Outline: Primes in APs and Sieve Theory

Hunter Liu and Daniel Rui 6/6/24

This document is a list of the questions we hope you can answer, and topics/themes we hope you will understand by the end of our talk! The material for the talk is drawn mostly from Matomäki&Merikoski&Teräväinen (MMT) "Primes in arithmetic progressions and short intervals without L-functions", and Terry Tao's 254A Notes.

Overarching Questions:

1. Fix an integer (the *modulus*) $q \geqslant 2$. For a *residue class* $a \in (\mathbb{Z}/q\mathbb{Z})^\times$, how many primes p in the arithmetic progression (AP) $a\ (q)$? (Dirichlet: infinitely many.) How soon can I be guaranteed to see a prime $p \equiv a\ (q)$? (Linnik: first such prime $\ll q^L$ for some $L > 0$.)
2. How does the "conspiracy" of quadratic characters $\chi : (\mathbb{Z}/q\mathbb{Z})^\times \to \{\pm 1\}$ with certain "extreme" functions, like the Liouville function λ (i.e. "the presence of a Siegel zero") impact the study of the above questions?
3. What are the ideas behind sieve theory? What are the main results/"out-of-the-box tools"? How can I apply them to some concrete examples? What are the fundamental limitations? How can we push past those limitations?

Part I: Dirichlet

1. How Euler used the Euler product factorization $\zeta(s) \coloneqq \sum_n \frac{1}{n^s} = \prod_p (1 - \frac{1}{p^s})^{-1}$ to prove infinitude of primes, more precisely (by taking log of Euler product and manipulating) $\sum_p \frac{1}{p} = \infty$. ($\sum_{n \leqslant x} \frac{1}{n} = \log x + O(1)$, so can write $\log \zeta(1) = \log \infty$, and so $\sum_p \frac{1}{p} = \log\log \zeta(1) = \log\log\infty$. This is obviously nonsense, but actually one can show rigorously that $\sum_{p \leqslant x} \frac{1}{p} = \log\log x + O(1)$; called Mertens2.)
2. What goes wrong with $\sum_p \frac{1_{p \equiv a\ (q)}}{p}$? Why is complete multiplicativity important?
3. Fourier analysis on abelian group $G = (\mathbb{Z}/q\mathbb{Z})^\times$ to bring back multiplicativity.

$$\hat{f}(\chi) \coloneqq \frac{1}{|G|} \sum_{x \in G} f(x)\overline{\chi(x)} \quad \rightsquigarrow \quad f(x) = \sum_{\chi \in \hat{G}} \hat{f}(\chi)\chi(x) \quad \Longrightarrow \quad 1_{x=a} = \frac{1}{|G|} \sum_{\chi \in \hat{G}} \chi(x)\overline{\chi(a)}$$

4. What is the principle character χ_0?
5. Why is $L(1, \chi) \neq 0$ important?
6. THEME: Duelling Conspiracies
7. What are quadratic characters $\chi : (\mathbb{Z}/q\mathbb{Z})^\times \to \{\pm 1\}$ and why are they especially important?
8. What "extreme behaviors" might quadratic characters χ "conspire" with, and how would that impact $L(1, \chi)$?
9. THEME: Siegel Zeroes/Conspiracies in Quadratic Characters
10. (((If cover 11.5 later, discuss ruling out $L(1, \chi) = 0$ by assuming f.s.o.c. $\zeta(s)L(s, \chi) = \sum_n \frac{[1 * \chi](n)}{n^s}$ is holomorphic at $s = 1$, Landau's theorem for Dirichlet series with non-negative coefficients.)))

Part II: Sieve Theory

1. Basic setup of sieves, via most basic example of $\sum_{\sqrt{x} \leqslant p < x} 1$; i.e. sieving $a_n = 1 \cdot 1_{\sqrt{x} \leqslant n < x}$ up to *sifting level* $z = \sqrt{x}$.

2. Notation: $P(z) = \prod_{p<z} p, \qquad X_d = Xg(d) + r_d, \qquad V(z) = \prod_{p<z}(1-g(p))$

3. THEME: PHILOSOPHY OF SIEVE THEORY: what can I know about the sifted sum of a_n, ONLY knowing the X_d information? How best to utilize this information?

4. THEME: TRUNCATION/MAIN-TERM VS. ERROR-TERM TRADEOFF.

5. Level of distribution D, and parameter $s = \frac{\log D}{\log z}$ (and its role in the Fundamental Lemma of Sieve Theory)

6. How to use Buchstab identity (and iterations of it) to truncate sums while **maintaining inequalities** (upper/lower bound).

7. Where does the first error term $S_2 = \sum_{p_2<p_1<z,\ \neg A_2(p_1,p_2)} 1_{p_1 p_2 | n} 1_{p_*(n)=p_2}$ of the Buchstab iteration come from? What does it have to do with products of 3 primes?

