Primes with restricted digits in progressions

Kunjakanan Nath

Université de Montréal

Online Seminar, 30 October 2020

Outline

Distribution of prime numbers I

Question

How many prime numbers are there? Are there finitely or infinitely many primes? If infinite, how many primes are there up to x?

- Euclid (300 BC): There are infinitely many primes. Consider $N = p_1 \dots p_k + 1$ and use the Fundamental Theorem of Arithmetic.
- Euler (1737): $\sum_p 1/p$ diverges.
- Gauss/Legendre (19th century) conjectured that

$$\pi(x) := \{p \leq x\} \sim \mathrm{Li}(x) := \int_2^x \frac{dt}{\log t}.$$

- Chebyshev (1850) proved that there exist two positive constants C_1 and C_2 such that

$$C_1 \frac{x}{\log x} \leq \pi(x) \leq C_2 \frac{x}{\log x}.$$

Distribution of prime numbers II

- Riemann (1857) introduced the Riemann zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

$$\{\text{Distribution of primes}\} \longleftrightarrow \{\text{the zeros of the } \zeta(s)\}$$

- Hadamard and de la Valle Poussin (1896) independently proved the prime number theorem (PNT) building on the ideas of Riemann. They showed that

$$\pi(x) \sim \int_2^x \frac{dt}{\log t}.$$

Remark

The probability that a random integer $n \in (x, 2x]$ is prime is roughly $1/\log x$.

Distribution of prime numbers in arithmetic progressions I

Question

How many prime numbers are there of the form $4k+1$ and $4k+3$? In general, if $(c,d)=1$, are there infinitely many primes of the form $c+dk$?

- Dirichlet (1837) showed that there are infinitely many primes of the form $c+dk$ whenever $(c,d)=1$. He introduced the Dirichlet characters to prove this result.
- More generally, for $(c,d)=1$ we set

$$\pi(x;d,c) = \#\{p \le x : p \equiv c \mod d\}.$$

Can we expect

$$\pi(x;d,c) \sim \frac{\pi(x)}{\varphi(d)}?$$

Distribution of prime numbers in arithmetic progressions II

- Prime number theorem in arithmetic progression (Siegel-Walfisz theorem): There exists a constant $A > 0$ such that for $d \leq (\log x)^A$, we have
$$\pi(x; d, c) \sim \frac{\pi(x)}{\varphi(d)}.$$
- Can we extend the modulus d in some larger range? What do we expect? ($d \leq x^{1-\epsilon}$)
- Generalized Riemann Hypothesis (GRH) implies $d \leq x^{1/2-\epsilon}$.
- Bombieri-Vinogradov theorem implies $d \leq x^{1/2-\epsilon}$ on an average over d. This is as strong as GRH for applications in analytic number theory.

Distribution of prime numbers in arithmetic progressions III

Question

What does average over d mean?

There exist constants $A, C > 0$ such that

$$\sum_{d \leq D} \max_{(c,d)=1} \left| \pi(x; d, c) - \frac{\pi(x)}{\varphi(d)} \right| \leq C \frac{x}{(\log x)^A},$$

for $D \leq x^{1/2-\epsilon}$.

- BV theorem implies that for most $d \leq x^{1/2-\epsilon}$ we have $\pi(x; d, c) \sim \pi(x)/\varphi(d)$.
- D is called the level of distribution. $D \leq x^{1-\epsilon}$ (Elliot-Halberstam conjecture).
- Applications: Recent work of Maynard and Tao (2013) on bounded gaps between primes.

Primes with restricted digits I

Question

Given any subset $\mathcal{B}$ of the natural numbers, what can we say about the distribution of primes in the set $\mathcal{B}$?

