MATH 250: THE DISTRIBUTION OF PRIMES

ROBERT J. LEMKE OLIVER

0. Motivation and goals for the course

General comments on the course

There are a few overarching goals for this course. Our main aim is to say real, interesting, and modern things about prime numbers in a way that is accessible to and enjoyable for people who are not analytic number theorists. (Though you are certainly encouraged to see the light during the course of the semester!) This means that our main focus in proving theorems will be to convey the ideas behind the proof without getting lost in the weeds of proving technical analytic estimates. That said, our proofs will have real content – the material deserves an honest presentation – but we will black box many results from analysis, and *we will assume the Riemann hypothesis whenever convenient.*

Wait, what?

The Riemann hypothesis is true, or at least near enough to true, that you should think about its role in a proof roughly as, "Here's how things really work." In this language, making a proof not rely on the Riemann hypothesis amounts to showing that the Riemann hypothesis is close enough to true, to showing that reality can't conspire against us too badly. This is a beautiful and important theory, but it is not the story we're telling here. For readers who want a complete, unconditional treatment of the prime number theorem and its cousins, we strongly recommend Davenport's classic *Multiplicative number theory.*

There are some particular mathematical themes we'll be exploring in this course. The central theme will be that *knowledge of the distribution of primes in arithmetic progressions leads to profound consequences.* Related to this is the fact that many of the classical theorems about prime numbers, like the prime number theorem, are not endpoints. They are beautiful results that can stand on the merit of their statement to be sure, but they are also useful results that play important roles in other proofs. And of course, exactly the same can be said about recent developments in the theory; a striking example of this is Maynard's work on small gaps between primes (how close together can two prime numbers be?) directly informing his work on large gaps (how far apart?). We will present at least the first of these theorems in this course (Theorem 0.7 below). Time and interest permitting, we may also discuss the second.

Theorems we will prove

The first major theorem we'll prove in this course is the prime number theorem.

Theorem 0.1 (Prime number theorem, qualitative version)**.** *Let* $\pi(x) = \#\{p < x : p \text{ prime}\}$. *Then as* $x \to \infty$*,*

$$\pi(x) \sim \frac{x}{\log x},$$

Date: September 20, 2017.

that is,

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log x} = 1.$$

Notation. Whenever we write $\log x$, we mean the natural log; that is, $\log e = 1$ and $\frac{d}{dx}\log x = 1/x$.

Theorem 0.1 is the prime number theorem as you probably first saw it. It's a surprising result – what's the connection between prime numbers and $\log x$? – and also one of the crowning achievments of 19th century mathematics (the theorem was proved in 1898 by Jacques Hadamard and Charles de la Vallée Poussin). But notice that this statement is the "endpoint version" of the theorem. It's not actually useful for anything! Suppose you need to know how many prime numbers there are up to a million. Is the answer about $10^6/\log(10^6)$? Or is $x/\log x$ only a reasonably good approximation once x is massive, say on the order $10^{10^{10}}$? Theorem 0.1 is silent on the subject. It only tells you that the approximation is good once x is sufficiently large, without saying what "sufficiently large" actually means. To state a more precise version of the prime number theorem, we need some notation.

Notation. Given two functions $f(x)$ and $g(x)$, we write $f(x) = O(g(x))$ as $x \to \infty$ if there is some constant $C > 0$ such that $|f(x)| \leq Cg(x)$ for all sufficiently large x[1] . The constant C is referred to as the *implied constant*. We will also very frequently write

$$f(x) = g(x) + O(h(x))$$

to mean $f(x) - g(x) = O(h(x))$.

The following is the best known, unconditional[2] version of the prime number theorem.

Theorem 0.2. *As $x \to \infty$,*

$$\pi(x) = \mathrm{li}(x) + O\left(x \exp\left(-c\frac{(\log x)^{2/3}}{(\log\log x)^{1/5}}\right)\right),$$

where c is some positive constant, and

$$\mathrm{li}(x) = \int_2^x \frac{1}{\log t}\, dt.$$

There are a few things to discuss here. First off, the error term probably appears inscrutable. You should loosely process it as a function which is larger than $x^{0.99}$, or indeed $x^{1-\epsilon}$ for any $\epsilon > 0$, but smaller than $x/(\log x)^A$ for any $A \geq 1$. Exotic big-oh statements like this are common in analytic number theory, but if we're willing to assume the Riemann hypothesis (as we do in this course), things clean up quite a bit.

