Dirichlet's Theorem on Primes in Arithmetic Sequences
Math 129 Final Paper

Gregory Valiant

May, 2005

1 Introduction

Dirichlet's theorem on primes in arithmetic sequences states that in any arithmetic progression $m, m+k, m+2k, m+3k, \ldots$, there are infinitely many primes, provided that $(m,k) = 1$. Euler first conjectured a result of this form, claiming that every arithmetic progression beginning with 1 contained an infinitude of primes. The theorem as stated was conjectured by Gauss, and proved by Dirichlet in 1835. The proof given below follows the original proof rather closely. I believe an elementary proof was found, though it is fairly unenlightening and unintuitive.

2 Characters

Although in our proof of Dirichlet's theorem, we will only make use of 1-dimensional characters from $(\mathbb{Z}/k\mathbb{Z})^* \to \mathbb{C}$, some of their properties hold in the more general setting, without complicatating the proofs. Thus we will briefly consider general characters.

Definition 2.1 *Given a finite group G, and a representation $\rho : G \to GL(V)$, the* character *of the representation ρ is the map $\chi : G \to \mathbb{C}$ defined by*

$$\chi(g) = trace(R_g),$$

where R_g is the matrix representation given by some choice of basis for V.

Note that this is well-defined, since the trace will simply be the sum of the eigenvalues of R_g, and thus will be independent of choice of basis.

First we shall establish some basic properties of general characters, and then we will focus our attention to 1-dimensional (Dirichlet) characters.

Proposition 2.1 *As above, let G be a finite group, and ρ a representation of G on vector space V, with character χ.*

1. $\chi(1) = dim(V)$.

2. $\forall g, g' \in G, \chi(g') = \chi(gg'g^{-1})$. *(Thus characters are class functions.)*

3. $\chi(g^{-1}) = \overline{\chi(g)}$, *where* $\overline{z}$ *denotes the complex conjugate of* z.

Proof: Since ρ is a homomorphism from $G \to GL(V)$, it must send $1_G \to I$. Thus we have

$$\chi(1) = trace(I) = dim(V).$$

For part 2, again since ρ is a homomorphism, we have $R_{gg'g^{-1}} = R_g R_{g'} R_{g^{-1}}$. Now just note that since the characteristic polynomial of a linear operator is independent of basis, so is the trace, and thus $trace(R_g R_{g'} R_{g^{-1}}) = trace(R_{g'})$, as claimed.

For part 3, we rely on the finiteness of G. Since $|G|$ is finite, for $g \in G$, let d denote its finite order. Since $I = R_{g^d} = (R_g)^d$, R_g must be a matrix of order at most d, and thus its eigenvalues $\lambda_1, \ldots, \lambda_n$ are all roots of unity, and thus have complex norm equal to 1. In particular, $\lambda_i^{-1} = \overline{\lambda_i}$. Thus we have

$$\chi(g^{-1}) = \lambda_1^{-1} + \ldots + \lambda_n^{-1} = \overline{\lambda_1} + \ldots + \overline{\lambda_n} = \overline{\chi(g)}.$$

Throughout the remainder of this paper, we will restrict our attention to 1-dimensional characters from $(\mathbb{Z}/k\mathbb{Z})^* \to \mathbb{C}$. The first main difference to note is that in the case of 1-dimensional characters (characters derived from 1-dimensional representations), each character χ is a homomorphism from $G \to \mathbb{C}^*$, in particular $\chi(g)\chi(h) = \chi(gh)$. This observation yields the following useful proposition.

Proposition 2.2 *In our restricted setting, with* $G = (\mathbb{Z}/k\mathbb{Z})^*$*, the values attained by characters are* $|G|^{th}$ *roots of unity.*

(Recall that $|G| = \phi(k)$, where $\phi(k)$ is Euler's totient function, defined to be the number of positive integers less than or equal to k.)

It is worth noting that these characters form a group, with the law of composition being multiplication of functions:

$$\chi\chi'(g) = \chi(g)\chi'(g).$$

We shall denote this *character group* by $\hat{G}$.

Proposition 2.3 $G \cong \hat{G}$.

Proof: By the structure theorem for abelian groups, we have

$$G \cong C_{d_1} \oplus \ldots \oplus C_{d_n},$$

where C_m denotes the cyclic group with m elements, and $d_i > 1$, $d_1|d_2 \ldots$. Thus G will be generated by the n-element set $\{(0,\ldots,0,1,0,ldots,0)\}$. Furthermore, by the multiplicativity, a character is uniquely determined by its value on each of the generators. Now, we just need to consider the possible values that a character can attain on a given generator. By the multiplicativity, a character must take the generator with a 1 in the i^{th} component to one of the d_i^{th} roots of unity, and any choice of one such root will yield a unique and valid character. Thus given a character that takes the j^{th} generator to $e^{\frac{2\pi i m_j}{d_j}}$, we can represent it as the list $(m_1,\ldots,m_n)$. Our bijection is now clear. Futhermore, it is compatible with multiplication, since exponents add.

