Estimating the Number of Primes In Unusual Domains

Johnathan Cai, Ryan Diehl, William Gasarch, Ian Kim, Rohan Sinha

October 16, 2025

Abstract

The Prime Number Theorem states that the number of primes in $\{1, \ldots, x\}$, denoted $\pi(x)$, is approximately $\frac{x}{\ln(x)}$. In this paper, we investigate the distribution of primes for domains other than $\mathbb{N}$. First we look at $\mathbb{A}_d = \{x \colon x \equiv 1 \pmod{d}\}$. We give a heuristic argument to form a conjecture on the number of *congruence monoid primes* in $\mathbb{A}_d$ that are $\leq x$. We then provide empirical evidence that indicates our conjecture is close but may need some correction. Second, we do similar calculations for the Gaussian Integers. Third, we discuss the difficulty of these types of questions for quadratic extensions of Z.

Keywords: Prime Number Theorem, Gaussian Integers, Circles, and Primes.

1 Introduction

In 1896, Jacques Hadamard and Charles-Jean de la Vallée Poussin both independently discovered that the number of primes $\leq x$ was roughly $\frac{x}{\ln(x)}$ using complex analysis. This was built off of earlier work done by Pafnuty Chebyshev in the 1850s. Later proofs have been created for the Prime Number Theorem by Paul Erdős and Atle Selberg in 1948. Their proof only uses calculus; however, they are still difficult. In 1980 Newman [1] (also see Zagier [3]) gave a short proof that only used a little complex analysis.

In Section 2, we define the *Congruence Monoid* and discuss an analog of the Prime Number Theorem for them. In Section 3, we define the *Integral Domain* and related notions. In Section 4, we define the Gaussian Integers and discuss analogs of the prime number theorem in them. In Section 5, we discuss other integral domains. In Section 6, we recap the open problems encountered.

2 Congruence Monoids

2.1 Definitions For Congruence Monoids

Def 2.1 [Prime] Let $\mathbb{D} \subseteq \mathbb{N}$. A *prime* in $\mathbb{D}$ is a number $p \in \mathbb{D}$ such that $p > 1$ and the only positive divisors of p that lie in $\mathbb{D}$ are 1 and p itself.

Def 2.2 [Congruence Monoid Prime] Let $d \in \mathbb{N}$ with $d \geq 2$. Define the set

$$\mathbb{A}_d = \{n \in \mathbb{N} \colon n \equiv 1 \pmod{d}\}.$$

An element $p \in \mathbb{A}_d$ is called a *congruence monoid prime* if $p \neq 1$ and, whenever $p = ab$ for some $a, b \in \mathbb{A}_d$, then either $a = 1$ or $b = 1$. Factorizations involving elements outside $\mathbb{A}_d$ are not considered.

Examples Lets look at $\mathbb{A}_4$.

1. We write down the first few elements:

$$1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41, 45$$

2. All of the numbers that are prime in $\mathbb{N}$ are prime in $\mathbb{A}_4$. So thats $5, 13, 17, 29, 37$.

3. What about the numbers that are not prime in $\mathbb{N}$?

   9 is prime. Even though $9 = 3 \times 3$ note that $3 \notin \mathbb{A}_4$.

   21 is prime. Even though $21 = 3 \times 7$, note that $e \notin \mathbb{A}_4$.

   25 is not prime since $25 = 5 \times 5$ and $5 \in \mathbb{A}_4$.

   33 is not primes since $33 = 3 \times 11$ and $3 \notin \mathbb{A}_4$.

   45 is not prime since $45 = 5 \times 9$ and $5, 9 \in \mathbb{A}_4$.

4. Note that there are numbers that are not primes in $\mathbb{N}$ but are primes in $\mathbb{A}_4$.

We want to define an analog of $\pi(x)$ for $\mathbb{A}_d$. We first carefully define $\pi(x)$.

Def 2.3 $\pi(x)$ is the number of elements of $\{1, \ldots, x\}$ that are prime, i.e.,

$$\pi(x) = \#\{x \in \{1, \ldots, x\} \colon p \text{ is prime}\}.$$

Note that the domain of interest is $\{1, \ldots, x\}$. We define $\pi_d(x)$, the analog of $\pi(x)$ for $\mathbb{A}_d$, noting that the domain of interest is $\mathbb{A}_d \cap \{1, \ldots, x\}$.