- In general, it is too difficult. For example, take the set $\{n^2+1 : n \in \mathbb{N}\}$ (Open problem).
- How about prime numbers with no digit 3 in its decimal expansion? For example, $2, 5, 7, 11, 17, 19, 29, 41, 47, 59, 61, 67, \dots$
- In general, for any integer $b \geq 2$ and $a_0 \in \{0, 1, 2 \dots, b-1\}$, set

$$\mathcal{A} = \Big\{ \sum_{j \geq 0} n_j b^j : n_j \in \{0, 1, \dots, b-1\} \setminus \{a_0\} \Big\}.$$

  Note that $\mathcal{A}$ is the set of integers with no digit a_0 in its b-adic expansion. In particular, if we set $b = 10$ we get our decimal representation.

Primes with restricted digits II

- Let us start with some heuristics for the simple case with $b = 10$ and $a_0 = 3$ in $\mathcal{A}$. Set $x = 10^k$, where k is an integer and $k \to \infty$. Then for any, $n < x = 10^k$, we have the following decimal representation

$$n = 10^{k-1} n_{k-1} + 10^{k-2} n_{k-2} + \ldots + 10 n_1 + n_0,$$

  where each $n_j \in \{0, 1, \ldots, 9\}$.
- If $n \in \mathcal{A}$ then $n_j \in \{0, 1, \ldots, 9\} \setminus \{3\}$.
- Each n_j has roughly 9 choices.
- So we expect that the number of integers up to $x = 10^k$ with no digit $a_0 = 3$ is $\approx 9^k$.
- By the PNT, the probability of $n < 10^k$ being a prime is $\approx 1/\log 10^k$.
- This heuristic implies that

$$\{p < 10^k : p \text{ does not contain } 3\} \approx \frac{9^k}{\log 10^k}.$$

Primes with restricted digits III

Remarks

- If $x = 10^k$ then $9^k = 10^{\log_{10} 9^k} = (10^k)^{\log 9/\log 10} = x^{\gamma_{10}}$, where $\gamma_0 = \log 9/\log 10 < 1$.
- This means that the integers with no digit 3 in its decimal expansion is a sparse set. Sparse sets are always difficult to handle.
- We can also generalize the heuristic and expect

$$\{p < x : p \in \mathcal{A}\} \approx \frac{x^{\gamma_b}}{\log x},$$

  where $\gamma_b = \log(b-1)/\log b$.
- Can we prove it rigorously?

Primes with restricted digits IV

- Maynard (2016) showed that for any $x \geq b \geq 10$, there exists two constants C_1 and C_2 such that

$$C_1 \frac{x^{\gamma_b}}{\log x} \leq \{p \in \mathcal{A}\} \leq C_2 \frac{x^{\gamma_b}}{\log x},$$

where

$$\mathcal{A} = \Big\{ \sum_{j \geq 0} n_j b^j : n_j \in \{0, 1, \ldots, b-1\} \setminus \{a_0\} \Big\},$$

and $\gamma_b = \log(b-1)/\log b$.

- For $b \geq 2 \times 10^6$, Maynard (2015) showed that

$$\{p \in \mathcal{A}\} \sim \kappa_b(a_0) \frac{x^{\gamma_b}}{\log x},$$

for some constant $\kappa_b(a_0)$.

Primes with restricted digits in arithmetic progressions I

Question

How are primes p with no digit a_0 in its b-adic expansion distributed in an arithmetic progression $p \equiv c \pmod{d}$ for $(c, d) = 1$?

- Can we expect a BV type theorem for primes with no digit a_0 in arithmetic progression?

For technical convenience, we will work with $\Lambda(n) = \log p \cdot 1_{n=p^m}$ for some m. Essentially, we will count primes with weights $\log p$.

Remark

PNT $\Leftrightarrow \sum_{n \le x} \Lambda(n) \sim x$ as $x \to \infty$.

