Bombieri- Vinogradov Theorem

Lily Anne Baker

April 2024

1 Introduction

Generating functions are a powerful tool in number theory that can be used to solve recurrence relations. In this talk, I will discuss the purpose of generating functions and how they are used. I will begin with a basic definition of generating functions, introduce four operations on generating functions, and explain how to find both generating and closed functions using the Fibonacci sequence. Lastly, I will talk about asymptotic approximations of the Fibonacci Sequence.

2 Prime Number Theory Recap

Prime number theory recap This presentation greatly references the prime number theory. Here's a quick refresher on what that is. The PNT states that the number of primes up to x is approximately x/log x. This is equivalent to the statement that Ψ(x) x (sigh x is equivalent to x) where Ψ (x) is the sum of f $\wedge$(n) from 1 to x and in turn $\wedge$(n) is log p if n = p^k for some k 1 and 0 otherwise, so a sort of "weighted prime counting function." Now what if we are dealing with primes congruent to a (mod q) (aka, primes that when divided by q have the same remainder)? The PNT states that the number of primes congruent to a (mod q) is about $\frac{1}{\varphi(q)} * \frac{x}{logx}$ primes up to x. Remember that φ(q) ("fi") is Euler's totient function, which counts the number of ints up to q that are co-prime with q.

3 What we're interested in

This presentation is focused on bounding the ERROR TERM. The error term measures the difference between the actual number of primes up to x vs the predicted number of primes based on q. (AKA, it tells us how good our prediction is when counting prime numbers). Here, our error term is: $E * (x, q) = \Psi(x; a, q) - \frac{x}{\varphi(q)}$ is the estimate for the number of primes congruent to a mod p where $\Psi(x)$ is the sum of (n) over $n \leq x$ which are congruent to a modulo q. In this presentation, we want to find the error in this estimate. How can we bound $E(x; a, q) = \Psi(x; a, q) - x/\varphi(q)$?

4 Generalized Reimann Hypothesis

GRH (generalized Riemann hypothesis). Is a major open conjecture. The GRH implies that $E*(x;q) = O\sqrt{x}(logx)^2)$. When we assume the GRH is true, we can say that the sequence of prime numbers is predictable/orderly. This equation above is the error term in the PNT. The error term E(x;q) represents how closely we can predict the number of primes up to a certain point x using a given mathematic function q. We use the GRH to estimate this error term. So, when we assume the GRH is true, we are saying that the error should be of the order of $\sqrt(x)logx^2$. In other words, the error term E(x;q) should grow no faster than $\sqrt(x)logx^2$ as x increases. This is good, bc if our error term were to grow too quickly, then our predictions would become less accurate. Overall, the GRH tells us that if we average over q, we almost get the bound that the GRH would predict: the sum of $E*(x;q)$ over $q \leq Q$ is $O\sqrt{x}Q(logx)^5)$, provided we put some suitable bounds on Q. So in particular, the "average value: of $E*(x;q)$ is $O(x^(1/2)(logx)^5)$, which is just a couple of log factors away from the GRH bound.

5 Presentation outline

Now that I've given a brief explanation of the prime number theory, our goals for the presentation, and the GRH, I will walk you through our next steps. As I said, our goal is to bound the error term and see how accurate it really is. To do this, we will talk about the large sieve. Following my explanation of the large sieve, I will introduce Dirichlet characters. If you guys are unfamiliar with Dirichlet characters, do not worry, I am only going to give a brief explanation of them. They're super basic arithmetic functions with good properties that often show up when we're worrying about primes in arithmetic progressions. Next, I will go into the modifications of the large sieve, and how we can apply it to the BVT.

By the end of this presentation, you should have a general understanding of the large sieve and the bvt. You hopefully will have a deeper understanding of prime number distribution, and have the tools for analyzing prime number distributions. As a disclaimer, I will abbreviate the generalized Reiman hypothesis as GHR, the prime, and the Bombieri- Vinogradov Theorem as BVT (written on board). Please let me know if you have any questions so far.