Def 2.4 [Congruence Monoid Prime Count] $\pi_d(x)$ is the number of elements of $\mathbb{A}_d \cap \{1, \ldots, x\}$ that are prime in $\mathbb{A}_d$, i.e.,

$$\pi_d(x) = \#\{x \in \mathbb{A}_d \cap \{1, \ldots, x\} \colon p \text{ is prime in } \mathbb{A}_d\}.$$

2.2 Congruence Monoid Prime Estimation

How does $\pi_d(x)$ compare to $\pi(x)$?

- The domain for $\pi_d(x)$ is smaller than that for $\pi(x)$. This suggests that $\pi_d(x) \leq \pi(x)$.
- There are primes in $\mathbb{A}_d$ that are not primes in $\mathbb{N}$. Hence this suggests that $\pi(x) \leq \pi_d(x)$.

Balancing these effects, we propose the estimation

$$\pi_d(x) \approx \frac{x}{d \ln(x)^{1/d}}.$$

Our empirical results from computational simulations support this approximation, and its accuracy can be quantified by the normalized ratio

$$R_d(x) = \frac{\pi_d(x)}{x/\big(d \ln(x)^{1/d}\big)}.$$

Values $R_d(x) \approx 1$ indicate close agreement between the model and observed data.

The table below displays the accuracy of the estimation across several values of d for primes $\leq x = 10^4$. The mean absolute percentage deviation (MAPE) measures the accuracy of a forecasting model, which in this case is our estimation. The corresponding graphs provide a visual comparison of $\pi_d(x)$ and the estimation over the full range of x for a few values of d from the table. Complete data can be viewed here.

d	Largest Prime	Actual Count	Estimate	R_d	$\lvert R_d - 1 \rvert$	MAPE (%)
3	10000	1380	1590.21	0.86781	0.13219	9.05
5	9996	1210	1282.34	0.94358	0.05642	3.81
7	9997	1009	1039.97	0.97022	0.02978	2.45
9	10000	851	868.19	0.98020	0.01980	2.28
11	10000	745	742.93	1.00279	0.00279	2.88
13	9998	653	648.33	1.00720	0.00720	3.10
21	9997	438	428.29	1.02268	0.02268	3.88
50	9951	196	190.38	1.02953	0.02953	3.03

Table 1: Comparison of Actual and Estimated D_d-Prime Counts up to 10^4

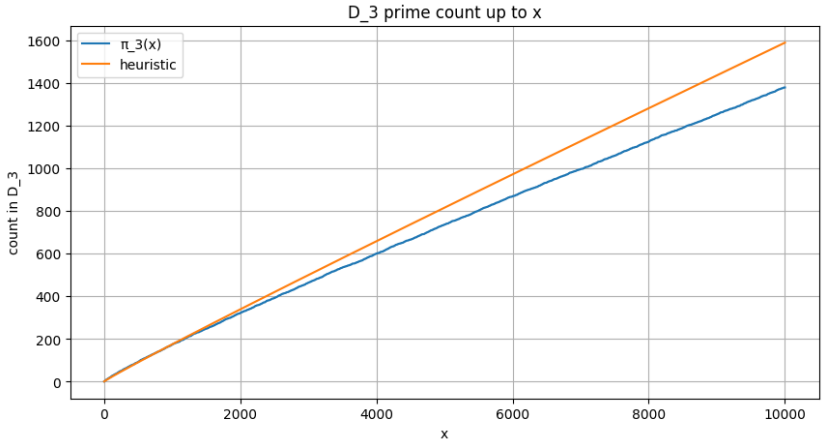

Figure 1: Actual prime count $\pi_3(x)$ (blue) versus the estimate (orange) for $d = 3$.

The estimation very minorly underestimates until $x \approx 800$, then increasingly starts to overestimate.