8. Multiplying the Buchstab iteration formulas by a_n and summing $\sum_{n<x}$, arrive at

$$S(\mathcal{A}, z) = \sum_{\substack{d|P(z) \\ d \leqslant D}} \underbrace{[\mu(d) 1_{\underline{\mathcal{D}}}(n)]}_{=:\underline{\lambda_d}} \cdot X_d + S_2(\mathcal{A}, D, z) + S_4 \dots$$

   where (the D, z dependence is in the details of the predicate $A_2(p_1, p_2; D, z)$)

$$S_2(\mathcal{A}, D, z) = \sum_{\substack{p_2 < p_1 < z \\ \neg A_2(p_1,p_2)}} \sum_{\substack{n<x \\ p_1 p_2 | n \\ p_*(n) = p_2}} a_n$$

9. Statement of Sieve Black Box (SBB) for $\kappa = 1$ (key restriction: $s \geqslant 2$). (((If do 11.5, then need $\kappa = 2$ as well.)))

10. THEME: PARITY PROBLEM: $(1 \pm \lambda(n))$ have the same X_d information, so sieve theory can not tell them apart. Why does that lead to the lower bound in the Sieve Black Box (SBB) of $f(2) = 0$?

11. Back to primes in AP (MMT Prop. 11.4), using SBB on $a_n = 1_{n \equiv a\ (q)}$. Must do "Split Sum Trick" to lower sifting level z so that $s = \frac{\log D}{\log z}$ passes SBB threshold $s \geqslant 2$.

12. Key assumption of $\sum \frac{1-\psi(p)}{p}$ being small, i.e. $\psi(p) = 1$ often!

13. (((If have time, do 11.5, which finds primes in APs using the "opposite extreme" of $\sum \frac{1+\psi(p)}{p}$ being small, i.e. $\psi(p) = -1$ often.)))

Part III: Linnik

1. Consider $\mathcal{A}^{(t)} := (1_{n \equiv a\ (q)})_{n \leqslant t} =: (a_n)$

2. Want to apply SBB to the $\underline{\lambda_d}$ term in the formula for $S(\mathcal{A}^{(x)}, \sqrt{x})$ in II(8) above. Sadly Error-term can only take up to level of distribution $D = x^{1-}$, so must do "Split Sum Trick" and then apply SBB (that is where integral term and error term come from), and get

$$\sum_{\sqrt{x} \leqslant p < x} a_p = S(\mathcal{A}^{(x)}, \sqrt{x}) \geqslant -\boxed{\int \ldots} + S_2(\mathcal{A}^{(x)}, \underbrace{x^{1-}}_{=:D}, \sqrt{x}) + [\text{error}]$$

3. Turns out right sum to look at is "logarithmic sum" $\sum_{\sqrt{x} \leqslant p < x} \frac{a_p}{p}$ instead of $\sum_{\sqrt{x} \leqslant p < x} a_p$. Relate the two sums via partial summation/FTC (FTC will lead to negative integral term becoming double integral).

4. The $S_2(\mathcal{A}, \ldots)$ term survives the summation by parts mostly intact (minor variations, with the one major change being the transition from a_n to $\frac{a_n}{n}$), transforming to

$$S_2(\mathcal{A}, \sqrt{x}) = \sum_{\substack{n<x \\ p_1p_2|n \\ p_*(n)=p_2 \\ p_2<p_1<\sqrt{x} \\ \neg A_2(p_1,p_2;\ldots)}} a_n \quad \rightsquigarrow \quad \tilde{S}_2(\mathcal{A}) = \sum_{\substack{\sqrt{x} \leqslant n<x \\ p_1p_2|n \\ p_*(n)=p_2 \\ p_2<p_1<\sqrt{n} \\ \neg A_2'(p_1,p_2;\ldots)}} \frac{a_n}{n}.$$

5. Can restrict RHS sum further (thus producing lower bound of $\tilde{S}_2$) to only consider $n = p_1p_2p_3$, i.e. **product of exactly 3 primes**. So in total we have

$$\sum_{\sqrt{x} \leqslant p < x} \frac{1_{p \equiv a\ (q)}}{p} \geqslant -\boxed{\int\int \ldots} + [\text{error}] + \sum_{\substack{\sqrt{x} \leqslant p_1p_2p_3 < x \\ p_*(n)=p_2 \\ p_2<p_1<\sqrt{p_1p_2p_3} \\ \ldots}} \frac{1_{p_1p_2p_3 \equiv a\ (q)}}{p_1p_2p_3}.$$

6. Further manipulations lead to

$$[\text{RHS 3-prime sum}] \geqslant \frac{1}{3} \sum_{x^{1/6} \leqslant p_1,p_2,p_3 < x^{1/3}} \frac{1_{p_1p_2p_3 \equiv a\ (q)}}{p_1p_2p_3} + [\text{error}]$$

(The error comes from the fact that in one sum we only have $p_1 \neq p_2$, but in the other we could have $p_1 = p_2$. So we need to figure out how the cases where not all $p_1, \ldots, p_3$ are distinct contribute.)