6 The Large sieve

The term large sieve was coined by Linnik, a Ukrainian- Soviet mathematician. He is creditable for the "large sieve", and many other advancements in Prime number theory. Taking a look at the "large sieve", the sieve itself is not that large. Some may say it doesn't even classify as a sieve, but rather just an inequality. Linnik considered the large sieve to be large in the sense that it excludes a lot more congruence classes (mod p) than other sieves. For example, take a look at some of the sieves we have talked about so far in class.

let's look at the Brun's sieve. The Brun's sieve is combinatorial, whereas the large sieve is fundamentally an analytic result. Additionally, the Bruns sieve is specifically tailored to study twin primes and prime constellations, such as prime pairs with a fixed difference of 2n+1. It targets specific patterns within the prime number sequence. The large sieve is conceptually broader in its application and impact on prime number analysis.

Let S(x) be a trigonometric polynomial:

$$\sum_{M+1}^{M+N} a_n e(nx)$$

- N>0 and N and M are integers
- a_n are arbitrary complex numbers
- $e(x)^{2\pi ix}$

In its basic form, the large sieve of Linnik is an inequality of the form

$$\sum_{r=1}^{R} |S(x_r)|^2 \leqslant \Delta(N,\delta) \sum_{M+1}^{M+N} |a_n|^2.$$

- Here the Xr are arbitrary real numbers that are distinct modulo 1 - R is the number of residue classes modulo n - X_r represents a set of residue classes
- $S(X_r)$ represents the sequence formed by the elements in Xr
- δ is a measure of their spacing, given by:

$$\delta = \min_{r \neq s} \|x_r - x_s\|,$$

Where the minimum is taken over all possible pairs r, s with r$\neq$ s and $\|x\|$denotes the distance from x to the nearest integer. The factor Δ(N,δ) must depend on both N and δ.

In fact, Δ(N,δ) $\geq N$, for if an=1 for all n and x1=1, then the sum on the left is ¿= N2, while the sum on the right is equal to N. On the other hand,

$$\int_0^1 \sum_{r=1}^{R} |S(x_r + x)|^2 \, dx = R \sum_{M+1}^{M+N} |a_n|^2.$$

If it happens that X_r are equally spaced then $\delta = R - 1$, so we deduce that Δ(N,δ) $\geq \delta^{-1}$. So what does this mean? X_r are arbitrary real numbers that

are distinct modulo 1. When I say they are "equally spaced", I am referring to a scenario where the difference between consecutive points is constant.
Lets look theorem 1, where S(x) and δ be as in the first and third equation respectively

$$\sum_{r=1}^{R} |S(x_r)|^2 \leqslant (N + \delta^{-1}) \sum_{M+1}^{M+N} |a_n|^2. \tag{1.4}$$

Moreover, if

$$\delta_r = \min_{s \neq r} \|x_r - x_s\| \tag{1.5}$$

for all r, then

$$\sum_{r=1}^{R} (N + \tfrac{3}{2}\delta_r^{-1})^{-1} |S(x_r)|^2 \leqslant \sum_{M+1}^{M+N} |a_n|^2. \tag{1.6}$$

The inequality from 1.4 contains several previous formulations of the large sieve. In particular,Gallagher, Bombieri and Davenport showed that $\pi N+\delta^{-1}$, 2 $\max(N, \delta^{-1})$and $N+2\delta^{-1}$ are all permissible expressions for Δ(N,δ). Moreover, Bombieri and Davenport have given examples in which Δ(N,δ) $= N + \delta^{-1} - 1$, so that 1.4 is extremely sharp. The weighted sieve in the last equation 1.6 is fundamentally more delicate than 1.5. The weights are particularly useful in arithmetic applications because the Farey fractions are irregularly spaced. (formally, the farey fractions of order n, denoted by Fn, are the set of irreducible fractions $\frac{a}{b}$ st $0 \leq 0 \leq b \leq n$ and gcd (a,b)=1.

Using this theorem, we will do Corollary 1. Let N (fancy) be a set of Z integers in an interval [M+1, M+N]. For each prime p let ω(p) denote the number of residue classes mod p which contain no element of N. Then

$$Z \leqslant L^{-1},$$

$$L = \sum_{q \leqslant z} (N + \tfrac{3}{2}qz)^{-1} \mu(q)^2 \prod_{p|q} \frac{\omega(p)}{p - \omega(p)}$$

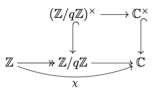
Here, the error term occurs as the term $\frac{3}{2}$ qz instead of cz^2 (which happens in Corollary 4.3). This leads to significant improvements when ω(p) is small. Remember as I said earlier, it's better to have a small value so that over time, the estimate does not fall inaccurate.
Notation wise, the fancy script N is the set of sieved integers with size Z. the primes between M and M+n, so this gives a similar upper bound of sieved sets to what we've seen before, though there are different bounds and details like omega(p). This corollary will be important for the application to the BVT later on.