[1]If the restriction to sufficiently large x bothers you, as it reasonably might after we railed against its implicit presence in Theorem 0.1, take comfort in the fact that most functions we will care about may as well be assumed to have domain $[1, \infty)$. (You don't really gain anything by looking at $\pi(x)$ for $x < 1$.) Suppose that the inequality $|f(x)| \leq Cg(x)$ holds for all $x \geq N$. Since the interval $[1, N]$ is compact, if the function $|f(x)/g(x)|$ is either continuous or bounded above by a continuous function, it will have a unique maximal value, M, on this interval. The inequality $|f(x)| \leq \max(M, C)g(x)$ will then hold for all $x \geq 1$. Thus, a "big-oh" statement can typically be assumed to hold for all x at the expense of having a larger implied constant.

[2]i.e., not relying on the Riemann hypothesis

Theorem 0.3. *Assume the Riemann hypothesis. Then as $x \to \infty$,*

$$\pi(x) = \mathrm{li}(x) + O\left(x^{1/2} \log x\right).$$

The error term in Theorem 0.3 can be loosely described as saying, half the digits of the approximation are correct. For example, the one millionth prime number is 15,485,863, and $\mathrm{li}(15485863) \approx 1000409.5$. This is a small instance where we can visually verify the statement, "The Riemann hypothesis is close enough to true."

A second issue you might have with the prime number theorem is more fundamental: why is the prime number theorem an approximate statement and not an exact one? Wouldn't an exact formula for $\pi(x)$ be more desireable? In fact, we have exact formulas for $\pi(x)$, which we call "explicit formulas." We develop such a formula in a later lecture, but we will see there that this formula involves infinitely many terms. This is true for a fundamental reason – the primes are inherently too complicated to be encapsulated by a single, finite expression (we justify this somewhat circularly after we have developed the explicit formula) – but it means that the explicit formula is hard to use in its raw, exact form. Even convergence of the formula is not obvious! The prime number theorem separates the wheat from the chaff in the explicit formula and thereby makes the formula useful.[3]

A third issue one might have with the prime number theorem is the presence of the function $\mathrm{li}(x)$ instead of a more "natural" function. As it happens, the analytically "right" way to count primes is not as direct as with $\pi(x)$, but instead to weigh primes differently according to their logarithm. (We explain this in a future lecture.) We will ultimately prove the prime number theorem in the form

$$\sum_{p \le x} \log p = x + O(x^{1/2} (\log x)^2).$$

The right-hand side in this formula is much nicer, accounting for the fact that the left-hand side is more amenable to study. In uncoupling the weights $\log p$ to return to $\pi(x)$, the right-hand side is forced to become $\mathrm{li}(x)$.

The next major theorem we will prove concerns primes in arithmetic progressions.

Theorem 0.4 (Dirichlet's theorem)**.** *Let $q \ge 2$ and a be coprime integers. Then there are infinitely many primes $p \equiv a \pmod q$.*

Dirichlet's theorem is beautiful, answers a natural question, and is more than capable of standing on its own terms. But again, this is an "endpoint" theorem! It doesn't tell us how likely primes are to be congruent to $a \pmod q$, just that there are infinitely many. In other words, we want an analogue of Theorem 0.3 for these sets of primes. Luckily, we have just such a thing!

Theorem 0.5 (The prime number theorem for primes in arithmetic progressions)**.** *Assume the generalized Riemann hypothesis* [4]*. If $q \ge 2$ and a are coprime integers, let $\pi(x; q, a) :=$*

[3]This is a somewhat reductive viewpoint, of course. There are many more refined questions about prime numbers one may ask that require a sophisticated understanding of the explicit formula that doesn't discard infinitely many terms as chaff. A notable example of this is work on primes in short intervals, where our best unconditional knowledge far exceeds what can be obtained from Theorem 0.2, getting strikingly close to what would be obtained from Theorem 0.3. We do not pursue these topics in this course, however.

[4]There are unconditional versions of this theorem, with the same quality error terms as Theorem 0.2, but with the possible presence of a secondary main term which is expected to not exist but can't be ruled out.

$\#\{p \leq x : p \equiv a \pmod q\}$. *Then*

$$\pi(x; q, a) = \frac{1}{\phi(q)}\mathrm{li}(x) + O(x^{1/2}\log x),$$

where $\phi(q) = \#\{1 \leq a \leq q : \gcd(a, q) = 1\}$.

Theorem 0.5 shows in a concrete way that each residue class $a \pmod q$ is equally likely; we say the primes are equidistributed $\pmod q$ for each q. The main theme we will explore in the latter part of this course is that this equidistribution implies many striking and sometimes quite surprising consequences. The first application we will see is a theorem due to Daniel Shiu.