For G as defined above, If we wish to extend a character χ_G from a map $G \to \mathbb{C}$ to a map $\chi_{\mathbb{N}} : \mathbb{N} \to \mathbb{C}$, we can do so easily by the following definition:

$$\chi_{\mathbb{N}}(n) = \begin{cases} \chi_G(n \mod k) & \text{if } (n,k) = 1 \\ 0 & \text{otherwise} \end{cases}.$$

Note that this extension is still multiplicative, and the product of the extensions of two characters is the extension of the products. We will frequently neglect to specifiy whether we are dealing with an actual character, or its extension, and furthermore, will refer to both maps as 'characters'.

We will now define a special character that appears frequently enough to deserve a special treatment:

Definition 2.2 *The* principle character*, which we will generally denote χ_0 is the extension of the trivial character, thus*

$$\chi_0(n) = \begin{cases} 1 & (n,k) = 1 \\ 0 & (n,k) \neq 1 \end{cases}.$$

Proposition 2.4

$$\sum_{g \in G} \chi_1(g)\overline{\chi_2(g)} = \begin{cases} 0 & \text{if } \chi_1 \neq \chi_2 \\ |G| & \text{if } \chi_1 = \chi_2 \end{cases}$$

Proof: Considering the first case, if $\chi_1 = \chi_2$, then $\chi_1\overline{\chi_2} = \chi_0$, in which case our claim clearly holds.

Now, if $\chi_1 \neq \chi_2$, then $\chi_1\overline{\chi_2} \neq \chi_0$, and thus there must be some $g' \in G$ for which $\chi_1(g')\overline{\chi_2(g')} \neq 1$. Consider the following:

$$\begin{aligned} \chi_1(g')\overline{\chi_2(g')} \sum_{g \in G} \chi_1(g)\overline{\chi_2(g)} &= \sum_{g \in G} \chi_1(g'g)\overline{\chi_2(g'g)} \\ &= \sum_{g'' \in G} \chi_1(g'')\overline{\chi_2(g'')}. \end{aligned}$$

The last step holds since as g ranges over all of G, the product $g'g$ also ranges over all of G. Our proposition follows trivially.

Proposition 2.4 does hold in the general setting, although the proof is slightly more involved. Note that a trivial consequence of proposition 2.4 is that

$$\sum_{g \in G} \chi(g) = \begin{cases} 0 & \text{if } \chi \neq \chi_0 \\ |G| & \text{if } \chi = \chi_0 \end{cases}$$

We will conclude this section with a similar proposition to the above:

Proposition 2.5 $\sum_{\chi \in \hat{G}} \chi(g) = \begin{cases} |G| & \textit{if } g = 1 \\ 0 & \textit{otherwise} \end{cases}$

Proof: In the case that $g = 1$, the proposition follows trivially, since each χ is a homomorphism, and thus must map $1 \to 1$. If $g \neq 1$, then there is some $\chi' \in \hat{G}$ such that $\chi'(g) \neq 1$. Now, just observe that

$$\chi'(g) \sum_{\chi \in \hat{G}} \chi(g) = \sum_{\chi \in \hat{G}} \chi'\chi(g) = \sum_{\chi \in \hat{G}} \chi(g),$$

since as χ ranges over all of $\hat{G}$, so does the product $\chi'\chi$.

3 Basic Properties of the Zeta Function

It will be useful to establish some very basic properties of the Riemann Zeta function, and Euler's product formula. We will start with a definition:

Definition 3.1 *For $s > 1$, the* Riemann zeta function *is defined as*

$$\zeta s = \sum_{n=1}^{\infty} n^{-s}.$$

First, we should note that by the integral test, $\zeta(s)$ converges absolutely for $s > 1$. We would like to extend *zeta* at least to the half-plane $Re(s) > 0$.

Proposition 3.1 *The zeta function can be extended analytically to the entire plane, with the exception of a simple pole at $z = 1$, where it has a residue of 1. Furthermore, the zeros of the zeta function consist of the* trivial *zeros at $-2, -4, -6, \ldots$, and the zeros with $0 \leq Re(s) \leq 1$. (The second set of zeros are with $Re(s) = 1/2$, if we believe the Riemann hypothesis.)*

The above proposition follows from extending the gamma function to the entire plane, and then extending zeta by the 'method of moments'. We shall omit the proof of the above, because it is outside the main scope of this course. For a detailed proof of the above, see Bak and Newman, section 18.2.