Primes with restricted digits in arithmetic progressions II

Theorem 1 (N., 2020+)

Let $\delta > 0$, let b be an integer that is sufficiently large in terms of δ, and let $\mathcal{A} = \{\sum_{j\geq 0} n_j b^j : n_j \in \{0,\ldots,b-1\} \setminus \{a_0\}\}$ be the set of integers with no digits a_0 in its b-adic expansion and let $D < x^{1/3-\delta}$. Then for any constant $A > 0$ we have

$$\sum_{\substack{d \leq D \\ (c,d)=1}} \Bigg| \sum_{\substack{n<x \\ n \equiv c \pmod{d}}} \Lambda(n) 1_{\mathcal{A}}(n) - \frac{\beta_b(a_0)}{\varphi(d)} \sum_{\substack{n<x \\ (n,d)=1}} 1_{\mathcal{A}}(n) \Bigg| \leq C(b) \frac{x^{\gamma_b}}{(\log x)^A}, \quad (1)$$

where $\gamma_b = \log(b-1)/\log b$ and $\beta_b(a_0)$ is a constant depending on b and a_0.

Note that we can only take $D \leq x^{1/3-\delta}$. However, in the BV theorem D can be up to $x^{1/2-\epsilon}$.

Primes with restricted digits in arithmetic progressions III

Theorem 2 (N., 2020+)

Let $\delta > 0$, let b be an integer that is sufficiently large in terms of δ, and let $\mathcal{A} = \{\sum_{j\geq 0} n_j b^j : n_j \in \{0,\ldots,b-1\} \setminus \{a_0\}\}$ be the set of integers with no digits a_0 in its b-adic expansion and

$$D_1 D_2 \leq x^{4/9-\delta}, \quad D_1 D_2^{3/2} \leq x^{1/2-\delta}, \quad D_1 \leq x^{1/3-\delta},$$

then for any $A > 0$, we have

$$\sum_{d_1 \leq D_1} \sum_{\substack{d_2 \leq D_2 \\ (d_1,d_2)=1 \\ (c,d_1 d_2)=1}} \left| \sum_{\substack{n<x \\ n\equiv c \pmod{d_1 d_2}}} \Lambda(n) 1_{\mathcal{A}}(n) - \frac{\beta_b(a_0)}{\varphi(d_1 d_2)} \sum_{\substack{n<x \\ (n,d_1 d_2)=1}} 1_{\mathcal{A}}(n) \right| \leq C(b) \frac{x^{\gamma_b}}{(\log x)^A},$$

where $\gamma_b = \log(b-1)/\log b$ and $\beta_b(a_0)$ is a constant.

Applications

- Titchmarsh type divisor problem: $\sum_{p \leq x} 1_{\mathcal{A}}(p)\tau(p-1)$
- Primes of the form $p = 1 + m^2 + n^2$ and $p \in \mathcal{A}$.

High level summary of proof

Recall that we want to estimate the sum

$$\sum_{\substack{n<x \\ n\equiv c \pmod{d}}} \Lambda(n)1_{\mathcal{A}}(n).$$

- Use the circle method to convert the sum into its Fourier transform.
- Divide the sum into two parts: 'major arcs' and 'minor arcs'.
- For major arcs, use the distribution of $\mathcal{A}$ and Λ in arithmetic progressions.
- For minor arcs, use the $L^\infty - L^1$ cancellation idea.
- Key ingredients:
  - L^1 bound for the Fourier transform of $\mathcal{A}$ (follows from Maynard's work)
  - L^∞ bound for the Fourier transform of Λ in arithmetic progressions on an average over moduli $d \leq D$.
- The key challenge is to take D as large as possible. We use Vinogradov's idea to estimate those sums.