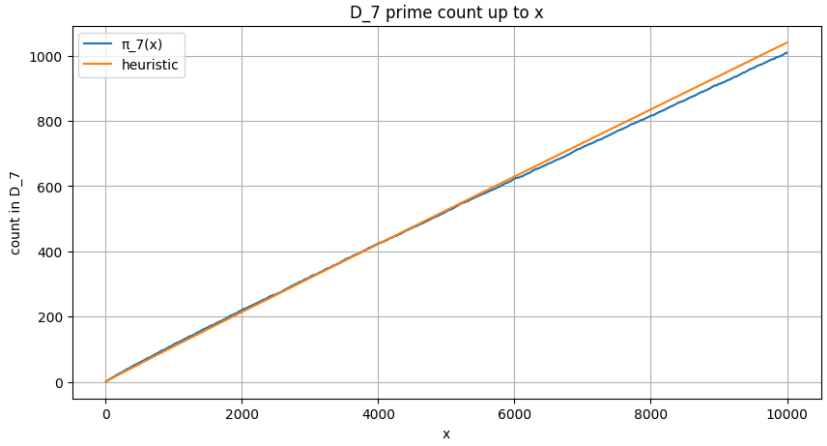

Figure 2: Actual prime count $\pi_7(x)$ (blue) versus the estimate (orange) for $d = 7$.

Similar to figure 1, the estimate very minorly underestimates until $x \approx 4100$ and then increasingly starts to overestimate.

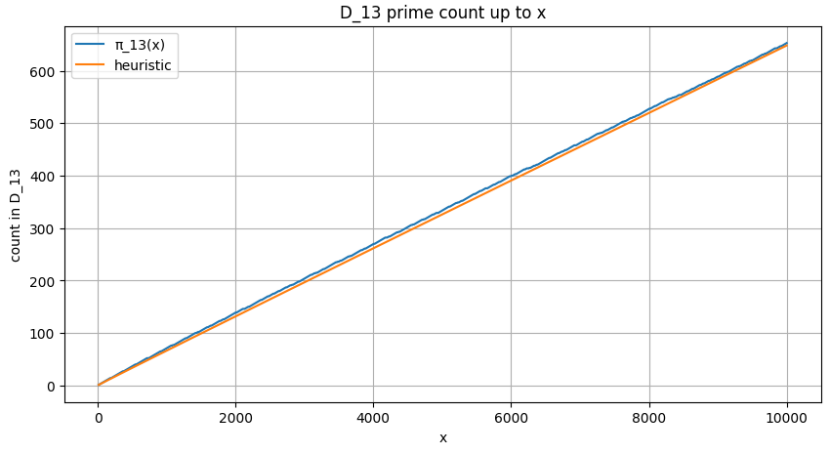

Figure 3: Actual prime count $\pi_{13}(x)$ (blue) versus the estimate (orange) for $d = 13$.

Same phenomenon described above except at $x \approx 13800$ (not shown in graph).

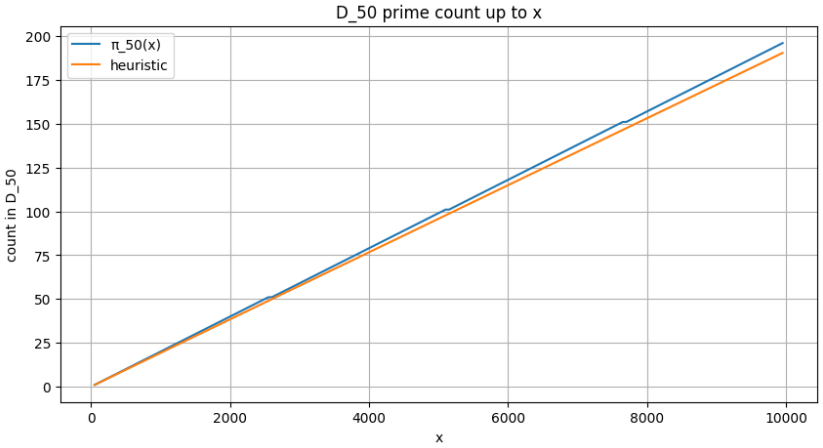

Figure 4: Actual prime count $\pi_{50}(x)$ (blue) versus the estimate (orange) for $d = 50$.