The other vital property of the zeta function, is its ability to be expressed as a product over primes. Throught the remainder of this paper, p will refer to a prime number, and a summation (or product) over p will refer to summation (or product)over all primes.

Proposition 3.2 (Euler's product formular) *At least for $s > 1$,*

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}.$$

Proof: First note that when the product on the right is expanded out, we have a summation of terms of the form $\frac{1}{p_1^{a_1} p_2^{a_2} \ldots p_r^{a_r}}$, with the p_i distinct. Provided that both sides converge, the above statement is equivalent to the statement that each positive integer has a unique prime factorization. Should I deal with convergence really explicitly, or can I just state that it works? (This is the only time in the paper that I would need to wave my hands...)

Also note that given any multiplicative function f, as above, we have $\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \left(1 - \frac{f(p)}{p^s}\right)^{-1}$, wherever both sides converge.

4 Proof of Infinitude of Primes

Before embarking on a proof of Dirichlet's theorem, we will first prove that there are infinitely many primes by using identities involving the Zeta function. This minor aside will be useful for two main reasons: first, it will give us a relatively straightforward illustration of the techniques that will be used while establishing some useful relations, and second, it will allow us to claim a stronger result than just that there are an infinitude of primes in an arithmetic progression; we will be able to state that the 'density' of primes in the arithmetic progression $m + kn$ is independent of m (provided, of course, that $(m, k) = 1$). Thus, in other words, the primes are roughly equally distributed among the residue classes modulo k.

Our goal in this section is to prove the following:

Theorem 4.1 *For $s > 1$,*

$$\sum_p p^{-s} = \log \frac{1}{(s-1)} + O(1).$$

Corollary 4.1 *There are an infinitude of primes.*

Proof: Assume for the point of contradiction that there are only finitely many primes. Thus $f(s) = \sum_p p^{-s}$ would be a continuous function of s, and in particular would be bounded on the interval $s \in [0, 2]$. On the other hand, $\log \frac{1}{(s-1)} + O(1)$ gets arbitrarily large near $s = 1$. Thus we have our contradiction.

The following two Lemmas will be useful in proving Theorem 4.1.

Lemma 4.1 *For $s > 1$,*

$$\zeta(s) = \frac{1}{s-1} + O(1)$$

Proof: For $s > 0$, $f(x) = x^{-s}$ is strictly decreasing as a function of x, and thus for $n \in \mathbb{N}$, we have

$$(n+1)^{-s} \leq \int_n^{n+1} x^{-s} dx \leq n^{-s}.$$

Summing over positive $n \in \mathbb{N}$, we get that $\zeta(s) - 1 \leq \int_1^\infty x^{-s} dx \leq \zeta(s)$. The integral evaluates to $1/(s-1)$, from which our Lemma follows.

Lemma 4.2 *For $s > 1$,*

$$\log \zeta(s) = \sum_{n-1}^{\infty} \frac{\Lambda(n) n^{-s}}{\log n},$$

where $\Lambda(n)$ is the von Mangoldt function defined by

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \\ 0 & \text{otherwise} \end{cases}.$$

Proof: Taking logarithms of both sides of the Euler product identity $\zeta(s) = \prod_p (1 - 1/p^s)^{-1}$, and recalling the power series expansion that for $|z| < 1$, $\log \frac{1}{1-z} = \sum_{j=1}^\infty z^k/k$, we get the following:

$$\begin{aligned} \log \zeta(s) &= \log \left(\prod_p (1 - 1/p^s)^{-1} \right) \\ &= \sum_p \log \frac{1}{1 - p^{-s}} \\ &= \sum_p \sum_{j=1}^{\infty} \frac{p^{-ks}}{k}. \end{aligned}$$

By rearranging the terms in the above double sum we see that it is the same sum as that claimed in the statement of the Lemma. (And, since all terms are positive, we can do this rearangement without affecting the sum, or the convergence). Finally note that for all n $\frac{\Lambda(n)}{\log n} \leq 1$, and thus we have that the series in question must converge absolutely wherever the zeta function converges absolutely. Namely, we have that $\sum_{n=1}^\infty \frac{\Lambda(n) n^{-s}}{\log n}$ converges absolutely to $\log \zeta(s)$ for $s > 1$, as desired.