7. The above is a triple convolution of a function $g(b)$! Normalize to mean 1: $\mathbb{E}_{b \in G} g(b) = 1 + o(1)$ (using Mertens2 from I(1) above!). Along with this L^1-control, also have ("Brun-Titchmarsh") L^∞-control $g(b) \leqslant 2+$.

8. THEME: Triple Convolution Good.

9. State MMTProp9.1: $\delta = \frac{1}{2+}$, $g : G \to [0,1]$ with mean $\geqslant \delta$. (Here abelian group $G = (\mathbb{Z}/q\mathbb{Z})^\times$.) Assume mass ($\geqslant \frac{1}{2}\eta$) in ALL index 2 cosets containing a. Then, get lower bound on triple convolution at a, i.e. $(g * g * g)(a) \geqslant \ldots$

10. Fourier inversion (see I(3) above!) leads to $\frac{1}{|G|^2}(g * g * g)(a) = \sum_{\chi \in \hat{G}} \overline{\chi(a)} \hat{g}(\chi)^3$.

11. First, the principle character (see I(4) above!) χ_0 contributes $\geqslant \delta^3$. Try naive triangle inequality on the other χ, i.e. what's the worst case of $|\hat{g}(\chi)|$? (Trivial bound is $\leqslant \delta$.)

12. Calculate $\hat{g}(\chi)$, write as weighted sum of exponentials $\sum_{b\ (m)} e(\frac{b}{m}) \cdot w_b$, where $g(\bullet) \in [0,1] \implies w_b \in [0, \frac{1}{m}]$ (L^∞-control). Also, $\sum_{b=0}^{m-1} w_b = \delta$ (L^1-control). Remember $\delta = \frac{1}{2+}$.

13. By "m-spoke wheel with δ-chipped spokes, originally each of length $\frac{1}{m}$" illustration, and $\delta = \frac{1}{2+}$, we see that for $m \geqslant 3$ we get non-trivial bound, which ends up being

$$\max_{\chi \in \hat{G}: \chi^2 \neq \chi_0} |\hat{g}(\chi)| \leqslant \frac{1}{4}\sqrt{1 + (4\delta - 1)^2}.$$

At around $\delta = \frac{1}{2+}$, this better upper bound is $\leqslant 0.36$!

14. Again, THEME: Siegel Zeroes/Conspiracies in Quadratic Characters

15. Finally know how to split the triple convolution sum III(10):

$$\begin{aligned}\frac{1}{|G|^2}(g * g * g)(a) &= \sum_{\chi \in \hat{G}} \overline{\chi(a)} \hat{g}(\chi)^3 \\ &\geqslant \underbrace{\delta^3}_{\chi_0} - \sum_{\chi \in \hat{G}: \chi^2 \neq \chi_0} |\hat{g}(\chi)|^3 + \sum_{\substack{\chi \in \hat{G}: \chi^2 \neq \chi_0 \\ \chi(a)\hat{g}(\chi) < 0}} (\chi(a)\hat{g}(\chi))^3 + \underbrace{\sum_{\substack{\chi \in \hat{G}: \chi^2 \neq \chi_0 \\ \chi(a)\hat{g}(\chi) \geqslant 0}} (\chi(a)\hat{g}(\chi))^3}_{\geqslant 0} \\ &\geqslant \delta^3 - \max\left\{ \max_{\chi \in \hat{G}: \chi^2 \neq \chi_0} |\hat{g}(\chi)|, \quad \max_{\substack{\chi \in \hat{G}: \chi^2 \neq \chi_0 \\ \chi(a)\hat{g}(\chi) < 0}} (-\chi(a)\hat{g}(\chi)) \right\} \cdot \sum_{\chi \in \hat{G}: \chi^2 \neq \chi_0} |\hat{g}(\chi)|^2\end{aligned}$$

16. So we just need lower bound on $\chi(a)\hat{g}(\chi)$. And finally, MMTLemma9.2: we use assumption that there is mass ($\geqslant \frac{1}{2}\eta$) in ALL index 2 cosets containing a to conclude lower bound $\chi(a)\hat{g}(\chi) \geqslant \eta - \delta \iff (-\chi(a)\hat{g}(\chi)) \leqslant \delta - \eta$.

17. Final lower bound of

$$\frac{1}{|G|^2}(g * g * g)(a) \geqslant \delta^3 - \max\left\{\frac{1}{4}\sqrt{1 + (4\delta - 1)^2}, \quad \delta - \eta\right\} \cdot \left(-\delta^2 + \sum_{\chi \in \hat{G}} |\hat{g}(\chi)|^2\right).$$

Use Plancherel to bound last sum by $\leqslant \delta$ and WIN!