Same phenomenon described above except at $x \approx 330,000$ (not shown in graph).

Unlike the Prime Number Theorem, where the estimation becomes more accurate as x grows (in fact, becomes perfect as x approaches infinity), the accuracy of our estimation is not as straightforward. The estimation becomes increasingly accurate up to a certain value of x (almost exact) and then diverges from the actual count.

Our conjecture seems to be, as they used to say, *close but no cigar* (note that none of the authors smoke and three of them are high school students).

Open Problem 2.5

1. Modify our conjecture so that it better fits the data.

2. Prove that modification. Since the Prime Number Theorem was hard to prove this might be hard as well; however, the math needed for the Prime Number Theorem is known and may be useful.

3 Integral Domains

When studying primes in integral domains careful definitions of *unit*, *prime*, and *irreducible* are needed.

Def 3.1 An integral domain is a triple $\mathbb{D} = (D, +, \times)$ such that (1) D is a set, (2) $+$ and $\times$ are maps from $D \times D$ into D, and (3) $(D, +, \times)$ satisfies the following properties:

1. D is closed under $+$ and $\times$.
2. $+$ and $\times$ are both commutative and associative.
3. $\times$ is distributive over $+$. Hence, for all $a, b, c \in D$, $a \times (b + c) = a \times b + a \times c$.
4. There exists an element $0 \in D$ such that, for all $a \in D$, $a+0 = 0+a = a$.
5. For every $a \in D$ there exists $b \in D$ such that $a + b = 0$. We usually denote b by $-a$.
6. There exists an element $1 \in D$ such that, for all $a \in D$, $a \times 1 = a$.
7. For all $a, b \in D$, if $ab = 0$ then either $a = 0$ or $b = 0$.

Def 3.2 Let $\mathbb{D}$ be an integral domain.

1. An element $u \in \mathbb{D}$ is a *unit* there exists $v \in \mathbb{D}$ such that $uv = 1$. The units of Z are $\{-1, 1\}$.
2. An element $p \in \mathbb{D}$ is *prime (in $\mathbb{D}$)* if (a) p is not a unit, and (b) if p divides ab then either p divides a or p divides b. The primes of Z are the usual primes and their negations. For example, 3 and -7 are both primes in Z.
3. An element $r \in \mathbb{D}$ is *irreducible (in $\mathbb{D}$)* if (a) r is not a unit, and (b) if $r = ab$, then either a or b is a unit. In Z irreducibles and primes are the same. This equivalence is false in some other integral domains.

Def 3.3 Let $\alpha \notin \mathbb{Q}$. Then $\mathsf{Z}[\alpha]$ is the set $\{a + b\alpha \colon a, b \in \mathsf{Z}\}$

4 Gaussian Prime Estimation

Def 4.1 The *Gaussian Integers* is the set $\mathsf{Z}[i]$ where $i = \sqrt{-1}$.

The Gaussian Integers are an integral domain. Hence the definitions of units, primes, and irreducible from Definition 3.2 apply to them. In the Gaussian integers (a) the units are $\{1, -1, i, -i\}$, and (b) primes and irreducibles are the same.

We want to study an analog of the Prime Number Theorem for the Gaussian Integers. Recall that the usual Prime Number Theorem is about the number of primes in $\{1, \ldots, x\}$. Since $\mathbb{N}$ is not an integral domain, but Z is, lets rewrite that as the prime number theorem taking into account that the domain is Z:

The number of primes in $\{y \in \mathsf{Z} \colon 0 \leq |y| \leq x\}$ *is approximately* $\frac{x}{\ln(x)}$. We take $0 \leq |y|$ since in the usual Prime Number Theorem we do not count both a prime and its negation. For example, we don't count both 7 and -7 as primes.

For the Gaussian Integers we need (1) a notion of size analogous to absolute value for Z, (2) a way to not count p, $-p$, ip, and $-ip$.

Def 4.2

1. Let $a, b \in \mathsf{Z}$. Then the *norm* of $a + bi$ is $a^2 + b^2$. This is denoted by $N(a + bi)$. This will not be our notion of size; however, $\sqrt{N(a + bi)}$ will be.