We are now ready to give our proof of Theorem 4.1:
Proof of Theorem 4.1: From Lemmas 4.1 and 4.2, we have the following:

$$\sum_p \sum_{j=1}^{\infty} \frac{p^{-ks}}{k} = \sum_{n=1}^{\infty} \frac{\Lambda(n) n^{-s}}{\log n} = \log \frac{1}{s-1} + O(1),$$

Now, we wish to show that the contribution to this sum from the higher prime powers is finite. This is not hard:

$$\sum_p \sum_{j=2}^{\infty} \frac{p^{-ks}}{k} \leq \sum_p \sum_{j=2}^{\infty} p^{-ks}.$$

Now, just note that each of the inner sums is a geometric series, which sums to $\frac{p^{-2s}}{1-p^{-s}}$, and thus, since $s > 1$, we have

$$\sum_p \sum_{j=2}^{\infty} \frac{p^{-ks}}{k} \leq \sum_p \frac{1}{p^s(p^s-1)} \leq \sum_p \frac{1}{p(p-1)} \leq \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

Thus we have established that the higher powers of primes contribute a constant amount to the original sum, thus $\sum_p p^{-s} = log\frac{1}{s-1} + O(1)$, as desired.

5 Dirichlet L-series

Before proceeding to the proof, it will be helpful define, and establish some basic properties of Dirichlet L-series.

Definition 5.1 *For some character χ, the corresponding* Dirichlet L-series*, $L(s, \chi)$, is defined to be*

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

Note that $L(s, \chi)$ is certainly analytic in for $s > 1$, as can be seen by comparing it termwise with the *Zeta* function.

The following two propositions will help clarify the behavior of Dirichlet L-series.

Proposition 5.1 *$L(s, \chi_0)$ can be analytically extended to the half-plane $Re(s) > 0$, with the exception of a simple pole with residue $\phi(k)/k$ at the point $s = 1$.*

Proposition 5.2 *For $\chi \neq \chi_0$, $L(s, \chi)$ can be analytically extended to the half-plane $Re(s) > 0$.*

Proof of Proposition 5.1 Since characters are multiplicative, we can use Euler's Product Formula, yielding $L(s,\chi_0) = \prod_p \left(1 - \frac{\chi_0}{p^s}\right)^{-1}$. Now, recalling that $\chi_0(p) = 0$ if $p|k$, and $\chi_0(p) = 1$ otherwise, the above yields

$$L(s,\chi_0) = \prod_{p\nmid k}\left(1-\frac{1}{p^s}\right)^{-1} = \prod_{p}\left(1-\frac{1}{p^s}\right)^{-1}\prod_{p|k}\left(1-\frac{1}{p^s}\right) = \zeta(s)\prod_{p|k}\left(1-\frac{1}{p^s}\right).$$

Since $\zeta(s)$ can be analytically extended to the half-plane with a simple pole at $s = 1$, with residue 1, and the second component is analytic in this region, the product is analytic in this half-plane with the exception of the point $s = 1$, where it has residue $\prod_{p|k}\left(1-\frac{1}{p}\right) = \frac{\prod_{p|k}(p-1)}{k} = \phi(k)/k$. To see why $\prod_{p|k}(p-1) = \phi(k)$, just note that $\phi(k)$ is multiplicative.

Proof of Proposition 5.2 For $\chi \neq \chi_0$, we have already shown that $\sum_{n=1}^{k}\chi(n) = 0$, and that $|\chi(n)|$ is either 0 or 1, and thus for any $x \geq 1$, $|\sum_{n\leq x}\chi(n)| \leq \phi(k)$, and, in particular, is bounded. Thus, since for any $s > 0$, the sequence $\{1/n^s\}$ is decreasing monatonically, by Dirichlet's Test for convergence, $L(s,\chi) = \sum_{n=1}^{\infty}\chi(n)n^{-s}$ converges absolutely, and thus is analytic in this region.

The final piece of machinery required before diving into the actual proof is the following proposition:

Proposition 5.3 *For $\chi \neq \chi_0$, $L(1,\chi) \neq 0$.*

Proof: As we will see, there are two very different cases to consider, according to whether χ is a real, or a complex character. The first case to consider is when χ is a real character (maps exclusively to $\mathbb{R}$).

5.1 The Real Character

First note that χ is a real character implies that it maps to ± 1 (the only real units), and thus in order for $\chi \neq \chi_0$ to be a homomorphism, it must map the quadratic residues to 1, and the quadratic nonresidues to -1. Simply put, $\chi(q) = \left(\frac{q}{k}\right)$, where $\left(\frac{\cdot}{\cdot}\right)$ is the Jacobi symbol (a natural extention of the Legendre symbol). There are various approaches that we could now take to conclude that $L(1,\chi) \neq 0$. The method we shall follow differs from Dirichlet's original method, though relies on similar tactics.