7 Dirichlet Characters

Eventually, we will use this information about the large sieve to solve the BVT. However, to do that, we need to understand Dirichlet Characters.

Let q $\in Z > 0$. A Dirichlet character of period q (function repeats itself every

q integers) is a function $x \longrightarrow \mathbb{C}$ included by a homomorphism $Z/qZ^* \longrightarrow \mathbb{C}^*$ (i.e, an irreducible character of (Z/qZ)*) in the following way:



equivalently, it is a function : $Z \longrightarrow \mathbb{C}$ which is:
1) Periodic modulo q. This condition states that the function Z(n) repeats itself every 1 set along the integers. If X(n) is known for some n, then X(n+q) = X(n) for all integers n.
2) Totally multiplicative. X(n) is totally multiplicative if it satisfies the property x(n+m)=x(n)x(m) for all integers n and m. This means that the function behaves multiplicatively under addition.
3) Satisfies x(1)=1
4) Satisfies x(n) =0 for all $n \in Z$ such that $(n, q) > 1$

As stated before, we denote the set of Dirichlet characters of period q by X_q. Since $(Z/qZ)^*$ is a finite abelian group, the irreducible characters of $(Z/qZ)^*$ are all one-dimensional and form a group (under multiplication) that is isomorphic to $(Z/qZ)^*$. This means that X_q, the set of Dirichlet characters of period q, forms a group under pointwise multiplication. Each Dirichlet char corresponds to an element of $(Z/qZ)^*$ and the multiplication of characters corresponds to the multiplication of characters in $(Z/qZ)^*$. Finally, note that the size of X_q, or the number of Dirichlet chars of period q in X_q is equal to the order of $(Z/qZ)^*$, denoted φ(q).

This visual definition may be a bit confusing. Here is an alternative definition:

Definition Dirichlet characters. Let G be the group of reduced residue classes modulo k. Corresponding to each character f of G we define an arithmetical function $\chi = \chi_f$ as follows:

$$\chi(n) = f(\hat{n}) \quad \text{if } (n, k) = 1,$$
$$\chi(n) = 0 \qquad \text{if } (n, k) > 1.$$

The function χ is called a Dirichlet character modulo k. The principal character χ_1 is that which has the properties

$$\chi_1(n) = \begin{cases} 1 & \text{if } (n, k) = 1, \\ 0 & \text{if } (n, k) > 1. \end{cases}$$

So now that I've explained what Dirichlet characters are, let's focus on their application. Dirichlet characters play a large role in the large sieve method. They are used to decompose sequences and sieve out unwanted terms. For example, let's make a sequence S(n) representing the number of primes congruent to a mod q up to n. We can sieve out terms from S(n) that do not satisfy our congruence condition by using our Dirichlet char Xa. Terms where $X_a(n) = 0$ will be sieved out. Removing those terms, we will be left with a sequence of terms that follow the congruence condition. This is one of the ways we use Dirichlet characters.

Suppose that for a Dirichlet character X:

$$S(\chi) = \sum_{n=M+1}^{M+N} a_n \chi(n)$$

Then,

$$\sum_{q \le Q} \frac{q}{\varphi(q)} \sum_{\chi \bmod q}^{*} |S(\chi)|^2 \le \lambda(N, Q) \sum_{n=M+1}^{M+N} |a_n|^2$$

holds with $\lambda(N, Q) \ll N + Q^2$.

This is a general inequality for Dirichlet characters that will help us prove the BVT. Additionally, Dirichlet characters are used in Vaughan's identity.

