DIRICHLET'S THEOREM ON PRIMES IN ARITHMETIC PROGRESSIONS

KUAT YESSENOV

ABSTRACT. In this expository paper we present a proof of the Dirichlet's theorem on the existence of infinitely many prime numbers in arithmetic progressions $\{a + bn \mid n \geq 0\}$ with relatively prime a and b. [1]

1. INTRODUCTION

We know that the prime numbers greater than 2 are odd. Alternatively, we can say that the arithmetic progression $\{1 + 2k \mid k \in \mathbb{Z}_{>0}\}$ contains infinitely many primes. One may wonder whether same is true for *arbitrary* arithmetic progressions provided some necessary conditions. It can be shown in elementary ways that it is true for progressions $\{\pm 1 + 4k \mid k \in \mathbb{Z}_{>0}\}$ (see [1].)

Our goal is to prove the following celebrated:

Theorem 1.1 (Dirichlet)**.** *Let a and m be relatively prime positive numbers. Then there exist infinitely many prime numbers p such that $p \equiv a \pmod{m}$.*

We will develop our two basic tools first: characters of finite abelian groups and L-series. We are following the proof in [1].

2. CHARACTERS

Consider a finite abelian group G.

Definition 2.1. *A* character *of G is a homomorphism $\chi : G \to \mathbb{C}^\times$ of G into the multiplicative group of complex numbers.*

The characters form a group $\widehat{G}$ with group operation $\chi_1\chi_2(x) = \chi_1(x)\chi_2(x)$ and identity character $\chi_0(x) = 1$ for any $x \in G$. This group is called the *dual* group of G. For any character χ, there is a corresponding character $\bar{\chi}$ obtained by complex conjugation such that $\bar{\chi}(x) = \overline{\chi(x)} = \frac{1}{\chi(x)}$.

Some of the properties of characters are true in the general setting of finite abelian groups although we will restrict our attention to one particular type of groups later.

Date: May 1, 2006.

[1]This is the final paper for the complex analysis class taught in Spring 2006 by professor Andreea Nicoara

Proposition 2.2. *The dual group $\widehat{G}$ is isomorphic to G. In particular, they have the same number of elements.*

Proof. We show $G \cong \widehat{G}$ by induction on the number n of elements of G. For a cyclic group G with generator x, $\chi(x)$ is an n-th root of unity. Clearly, any character χ is uniquely identified by $\chi(x)$. Pick any n-th root of unity σ. Then the characters $\chi_i(x^k) = (\sigma^i)^k$ are the only ones in the group $\widehat{G}$. Notice that $\chi_i = \chi_1^i$, so $\widehat{G}$ is also cyclic of order n and $G \cong \widehat{G}$.

Now suppose G is not cyclic. Then $G = G_1 \times G_2$. There is an isomorphism $\widehat{G} \cong \widehat{G}_1 \times \widehat{G}_2$ which sends a character χ of G to a pair of its restrictions to subgroups $(\chi|_{\widehat{G}_1}, \chi|_{\widehat{G}_2})$. Indeed, one may check that the inverse map is defined by the rule:

$$(\chi_1, \chi_2) \text{ maps to } \chi \in \widehat{G} \text{ such that } \chi(a_1, a_2) = \chi_1(a_1)\chi_2(a_2).$$

Applying the induction assumption, we see that $G_1 \cong \widehat{G}_1$ and $G_2 \cong \widehat{G}_2$, so $G \cong \widehat{G}$. □

Proposition 2.3. *For any character $\chi \in \widehat{G}$ and any element $a \in G$,*

$$\sum_{x \in G} \chi(x) = \begin{cases} \#G & \text{if } \chi = 1_{\widehat{G}} \\ 0 & \text{if } \chi \neq 1_{\widehat{G}}. \end{cases}$$

$$\sum_{\chi \in \widehat{G}} \chi(a) = \begin{cases} \#G & \text{if } a = 1_G \\ 0 & \text{if } a \neq 1_G. \end{cases}$$

Proof. Choose any $y \in G$ such that $\chi(y) \neq 1$. If there is no such one, then $\chi = 1_{\widehat{G}}$ and the first formula follows. Otherwise,

$$\chi(y) \sum_{x \in G} \chi(x) = \sum_{x \in G} \chi(yx) = \sum_{x \in G} \chi(x)$$

Then $\sum_{x \in G} \chi(x) = 0$.

The map $\chi \mapsto \chi(a)$ is a character of $\widehat{G}$. It is trivial only if $a = 1_G$. Now apply our previous formula to the dual group $\widehat{G}$ to obtain the second formula.

□

For the rest of the paper, we fix our attention on the group $G = (\mathbb{Z}/m\mathbb{Z})^\times$, the multiplicative group of residues mod m. We call characters of this group as *characters mod m*.

Definition 2.4. *The order of the group $(\mathbb{Z}/m\mathbb{Z})^\times$ is $\phi(m)$, the* totient function.