We shall define the following function:

$$f(n) = \sum_{d|n}\frac{n\chi(d)}{d}.$$

f is simply the Dirichlet convolution of χ, and the constant function 1. It is easy to see that since χ is multiplicative, so is f. (In fact, it is not hard to show that the Dirichlet

convolution of any two multiplicative functions is also multiplicative. We will now consider the behavior of f on prime powers. First, not that if p divides k, then $f(p^m) = 1$, for any m, since the only divisor of p^m that will have a nonzero contribution will be 1, which is a quadratic residue. Next, for $p \nmid k$, in the case that m is even, $f(p^m) \geq 1$, because each even divisor of m will contribute a $+1$, as will 1, and thus there will be at least one more contributions of $+1$ than -1 in the sum. Similarly, in the case that m is odd, we have $f(p^m) \geq 0$.

From the above properties of f, and f's multiplicativity, we can conclude that $f(n) \geq 0$ for all positive n. Also, $f(n^2) \geq 1$, since all the prime factors of n^2 will have even exponents. Consider now,

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n}^s,$$

the Dirichlet series related to f. We have the following:

$$\begin{aligned} F(s) &= \sum_{n=1}^{\infty} \\ &\geq \sum_{m=1}^{infty} \frac{f(m^2)}{m^{2s}} (\text{ since } f(n) \geq 0) \\ &\geq \sum_{m=1}^{\infty} \frac{1}{m^{2s}} (\text{ since } f(m^2) \geq 1) \\ &= \zeta(2s). \end{aligned}$$

Thus, since $\zeta(2s)$ has a simple pole at $s = 1/2$, $F(s)$ must have some sort of singularity for some $S \geq 1/2$. Now, just note that $F(s) = L(s, \chi)\zeta(s)$ (since for any n, each divisor of n will show up in exactly one of the cross terms, which will account for that term in the sum defining f). If $L(1, \chi) = 0$, then this zero would cancel the simple pole of the zeta function at $s = 1$, in which case $F(s)$ would be analytic on the entir half-plane $s > 0$. But this contradicts our earlier statment that $F(s)$ has a singularity at some $s \geq 1/2$, and thus $L(1, \chi) \neq 0$, as desired.

5.2 The Complex Characters

The second case to consider is when χ is a complex character, and thus $\chi \neq \overline{\chi}$. In this case, consider the quantity

$$P(s) = \prod_{\chi \in \hat{G}} L(s, \chi).$$

Taking the logarithm of both sides, and proceeding in a similar fashion to our proof of 4.2, using the expansion of $log\frac{1}{1-z}$, we get the following:

$$\begin{aligned}\log(P(s)) &= \log\left(\prod_{\chi\in\hat{G}}\prod_{p}\left(1-\frac{\chi(p)}{p^s}\right)^{-1}\right)\\ &= \sum_{\chi\in\hat{G}}\sum_{p}\log(1-\chi(p)/p^s)^{-1}\\ &= \sum_{\chi\in\hat{G}}\sum_{p}\sum_{j=1}^{\infty}\frac{chi(p^j)}{jp^{js}}.\end{aligned}$$

Recalling Proposition 2.5, the above is simply

$$\sum_{p}\sum_{j\geq 1|p^j\equiv 1 \pmod k}\frac{1}{jp^{js}}\geq 0.$$

Thus, since $\log(P(s))\geq 0$,for $s>1$, we must have $P(s)\geq 1$ for $s>1$. To conclude, assume for the point of contradtiction that a complex character χ was such that $L(1,\chi)=0$. Then, clearly, $L(1,\overline{\chi})=0$, in which case $P(s)$ would have a zero of multiplicity at least 2 contributed by $\chi,\overline{\chi}$, and thus the simple pole contributed to $P(s)$ by χ_0 can not offset this zero of multiplicity at least 2. And thus $P(1)=0$, which is not the case, thus no complex character can have $L(1,\chi)=0$. (Note that this nice argument could not have been used to prove the case for real characters, since if χ is a real character, then $\overline{\chi}=\chi$, and thus only one zero is contributed, which will be offset by the simple pole from χ_0.)

6 The Proof

We wish to prove the following theorem:

Theorem 6.1 (Dirichlet) $\sum_{p\equiv\alpha \pmod\beta} p^{-s}=\frac{log\frac{1}{s-1}}{\phi(\beta)}+O(1)$.

The proof of this theorem is now relatively straighforward given the strong machinery established above.
Proof of Theorem 6.1: First, observe that for any m such that $(m,k)=1$, we have

$$\sum_{p\equiv m \pmod k}\frac{\phi(k)}{p^s}=\sum_{p}\frac{\sum_{\chi\in\hat{G}}\chi(pm^{-1})}{p^s}=\sum_{\chi\in\hat{G}}\left(\chi(m^{-1})\sum_{p}\frac{\chi(p)}{p^s}\right).$$

We will now consider the contribution of each character to the total sum.