Sketch of proof I

The proof uses the circle method. It is based on the following identity:

$$\frac{1}{L}\sum_{0\le \ell<L} e^{2\pi i\ell/L} = 1_{\ell=0}. \tag{2}$$

Let us see how it will be helpful to estimate our sum. For $x = b^k$, we write

$$\sum_{\substack{n<x\\ n\equiv c \pmod{d}}} \Lambda(n)1_{\mathcal{A}}(n) = \sum_{\substack{n,m<x\\ n\equiv c \pmod{d}}} \Lambda(n)1_{\mathcal{A}}(m)1_{m=n}$$

$$= \sum_{\substack{n,m<x\\ n\equiv c \pmod{d}}} \Lambda(n)1_{\mathcal{A}}(m)\left(\frac{1}{x}\sum_{0\le\ell<x} e^{2\pi i(m-n)\ell/x}\right)$$

$$= \frac{1}{x}\sum_{0\le\ell<x}\left(\sum_{m<x} 1_{\mathcal{A}}(m)e^{2\pi i m\ell/x}\right)\left(\sum_{\substack{n<x\\ n\equiv c \pmod{d}}} \Lambda(n)e^{-2\pi i n\ell/x}\right)$$

Sketch of proof II

$$\sum_{\substack{n<x \\ n\equiv c \pmod d}} \Lambda(n)1_{\mathcal{A}}(n) = \frac{1}{x}\sum_{0\le\ell<x} \widehat{1}_{\mathcal{A}}(\ell/x)\widehat{\Lambda}_d(-\ell/x), \tag{3}$$

where

$$\widehat{1}_{\mathcal{A}}(\ell/x) = \sum_{m<x} 1_{\mathcal{A}}(m)e^{2\pi i m\ell/x}, \tag{4}$$

and

$$\widehat{\Lambda}_d(-\ell/x) = \sum_{\substack{n<x \\ n\equiv c \pmod d}} \Lambda(n)e^{-2\pi i n\ell/x}. \tag{5}$$

We have made the simple looking expression into a complicated one in (3). How will it help us?

Sketch of proof III

We write

$$\frac{\ell}{b^k} = \frac{a}{q} + \frac{\eta}{b^k}, \tag{6}$$

for some integers a and q so that $(a, q) = 1$.

- "Major arc": If q and η are "small" ($\leq (\log b^k)^A$), we use the estimates from the distributions of the set $\mathcal{A}$ and Λ in arithmetic progressions.
- "Minor arc": If both q and η are large ($> (\log b^k)^A$), we use the $L^\infty - L^1$ cancellation philosophy as done by Maynard in his work.

Sketch of proof IV (Minor arc)

- Minor arc: When ℓ is in minor arc, the idea is to take advantage of the d averaging and estimate our sum as

$$\approx \max_{\ell \text{ in minor arc}} \sum_{d \le D} \left| \frac{1}{x} \sum_{0 \le \ell < x} \widehat{1}_{\mathcal{A}}(\ell/x) \widehat{\Lambda}_d(-\ell/x) \right|$$

$$\approx \max_{\ell \text{ in minor arc}} \underbrace{\sum_{d \le D} \left| \widehat{\Lambda}_d(-\ell/x) \right|}_{L^\infty \text{ bound for } \Lambda} \cdot \underbrace{\left| \frac{1}{x} \sum_{0 \le \ell < x} \left| \widehat{1}_{\mathcal{A}}(\ell/x) \right| \right|}_{L^1 \text{ bound for } \mathcal{A}}.$$

- L^1 bound for $\mathcal{A}$ is small. In fact, for $x = b^k$, it behaves like $(\log b)^k$, which is "small" compared to b^k if b is large enough. Actually, this is not enough for our purpose as we need to take advantage of both q and η parameters in the decomposition of the ℓ/b^k.
- For the L^∞ bound for Λ, we use Vinogradov's idea to decompose Λ into various pieces (known as Type I and Type II sums).

Conclusion

- In general, it is known that circle method cannot be applied to a binary problem. For example
  - We can use the circle method to solve the ternary Goldbach conjecture (every odd number greater than 5 can be written as the sum of three primes).
  - But the circle method fails for the binary Goldbach conjecture (every even number greater than 2 can be written as the sum of two primes. Reason: L^1 bound for primes is too big.
  - The circle method is applicable for the missing digit and primes as L^1 bound for the set $\mathcal{A}$ is small. Note that this is a binary problem. We want n to be prime and at the same time missing a digit.

Thank you for your attention!