8 Vaughan's identity.

I am not going to spend a lot of time on Vaughan's identity, but I will say that its purpose is to express the error term $E^*(x; q)$ in terms of Dirichlet character sums. This identity allows us to apply the large sieve method separately to each Dirichlet character component and will help us bind the resulting error term later on. As you can see there are a lot of summation symbols. That is because Vaughn's identity allows us to rewrite the error term as a sum over characters mod q. After expressing the error term in terms of character sums, we can apply estimates or bounds to these character sums (which is where the GRH comes into play)

Theorem 1 (Vaughan's identity). Let $U, V \ge 1$. For all positive integers n we have

$$\Lambda(n) = \Lambda(n)1_{[1,U]}(n) + \sum_{\substack{ab=n \\ a \le V}} \mu(a) \log b - \sum_{\substack{ab=n \\ a \le UV}} \sum_{\substack{cd=a \\ c \le V, d \le U}} \mu(c)\Lambda(d) - \sum_{\substack{ab=n \\ a>U, b>V}} \Lambda(a) \sum_{\substack{c \mid b \\ c \le V}} \mu(c).$$

9 Application to the BVT

For the remainder of this presentation, I will prove the BVT using all of the information we have learned so far. There are a lot of steps that go into proving the BVT, many of which I will skip over because of time. But the overall idea is that we modify the large sieve using Dirichlet characters.

Our goal is to find an estimate for E*(x,q) not for a fixed value of q but on average over all q up to a certain bound. We will use the large sieve to show

that for all $A > 0$,

$$\sum_{q \leqslant Q} E^*(x,q) \ll x^{\frac{1}{2}} Q (\log x)^5,$$

Provided that $x^{\frac{1}{2}} (logx)^{-A} \leq \mathbb{Q}^{\frac{1}{2}}$

To carry out this proof, we will use the GRH, Vaughan's identity, and an important lemma.

Step 1) asusme the GRH. This is where the introduction comes in handy. I will not prove how we get to this equation because it is quite complicated. So, assuming the GRH, we have the following:

$$E(x;q,a) = O(x^{\frac{1}{2}+\varepsilon})$$

for all $\varepsilon > 0$.

We also let

$$E(x,q) := \max_{\substack{a \\ (a,q)=1}} |E(x;q,a)| \quad \text{and} \quad E^*(x,q) := \max_{y \leqslant x} E(y,q).$$

Step 2) Apply Vaughan's identity to express the von Mangoldt function as a sum involving Dirichlet characters. We use the large sieve to remove unwanted terms (numbers that don't fit our congruence conditions) from these sums. This is shown below:

This is what that looks like

$$\sum_{q \leq z} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\substack{\chi \pmod q \\ \chi \neq \chi_0}} \max_{y \leq x} \left| \sum_{n \leq y} \chi(n) \Lambda(n) \right| \ll (z^2 x^{\frac{1}{2}} + x + z x^{\frac{5}{6}}) (\log x)^5.$$

$$\Lambda(n) = \Lambda(n) 1_{[1,U]}(n) + \sum_{\substack{ab=n \\ b \leq V}} (\log a) \mu(b) - \sum_{\substack{abc=n \\ b \leq V \\ c \leq U}} \mu(b) \Lambda(c) - \sum_{\substack{ab=n \\ a > U \\ b > V}} \Lambda(a) \sum_{\substack{cd=b \\ d \leq V}} \mu(d),$$

Above, we express $\bigwedge N$ as a sum over certain indicator functions. As explained earlier, we bound these sums using the GRH. After that, we combine the bounds into one final bound. This bond turns out to be:

$$(z^{\frac{5}{2}} x^{\frac{1}{3}} + x + z x^{\frac{5}{6}}) log^5 x$$

This bound provides an estimate for the error term in the context of the BVT. it ensures that the error term does not grow too quickly with respect to the parameters x and z.

10 Conclusion

In conclusion, during this presentation, we covered many topics. The first, as a refresher, was the Prime number theory. The second was the generalized Riemann hypothesis, which we used in the proof of the BVT when finding bounds for the sums of the indicator functions. Third, I talked about the large sieve. This method helps us sift through sequences of primes in arithmetic progressions, removing terms that do not fit our congruence conditions. Dirichlet characters play a huge role in the large sieve method, especially when applying it to the BVT. I then introduce Vaughns identity. This identity is extremely important to proving the BVT. it allows us to express error terms in terms of character sums, which can then be sifted through (using the large sieve and Dirichlet characters) and bounded (using the GRH). Combining all of this information, we were able to find an estimate of the error term for the BVT.

Proving this theorem, we are able to deepen our understanding of prime numbers and arithmetic progressions. We are able to learn more about the distribution of prime numbers, something that is super important in the world of number theory and cryptography.