Considering χ_0, we have that $\chi_0(m^{-1})\sum_p\frac{\chi_0(p)}{p^s}=\sum_p p^{-s}=\log\frac{1}{s-1}$, directly from Theorem 1. Thus the contribution of the principle character is the dominant contribution, and in order to prove our theorem, we must show that $\sum_{\chi\neq\chi_0}\left(\chi(m^{-1})\sum_p\frac{\chi(p)}{p^s}\right)=O(1)$. The following lemma is clearly sufficient:

Lemma 6.1 *For* $\chi \neq \chi_0, \sum_p \frac{\chi(p)}{p^s} < \infty$.

Proof: We will proceed as we have done in the proof of Theorem 4.1, working backwards this time:

$$\begin{aligned}
\sum_p p^{-s} &= \sum_p \sum_{m=1}^{\infty} \frac{\chi(p^m)}{mp^{ms}} - \sum_p \sum_{m=2}^{\infty} \frac{\chi(p^m)}{mp^{ms}} \\
&= \sum_p \sum_{m=1}^{\infty} \frac{\chi(p^m)}{mp^{ms}} + O(1) \\
&= \sum_p \left(\log \left(1 - \frac{\chi(p)}{p^s} \right)^{-1} \right) + O(1) \\
&= \log(L(s,\chi)) + O(1).
\end{aligned}$$

We have already shown that $L(1,\chi) \neq 0$, and that $L(1,\chi)$ is analytic near 1, so $\log(L(s,\chi))$ is analytic near $s = 1$, thus is bounded near $s = 1$. Thus we have proved our proposition, and our proof of Dirichlet's theorem is complete.

7 Where To Now?

Well, now that we have the nice result above, there are several natural questions that we can ask.

The first obvious question, especially in light of our course in which we saw many instances of easy generalizations of properties of $\mathbb{Q}$ to other number fields, is 'Does a similar result apply in arbitrary number fields?'. The answer is YES. In the first part of the 20^{th} century, Nikolai Chebotarev generalized Dirichlet's theorem to arbitrary algebraic number fields that are Galois extensions. The Chebotarev Density Theorem gives a similar asymptotic behavior for the density of prime ideals of number fields. Slightly more specifically, for a finite Galois extension L/K, and a conjugacy class $C \subset Gal(L/K)$, the prime ideals of K for which their corresponding 'Frobenius automorphism' is an element of C, and which are unramified in L, have density $\frac{|C|}{|G|}$. In essence, the different conjugacy classes all have the 'same number' of primes. The proof of this general theorem is quite involved. The Chebotarev density theorem does show up in some practical applications, such as in computing the Galois groups associated to a given irreducible polynomial. We will not venture further in this direction here.

The second natural question is 'What are the error bounds on the asymtotic behavior?'. This question is particularly natural in light of the strong error bounds on the prime number theorem.

In an attempt to investigate the error bounds, I ran some tests of the fraction of primes less than 100,000,000 which lie in various residue classes modulo various primes. The most interesting observation, is that for a given prime p, although the difference

between the numbers of primes in the different residue classes modulo p stays relatively small (as we showed it should), there seemed to be consistently fewer primes in quadratic residue classes. The graph attached plots

$$f_q(x) = \sum_{p<x, p \text{ quad nonres.} \pmod q} 1 - \sum_{p<x, p \text{ quad res.} \pmod q} 1$$

in the case $q = 3$. Considering the primes less than 100,000,000 modulo each of the 24 primes less than 100, for 23 of them $f_q(100,000,000) > 0$, and for the one prime in this range ,61, for which this value was less than zero, the average value of $f_{61}(x)$ over the primes less than 100,000,000 was greater than zero. This was quite compelling, and deserved chasing down. It turns out that this sort of behavior was first noticed by Chebyshev in the special case of primes congruent to 1 and 3 modulo 4. This behavior is known as Chebyshev's bias, and is quite analogous to the behavior of the error function for the prime number theorem.

8 $\pi(x) - Li(x)$

The prime number theorem states that

$$\pi(x) \sim Li(x),$$

where $\pi(x)$ is defined to be the number of primes less than x, and $Li(x) = \int_2^x \frac{dt}{\log x}$. (This statement just means that the probability of n being prime is roughly $\frac{1}{\log n}$.)

From numerical tests, $\pi(x) - Li(x)$ tends to be less than zero. In fact, no x is known for which $\pi(x) - Li(x) > 0$. Nevertheless, in 1914, Littlewood proved that $\pi(x) - Li(x)$ changes sign infinitely often. The best we can do for an example of where this function is negative, is the bound proven in 1987 by te Riele, stating that $\pi(x) - Li(x) > 0$ for some $x < 10^{370}$.

