(x)).$$

This is exactly what we wanted. ∎

The main point of this section is to prove the following statement.

Theorem A.2. *Let α be a nonnegative real number. Then over all sequences $\mathcal{P} \subseteq \mathbb{Z}$ with probability weighted by Cramér's model, we have that*

$$\sum_{\substack{p\leq x\\ p\in\mathcal{P}}} p^\alpha = \int_2^x \frac{t^\alpha}{\log t}\,dt + O_\varepsilon\left(x^{\alpha+1/2+\varepsilon}\right),$$

with probability at least $1 - x^{-2}$.

Proof. This argument is not original. To use Cramér's model, we will have to look directly at individual integers, so we rewrite the sum as

$$\sum_{\substack{p\leq x\\ p\in\mathcal{P}}} p^\alpha = \sum_{n\leq x} n^\alpha 1_\mathcal{P}(n),$$

where $1_\mathcal{P}$ is the $\mathcal{P}$-indicator. This will be easier to analyze with an expected value of 0, so we remove the main term from the sum by writing

$$\sum_{n\leq x} n^\alpha 1_\mathcal{P}(n) = \sum_{3\leq n\leq x} \frac{n^\alpha}{\log n} - \sum_{3\leq n\leq x} n^\alpha\left(1_\mathcal{P}(n) - \frac{1}{\log n}\right). \tag{A.1}$$

The main term, which is the first sum on the right-hand side, is estimated using Lemma A.1 as

$$\sum_{3\leq n\leq x} \frac{n^\alpha}{\log n} = \int_2^x \frac{t^\alpha}{\log t}\,dt + O\left(\frac{x^\alpha}{\log x}\right).$$

16

This error term is $O\left(x^\alpha\right)$, so it is safe. It remains to bound the error of Equation A.1, for which we set the random variables $X_n := n^\alpha \left(1_{\mathcal{P}}(n) - \frac{1}{\log n}\right)$ for $n \geq 3$, and $X := \sum_{3 \leq n \leq x} X_n$. We see that it suffices for $X = O_\varepsilon\left(x^{\alpha+1/2+\varepsilon}\right)$ with probability $1 - x^{-2}$.

We would like to bound X. On one hand, we note $\mathbb{E}[X_n] = 0$, so linearity of expectation gives $\mathbb{E}[X] = 0$. However, we would like something more sophisticated, so we will use Chebychev's inequality, for which we need to study

$$\mathbb{E}\left[|X|^k\right] = \mathbb{E}\left[\left|\sum_{3 \leq n \leq x} X_n\right|^k\right],$$

for some positive integer k to be fixed later. After expanding out this sum and using linearity of expectation, the independence of the X_n will imply that many terms have expected value 0. In fact, for a term

$$X_1^{a_1} X_2^{a_2} \ldots X_{\lfloor x \rfloor}^{a_{\lfloor x \rfloor}} \text{ such that } a_1 + a_2 + \cdots + a_{\lfloor x \rfloor} = k$$

to not have expected value 0, each a_n must either be 0 or bigger than 1. In particular, each nonzero exponent is at least two, so there are at most $k/2$ exponents with nonzero entries; then there are fewer than $k^{k/2}$ ways to assign the exponents.

So in total, there are fewer than $k^{k/2}\binom{\lfloor x \rfloor}{k/2} \leq C_k x^{k/2}$ terms to worry about, for some constant C_k depending on k. Further, each term consists of k terms of $X_\bullet = n^\alpha \left(1_{\mathcal{P}}(n) - \frac{1}{\log n}\right) = O\left(x^\alpha\right)$. So we have, in total, is bounded by

$$\mathbb{E}\left[|X|^k\right] \leq C_k x^{k/2} \left(x^\alpha\right)^k,$$

where C_k is some constant depending on k (but not x).

We are now ready to use Chebychev's inequality. Setting $\varepsilon > 0$ small, we see that

$$\mathbb{P}\left[|X| \geq x^\varepsilon \mathbb{E}\left[|X|^k\right]^{1/k}\right] \leq x^{-\varepsilon k}.$$

Using our bound, we see that

$$\mathbb{P}\left[|X| \geq C_k^{1/k} x^{\alpha+1/2+\varepsilon}\right] \leq x^{-\varepsilon k}.$$

With $\varepsilon > 0$ fixed, we now set $k \geq 2/\varepsilon$ so that $X = O_\varepsilon\left(x^{\alpha+1/2+\varepsilon}\right)$ with probability $1 - x^{-2}$. (Note the implied constant $C_k^{1/k}$ is now dependent on ε.) This is what we wanted. ∎