Theorem 0.6 ("There are arbitrarily long Shiu strings")**.** *Let $q \geq 2$ and a be coprime integers. For any integer $k \geq 2$, there exists a string of k consecutive primes $p_1, \ldots, p_k$ such that each $p_i \equiv a \pmod q$.*

For example, as a consequence of Theorem 0.6, we can find a million primes in a row, each of which ends in the digit 1. This might seem to contradict the equidistribution of Theorem 0.5, but in fact, if we believe that consecutive primes should behave roughly independently, then each possible pattern $p_1 \equiv a_1 \pmod q$, $p_2 \equiv a_2 \pmod q$, $\ldots$, $p_k \equiv a_k \pmod q$ should be equally likely (i.e., "random"). Loosely speaking, therefore, we should think of Shiu's theorem as saying, it is possible to toss a coin a million times in a row and get heads every time, albeit with an incredibly small (but positive!) probability[5]. What's more surprising is that Theorem 0.6 is almost the limit of our knowledge. For example, it is presently unknown whether there exist infinitely many pairs of consecutive primes p_1 and p_2 such that $p_1 \equiv 1 \pmod{10}$ and $p_2 \equiv 3 \pmod{10}$.

The next theorem we will prove is one of the most exciting recent developments in analytic number theory, that there exist bounded gaps between primes. The prime number theorem implies that the "average" spacing between consecutive primes should be on the order of $\log x$; that is, if we let p_n denote the n-th prime, then typically we should expect $p_{n+1} - p_n \approx \log n$. However, the primes can often be much closer than this! The twin prime conjecture asserts that $p_{n+1} - p_n = 2$ should hold infinitely often. This conjecture remains far out of reach, but in 2013, Yitang Zhang proved that there exists some constant C such that $p_{n+1} - p_n \leq C$ infinitely often. Not six months later, James Maynard and Terrence Tao independently reworked the technical underpinnings of Zhang's method (the so-called Goldston-Pintz-Yıldırım method) to make it simultaneously simpler and more powerful.

Theorem 0.7 ("Bounded gaps between primes"; Zhang, Maynard, Tao)**.** *Let $p_1, p_2, \ldots$ denote the sequence of primes in increasing order. For any integer $m \geq 1$, there exists a constant C_m such that*

$$\liminf p_{n+m} - p_n \leq C_m.$$

Zhang's proof applies only to the case $m = 1$ and yields the explicit value $C_1 = 70{,}000{,}000$. The Maynard-Tao proof (which we will present in this course) applies to any m, and originally yielded the value $C_1 = 600$; with subsequent efforts of the Polymath project, this has been reduced to $C_1 = 246$.

[5]In fact, we know that Shiu strings occur with positive probability thanks to work of Maynard, though this is not something that Shiu proved originally. We will prove Shiu's version in these notes.

Another exciting development occurred in 2013: Harald Helfgott completely resolved the *ternary Goldbach conjecture*. The Goldbach conjecture asserts that every even integer $n \geq 4$ can be written as the sum of two primes. This conjecture remains out of reach, but a close cousin is the ternary Goldbach conjecture, sometimes called the weak Goldbach conjecture, which asserts that every odd integer $n \geq 7$ can be written as the sum of three primes. Helfgott showed that this is indeed true for all $n \geq 7$, but his proof is too technical to reasonably be presented in any course, let alone this one with its de-emphasis on technical details. However, the basic idea of the proof goes back much earlier, to Vinogradov, who proved the following theorem, for which we will provide a modern proof due to Soundararajan.

Theorem 0.8 ("Ternary Goldbach is eventually true")**.** *There is some constant N_0 such that every odd integer $n \geq N_0$ is the sum of three primes.*

At this point in lectures, we will take stock of how much time is left in the course. We stress again: none of the theorems in this course are endpoints, even Theorems 0.6-0.8 that were motivated as applications of Theorem 0.5. Theorem 0.6 is proved using the Maier matrix method, an ingenious combinatorial device originally used to prove that the prime number theorem can be wrong when applied to extremely short intervals. Theorem 0.6 also naturally raises the question of how often other patterns arise, on which there is recent work of the author and Soundararajan. The work of Zhang, Maynard, and Tao relies on the Selberg sieve, a ubiquitous tool in analytic number theory, and the ideas behind the proof of Theorem 0.7 play a crucial role in Maynard's work on large gaps between primes, the dual problem to the twin prime conjecture. The proof of Theorem 0.8 is intimately connected to the Hardy-Littlewood conjectures, which also play a role in Theorem 0.7 and the work of Lemke Oliver and Soundararajan on consecutive primes. Using the Hardy-Littlewood conjectures, it is possible to develop a robust probabilistic view of the primes. Notably, assuming Hardy-Littlewood, Gallagher showed that the primes behave like a Poisson process with parameter $\log x$. Theorems 0.1-0.8 presented above will form the core material of this course, but we hope to develop some of the ideas of this paragraph further.