In contrast to these difficult results, the general behavior of $\pi(x) - Li(x)$ can be explained rather easily. $Li(x)$ essentially counts the number of prime powers that are at most x, whereas $\pi(x)$ counts the number of primes. The number of prime squares has the same magnitude as the error term in the prime number theorem (and the number of cubes, and higher powers, is negligeable compared to the number of squares), and thus we should expect that π will fall slightly short of $Li(x)$.

The above intuitive explanation can be solidified using the von Mangoldt 'explicit formula':

$$\sum_{n \le x} \Lambda(n) = x - \frac{\zeta'(0)}{\zeta(0)} - \sum_{z | \zeta(z)=0} \frac{x^z}{z}.$$

(The above can be derived pretty easily, starting with Perron's formula and techniques from complex analysis, but it is outside the scope of this paper.) $\frac{\zeta'(0)}{\zeta(0)} = \log(2\pi)$, and the sum on the right can be split up into the contribution from the trivial zeros of the

zeta function at the negative even integers, and the zeros with $0 < Re(z) < 1$. The contribution from the trivial zeros ends up being $\frac{1}{2}\log(1-\frac{1}{x^2})$. Thus, if we assume the Riemann Hypothesis, Mangoldt's explicit formula yields

$$\begin{aligned}\sum_{p^n \le x} \log(p) &= x - \log(2\pi) - \frac{1}{2}\log(1-\frac{1}{x^2}) - x^{1/2}\left(\sum_{\gamma|\zeta(1/2+i\gamma)=0} \frac{x^{i\gamma}}{1/2+i\gamma}\right) \\ &= x - x^{1/2}\left(\sum_{\gamma|\zeta(1/2+i\gamma)=0} \frac{x^{i\gamma}}{1/2+i\gamma}\right) + o(x^{1/2}).\end{aligned}$$

If we wish to count the density of primes, instead of the density of prime powers, then we should be considering $\sum_{p\le x}\log(p)$. This sum is simply

$$\sum_{p\le x}\log(p) = \sum_{p^n\le x}\log(p) - \sum_{p\le x^{1/2}}\log p + O(x^{1/3}) = x - x^{1/2}\left(1 + \sum_{\gamma|\zeta(1/2+i\gamma)=0}\frac{x^{i\gamma}}{1/2+i\gamma}\right) + o(x^{1/2}).$$

Thus, provided that $\sum_{\gamma|\zeta(1/2+i\gamma)=0}\frac{x^{i\gamma}}{1/2+i\gamma}$ is zero on average, which it is, the extra 1 will cause our true value to be slightly shy of the predicted value $Li(x)$. (The sum over nontrivial zeros is some slowly oscillating trigonometric series.)

9 Chebyshev's Bias

The explanation for the presence of slightly fewer primes in quadratic residue classes is very similar to the above. First we shall give the intuitive explanation, which closely paralells the intuitive explanation given above for the difference $\pi(x) - Li(x)$.

In our proof of Dirichlet's theorem, we were counting the number of prime powers that are in the desired residue class. (Recall the proof of Lemma 4.2.) Instead, what we are interested in is only the density of primes in the above residue classes. Since the contribution of prime cubes, and higher powers is dwarfed by the contribution of prime squares, it suffices to consider the magnitude of our error by assessing the contribution of the prime squares. For quadratic residue classes, the contribution from the prime squares that are in our class will need to be subtracted. Since this amount to be subtracted is of the same order as our error term, it will bias those residue classes.

As in the previous subsection, this intuitive explanation can be rigorized. Assuming the generalized Riemann hypothesis, the analogous formulae are the following:

$$\begin{aligned}\sum_{p^n\equiv q \pmod k, p^n\le x}\ln(p) &= \frac{x - x^{1/2}\left(\sum_\chi \chi(q)\sum_{\gamma|L(1/2+i\gamma,\chi)=0}\frac{x^{i\gamma}}{1/2+i\gamma}\right)}{\phi(k)} + o(x^{1/2}) \\ \sum_{p\equiv q \pmod k, p\le x}\ln(p) &= \sum_{p^n\equiv q \pmod k, p^n\le x}\ln(p) - \sum_{p^2\equiv q \pmod k, p^2\le x}\ln(p) + O(x^{1/3}) \\ &= \sum_{p^n\equiv q \pmod k, p^n\le x}\ln(p) - |\{y|y^2\equiv q \pmod k\}|x^{1/2} + o(x^{1/2}).\end{aligned}$$

Thus we have pinned down this rather interesting side note on the frequency of primes in various residue classes.