2. Let $r \in \mathbb{N}$. The *norm circle* of radius r in the set

$$\{a + bi \colon \sqrt{a^2 + b^2} \leq r\}.$$

Using the norm circle provides a clear and finite boundary, making it possible to study the distribution of Gaussian primes up to a specific size, analogous to counting up to x in the Prime Number Theorem. There is one more issue: we only look at $a, b \geq 0$ since that avoids the problem of counting a prime four times because of units.

Def 4.3 $\pi_G(r)$ is the number of primes in $a + bi \in \mathsf{Z}[i]$ such that (a) $a, b \geq 0$, and (b) $a + bi$ is in the norm circle of radius r.

We empirically found the estimate $\pi_G(r) \approx \frac{r^2}{2 \ln r}$

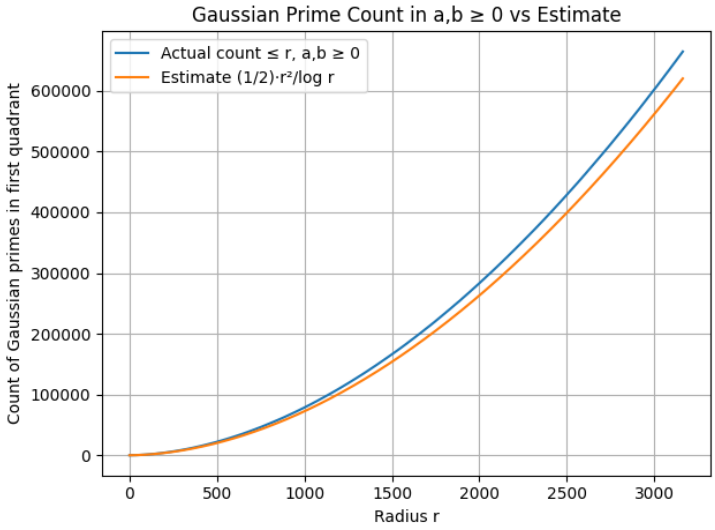

Figure 5: Actual Gaussian prime count for $0 < r^2 \leq 10^7$ (blue) versus the the estimate (orange).

Radius (r)	MAPE (%)
10^3	18.630
10^4	14.259
10^5	10.907
10^6	8.695
10^7	7.220

Table 2: Mean absolute percentage deviation (MAPE) of the estimate $\pi_G(r)$ at various radii.

Open Problem 4.4 Prove or disprove the estimate $\pi_G(r) \sim \frac{r^2}{2 \ln r}$. If the estimate is true then find the error term.

5 Other Integral Domains

We list integral domains and discuss if analogs of the Prime Number Theorem can be asked and what issues might arise. Much of the information in this

section is from Weintraub [2].
1) $\mathsf{Z}[\sqrt{-d}]$ where $d \in \mathbb{N}$ is squarefree. The standard norm is $N(a+b\sqrt{-d}) = a^2 + db^2$. There are two possible questions to ask:

- Approximate the number of primes of the form $a+b\sqrt{-d}$ where $a, b \geq 0$ and $a^2 + db^2 \leq r$, or
- Approximate the number of primes of the form $a+b\sqrt{-d}$ where $a, b \geq 0$ and $a^2 + b^2 \leq r$.

2) $\mathsf{Z}[\sqrt{d}]$ where $d \in \mathbb{N}$ is squarefree. Asking about primes in this integral domain is problematic for several reasons.
2a) For some values of d this is the wrong question.

In Section 4 we looked at $\mathsf{Z}[i]$. Why not $\mathsf{Z}[\frac{i}{2}]$? Lets start with $\mathbb{Q}(i)$. We look for an integral domain $\mathbb{D}$ such that $\mathsf{Z}[i] \subseteq \mathbb{D} \subseteq \mathbb{Q}(i)$ and (roughly): $\mathbb{Q}$ is to Z as $\mathbb{Q}(i)$ is to $\mathbb{D}$.

Def 5.1 Let $\mathbb{D}_1$ and $\mathbb{D}_2$ be integral domains such that $\mathbb{D}_1 \subseteq \mathbb{D}_2$.