Hence, by previous proposition there are $\phi(m)$ distinct characters mod m. It is useful to consider a character $\chi \in \widehat{G}$ as a function on all of $\mathbb{Z}$ in the following way:

$$\chi(n) = \begin{cases} \chi(\bar{n}) & \text{if } n \text{ and } m \text{ are relatively prime} \\ 0 & \text{otherwise} \end{cases}$$

where we denote $\bar{n}$ the image of n in the quotient $\mathbb{Z}/m\mathbb{Z}$.

A function $f : \mathbb{Z}_{>0} \to \mathbb{C}$ is said to be *multiplicative* if for any positive integers n and k, $f(nk) = f(n)f(k)$. Note that characters mod m are multiplicative.

3. Dirichlet L-series and ζ function

For a character χ mod m and $s \in \mathbb{C}$, *the Dirichlet L-series* is defined by

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

Definition 3.1. *The* zeta function *is defined for* $s \in \mathbb{C}$ *by* $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$.

These two functions share many similarities. We would like to explore properties of them which are crucial in the proof of the Dirichlet's theorem. First, we would like to have a test for convergence of the series of the similar type.

Proposition 3.2. *Assume* $\{a_i\}$ *are complex. Let* $F(s) = \sum_{i=1}^{\infty} \frac{a_i}{i^s}$. *If* $F(s)$ *converges for* $s = s_0$ *then it converges on all domain* $Re\{s\} > Re\{s_0\}$ *and it is holomorphic on this domain.*

Proof. Without loss of generality, we may assume that $s_0 = 0$ (by setting $a_i' = \frac{a_i}{i^{s_0}}$.) For a positive number $\alpha < \frac{\pi}{2}$ consider the compact set

$$C_\alpha = \{z \in \mathbb{C} \mid Re\{z\} \geq 0 \ , \ |\arg\{z\}| \leq \alpha\}$$

If we can show that $F(s)$ converges uniformly on any C_α then by Weierstrass theorem ([2] p. 174.) it is holomorphic on the union of the sets which contains the domain $Re\{s\} > 0$.

Indeed, denote the sum $\sum_{i=n}^{k} a_i$ by $A_{n,k}$. Since $F(s)$ converges at 0, by the Cauchy property, for any $\epsilon > 0$ there exists N such that for any $n, k > N$, $|A_{n,k}| < \epsilon$. Using Abel's summation formula, we obtain for $N < n < k$:

$$\left|\sum_{i=n}^{k} \frac{a_i}{i^s}\right| \leq \sum_{i=n}^{k-1} |A_{n,i}| \cdot |i^{-s} - (i+1)^{-s}| + \left|\frac{A_{n,k}}{k^s}\right| < \epsilon \left(\sum_{i=n}^{k-1} |i^{-s} - (i+1)^{-s}| + |k^{-s}|\right)$$

Now for $x = Re\{s\}$:

$$|i^{-s} - (i+1)^{-s}| = |e^{-s\log i} - e^{-s\log(i+1)}| = |-s\int_{\log i}^{\log(i+1)} e^{-ts}dt| \leq |s| \int_{\log i}^{\log(i+1)} e^{tx}dt = \frac{|s|}{x}(i^{-x} - (i+1)^{-x})$$

Notice that for $s \in C_\alpha$ there is a bound $M = \frac{1}{\cos\alpha}$ for $\frac{|s|}{x}$. Then

$$\left|\sum_{i=n}^{k} \frac{a_i}{i^s}\right| < \epsilon(M(n^{-x} - k^{-x}) + k^{-x}) \leq \epsilon(M+1)$$

for any $s \in C_\alpha$. So it converges uniformly and we are done. □

Proposition 3.3. *Let $f : \mathbb{Z}_{>0} \to \mathbb{C}$ be a multiplicative function. Consider the series $F(s) = \sum_{i=1}^{\infty} \frac{f(i)}{i^s}$.*

(a) *If the partial sums $A_k = \sum_{i=1}^{k} f(i)$ are bounded, then $F(s)$ converges for $Re\{s\} > 0$.*
(b) *If the coefficients $\{f(i)\}$ are bounded, then $F(s)$ converges absolutely for $Re\{s\} > 1$, and we have*

$$F(s) = \prod_{p \in P} \frac{1}{1 - f(p)p^{-s}}$$

Proof. For the first part, we use Abel's summation. Assume M is a bound for $|A_i|$ for all i. Then for $n < k$:

$$\left|\sum_{i=n}^{k} \frac{f(i)}{i^s}\right| \leq \sum_{i=n}^{k-1} |A_i - A_{n-1}| \cdot |i^{-s} - (i+1)^{-s}| + \left|\frac{A_k - A_{n-1}}{k^s}\right| \leq 2M \left(\sum_{i=n}^{k-1} |(i^{-s} - (i+1)^{-s})| + |k^{-s}|\right)$$

By previous proposition, it suffices to consider the case when s is real: $\left|\sum_{i=n}^{k} \frac{f(i)}{i^s}\right| \leq \frac{2M}{n^s}$ (remove absolute values above.) For $s > 0$ the Cauchy criterion implies that $F(s)$ converges.