10 Final Thoughts

It is quite nice that many of the arguments, and facts about primes apply with minimal adaptation to primes in individual residue classes. This might make us suspect that other facts about the general distribution of primes might also apply in some form to primes in a residue class. For example, we have strict bounds on the maximum gaps between primes. It is not too hard to show that for any n, there is at least one prime between n and $2n$. Can we make a similar argument regarding primes in arithmetic progressions? What about twin primes? For any arithmetic progression $nk+q$, do there exist an infinitude of prime pairs, where a prime pair is defined as $(nk+q, (n+1)k+q)$, in the case that k is even, or $(nk+q, (n+2)k+q)$ in the case that k is odd? The answer should probably be 'YES' in both cases, though I haven't really looked into either question.

11 my comments on my paper

I never actually defined 'density' rigorously, but I think this is fine, since it coincides with the intuitive sense. I leave out most things that involve lots of complex analysis. This should also be okay. As a side note, I wrote my final paper for math 113 (Complex Analysis) on the prime number theorem, and some related questions, so for me this was an especially nice follow-up to that. Thanks for teaching me this semester. Your online course book was wonderful, and will be a very useful resource for me in the future. I'm sorry for not attending more classes this semester. (The course notes were so good, and class was early. . . :) Anyway, thanks again, and good luck in California.

12 References

Artin, *Algebra*, Prentice Hall, Upper Saddle River, 1991.

Davenport, *Multiplicative Number Theory* 2nd Ed, Springer-Verlag, New York, 1980.

Landau, *Elementary Number Theory*, trans. Goodman, AMS Chelsea Publishing, Providence, 1991.

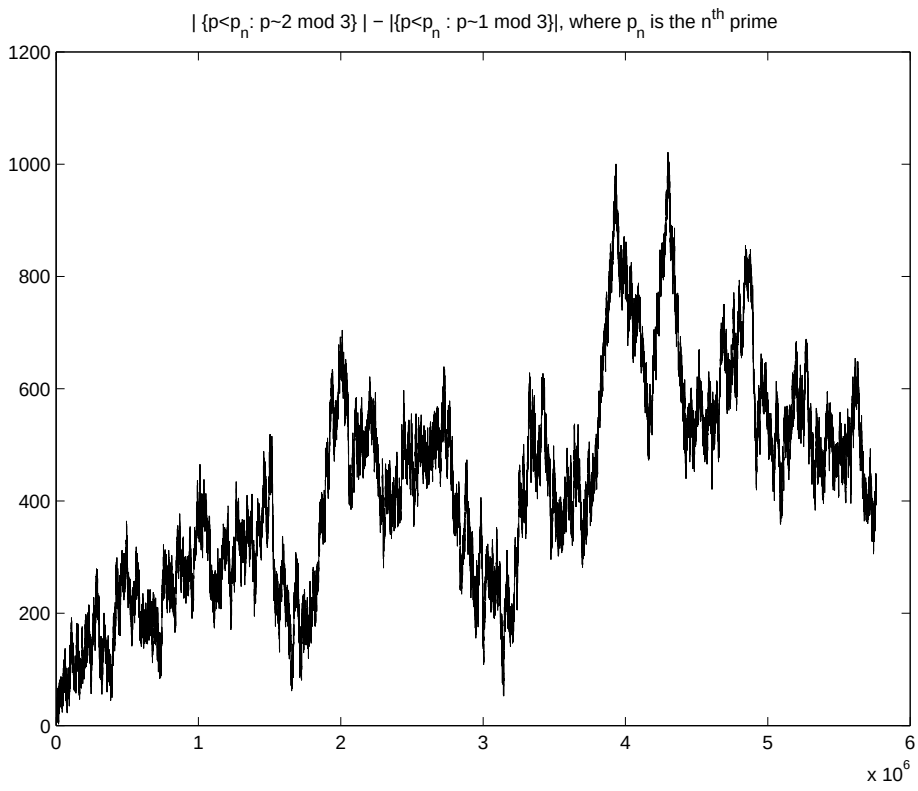
Chapter 6,7 of Chan Huat's course notes, avalable online at
http://www.math.nus.edu.sg/ chanhh/MA4263/Chapter6.pdf,
http://www.math.nus.edu.sg/ chanhh/MA4263/Chapter7.pdf

Niven, Zuckerman,Montgomery, *An Introduction to the Theory of Nunbers* 5th Ed, John Wiley & Sons, 1991.

Bak, Newman, *Complex Analysis* 2nd Ed, Springer, New York, 1997.

Vardi, *An Introduction to Analytic Number Theory.* Available online at
http://algo.inria.fr/banderier/Seminar/Vardi/index.html.

te Riele, *On the Sign Difference* $\pi(x) - Li(x)$ Mathematics of Computation, vol 48, n 177, 1987,pp 323-328.


| {p<p_n: p~2 mod 3} | − |{p<p_n : p~1 mod 3}|, where p_n is the n^th prime
1200
1000
800
600
400
200
0
0
1
2
3
4
5
6
x 10^6