1. $x \in \mathbb{D}_2$ is *integral over* $\mathbb{D}_1$ if x is the root of a monic polynomial with coefficients in $\mathbb{D}_1$.
2. The set of elements of $\mathbb{D}_2$ that are integral over $\mathbb{D}_1$ is the *integral closure of* $\mathbb{D}_1$ *in* $\mathbb{D}_2$.

Example 5.2

1. Z is the integral closure of Z in $\mathbb{Q}$.
2. $\mathsf{Z}[i]$ is the integral closure of Z in $\mathbb{Q}[i]$.
3. $\mathsf{Z}[\sqrt{5}]$ is *not* the integral closure of Z in $\mathbb{Q}[\sqrt{5}]$: $\frac{1+\sqrt{5}}{2}$ is integral over $\mathbb{D}$—it satisfies $x^2 - x - 1 = 0$.
4. The integral closure of Z in $\mathbb{Q}[\sqrt{5}]$ is $\mathsf{Z}[\frac{1+\sqrt{5}}{2}]$.
5. Let d be a square free integer.
   (a) If $d \equiv 2, 3 \pmod 4$ then the integral closure of Z in $\mathbb{Q}(\sqrt{d})$ is $\mathsf{Z}[\sqrt{d}]$.

(b) If $d \equiv 1 \pmod 4$ then the integral closure of Z in $\mathbb{Q}(\sqrt{d})$ is $\mathsf{Z}[\frac{1+\sqrt{5}}{2}]$.

Def 5.3 Let d be a squarefree integer. We define $\mathbb{O}(\sqrt{d})$ as follows:

1. If $d \equiv 2, 3 \pmod 4$ then $\mathbb{O}(\sqrt{d}) = \mathsf{Z}[\sqrt{d}]$.
2. If $d \equiv 1 \pmod 4$ then $\mathbb{O}(\sqrt{d}) = \mathsf{Z}[\frac{\sqrt{d}+1}{2}]$.

Here is the right question to ask:
is there an analog of the Prime Number Theorem for $\mathbb{O}(d)$?

2c) Number of Units.

1. If $d \geq 1$ then $\mathbb{O}(\sqrt{d})$ has an infinite number of units. This makes it hard to phrase an analog of the Prime Number Theorem.
2. $\mathbb{O}(\sqrt{-1})$ has four units (± 1, $\pm i$). $\mathbb{O}(\sqrt{-3})$ has six units (the six roots of unity). For all squarefree naturals $d \geq 5$, $\mathbb{O}(\sqrt{-d})$ has two roots of unit (± 1). Hence for these an analog of the Prime Number Theorem may be possible. But see the next item.

2d) For some $\mathbb{O}(\sqrt{d})$ primes and irreducibles are not the same. This may cause problems.

2e) For some d $\mathbb{O}(\sqrt{d})$ is not a unique factorization domain. This may cause problems.

We are *not* saying that formulating an analog of the Prime Number Theorem in $\mathbb{O}(\sqrt{d})$ is impossible; however, there are some difficulties to overcome.

3) What about adding cube-roots or higher fractional powers? What about adding more irrationals? These get into issues far harder than those encountered for quadratic extension.

6 Open Problems

We recap the open problems stated earlier.

1. Prove or disprove the suggested estimate for the number of primes in A_d. If proven then obtain an error term. If disproven the find the correct approximation.

2. Prove or disprove the suggested estimate for the number of primes in $\mathsf{Z}[i]$. If proven then obtain an error term. If disproven the find the correct approximation.

3. Formulate analogs of the Prime Number Theorem in $\mathsf{Z}[\sqrt{d}]$ for various values of d. Get empirical evidence to formulate conjectures.

References

[1] D. Newman. Simple analytic proof of the prime number theorem. *American Mathematical Monthly*, 87:693–696, 1980.

[2] S. Weintraub. *Factorization: Unique and otherwise*. CRC Press, 2008.

[3] D. Zagier. Newman's short proof of the prime number theorem. *American Mathematical Monthly*, 104:705–708, 1997.