Figure 10 and Figure 11 showcase the error term from Theorem A.2 on the primes. It appears that Cramér's model has successfully retrieved the correct magnitude of error term, though there are subtleties, such as the fact that the error largely looks negative.
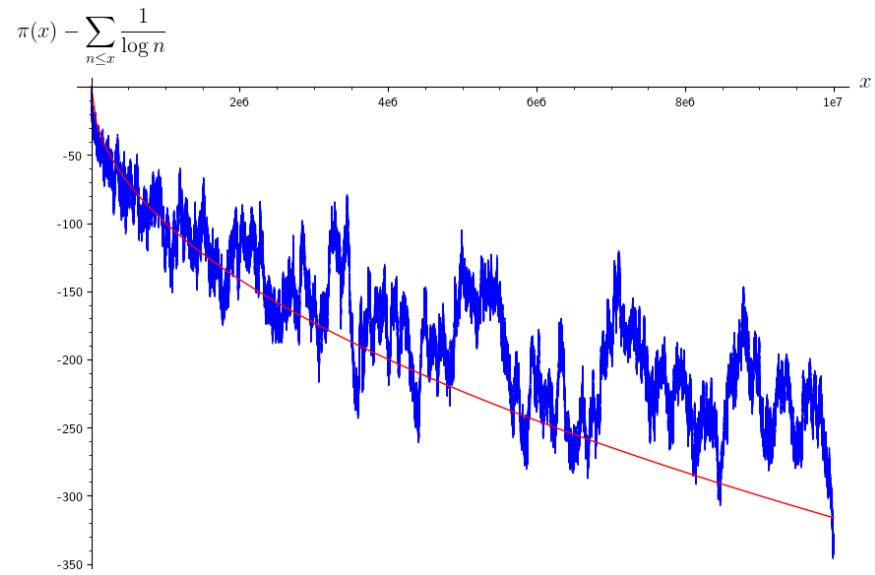
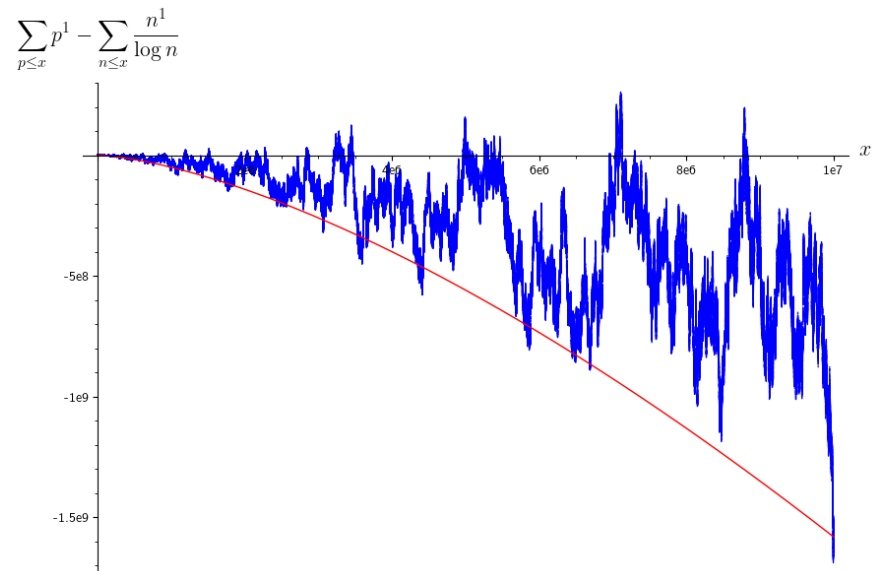Figure 10: Verification of the error bound in Theorem A.2; $y = -x^{1/2}/10$ in red

Figure 11: Verification of the error bound in Theorem A.2; $y = -x^{3/2}/20$ in red

References

[Pol49] Alphonse de Polignac. "Six propositions arithmologiques déduites du crible d'Ératosthène". fr. In: *Nouvelles annales de mathématiques : journal des candidats aux écoles polytechnique et normale* 1e série, 8 (1849), pp. 423–429. URL: http://www.numdam.org/item/NAM_1849_1_8__423_1/.

[Wes31] E. Westzynthius. *Über die Verteilung der Zahlen, die zu den n ersten Primzahlen teilerfremd sind.* German. Commentationes Helsingfors 5, Nr. 25, 37 S. (1931). 1931.

[Cra36] Harald Cramér. "On the order of magnitude of the difference between consecutive prime numbers". In: *Acta Arithmetica* 2.1 (1936), pp. 23–46. DOI: 10.4064/aa-2-1-23-46. URL: https://doi.org/10.4064/aa-2-1-23-46.

[Kov76] F. B. Koval'chik. "Density theorems for sectors and progressions". In: *Lithuanian Mathematical Journal* 15.4 (1976), pp. 618–631. DOI: 10.1007/bf00969793. URL: https://doi.org/10.1007/bf00969793.

[GPY07] D. A. Goldston, J. Pintz, and C. Y. Yildirim. *Primes in Tuples II.* 2007. arXiv: 0710.2728 [math.NT].

[Zha14] Yitang Zhang. "Bounded gaps between primes". English. In: *Ann. Math. (2)* 179.3 (2014), pp. 1121–1174. ISSN: 0003-486X. DOI: 10.4007/annals.2014.179.3.7.

[Sin15] Nilotpal Kanti Sinha. *On the asymptotic expansion of the sum of the first n primes.* 2015. arXiv: 1011.1667 [math.NT].