To show the second part, denote by M the bound for coefficients. Then notice

$$\left|\frac{f(i)}{i^s}\right| \leq \frac{M}{|i^s|} = \frac{M}{i^{Re\{s\}}}$$

It is well-known that if $Re\{s\} > 1$ then the sequence on the right converges, so $F(s)$ converges absolutely.

Let $S_N = \{s_1, \ldots, s_k\}$ for positive N be the of prime numbers less than N. Then

$$\prod_{p \in S_N} \frac{1}{1 - f(p)p^{-s}} = \prod_{p \in S_N} \left(\sum_{n=0}^{\infty} \frac{f(p^n)}{p^{ns}}\right) = \sum_{\substack{i \text{ divisible only} \\ \text{by primes in } S_N}}^{\infty} \frac{f(i)}{i^s}$$

As N approaches ∞, S_N approaches P, and the right hand side tends to $F(s)$. □

Throughout the paper we denote the set of all prime numbers by P. Since characters mod m and the constant function 1 are both multiplicative and bounded we have the following:

Corollary 3.4. *The L-series converges absolutely for $Re\{s\} > 1$, in which case*

$$L(s, \chi) = \prod_{p \in P} \frac{1}{1 - \chi(p)p^{-s}}$$

If the character $\chi \neq 1$, then it converges on the domain $Re\{s\} > 0$ and it is holomorphic. The ζ function converges absolutely for $Re\{s\} > 1$. On this domain $\zeta(s) = \prod_{p \in P} \frac{1}{1 - p^{-s}}$.

Proof. The only part we need to verify is that the partial sums $\sum_{i=1}^{k} \chi(i)$ are bounded for $\chi \neq 1$. But character on $\mathbb{Z}_{>0}$ is a periodic function and by property 2.3, the sum of its values over a period is zero. Hence, since the number of elements in a period is $\phi(m)$, for a partial sum $|\sum_{i=1}^{n} \chi(i)| \leq \phi(m)$. □

Proposition 3.5. *Define the function* $\Psi(s) = \zeta(s) - \frac{1}{s-1}$. *Then for* $Re\{s\} > 0$, $\Psi(s)$ *is holomorphic.*

Proof. Indeed,

$$\Psi(s) = \zeta(s) - \frac{1}{s-1} = \sum_{i=1}^{\infty} \frac{1}{n^s} - \int_1^{\infty} t^{-s} dt = \sum_{i=1}^{\infty} \int_n^{n+1} (n^{-s} - t^{-s}) dt$$

Set $\psi_n(s) = \int_n^{n+1} (n^{-s} - t^{-s}) dt$. Each of these functions is holomorphic. So if we can show that the series $\sum \phi_n$ converges uniformly on all compact sets in the domain $Re\{s\} > 0$, then according to Weierstrass theorem ([2] p. 174), Ψ is holomorphic on the domain. But the derivative of the integrand with respect to real number t is $\frac{s}{t^{s+1}}$, so:

$$|\phi_n(s)| \leq \sup_{n \leq t \leq n+1} |n^{-s} - t^{-s}| \leq \frac{|s|}{|n^{s+1}|} \leq \frac{|s|}{|n^{Re\{s\}+1}|}$$

Therefore, the series $\sum \phi_n(s)$ converges uniformly on every compact subset of the domain $Re\{s\} > 0$ since $Re\{s\} \geq \epsilon$ for some $\epsilon > 0$. □

Note that from the product expressions we have:

$$L(s, 1) = \zeta(s) \prod_{\substack{p \in P \\ p \mid m}} (1 - p^{-s})$$

From the result above, we can extend $L(s, 1)$ analytically to $Re\{s\} > 0$ so that it has a simple pole at 1.

When we consider the log of a holomorphic function, we mainly mean its principal branch, so that the Taylor series expansion $\log \frac{1}{1-z} = \sum_{k=0}^{\infty} \frac{z^k}{k}$ is valid.

Corollary 3.6. *We have* $\sum_{p \in P} p^{-s} \sim \log \frac{1}{s-1}$ *for real* $s \to 1+$ *in the domain of convergence.*

Proof. From the product formula above, we derive:

$$\log \zeta(s) = \sum_{p \in P} \sum_{i \geq 1} \frac{1}{i p^{is}} = \sum_{p \in P} p^{-s} + \psi(s)$$

where we denoted by $\psi(s)$ the rest of the terms $\sum_{p \in P, i \geq 2} \frac{1}{i p^{is}}$. We estimate the error term:

$$\psi(s) \leq \sum_{p \in P, i \geq 2} p^{-is} \leq \sum_{p \in P} \frac{1}{p^s (p^s - 1)} \leq \sum_{p \in P} \frac{1}{p(p-1)} \leq \sum_{n \geq 2} \frac{1}{n(n-1)} = 1$$

Hence, the term $\psi(s)$ is bounded. Since $\zeta(s) = \frac{1}{s-1} + \Psi(s)$ has a pole at 1 for Ψ holomorphic at 1, we get the asymptotic relation. □

The property above exposes an interesting fact about the distribution of prime numbers. For out purpose we need to look at a particular *subset* of P and for that we use L-series. It turns out to be essential that $L(1,\chi)$ does not vanish for $\chi \neq 1$. This requires some additional theorems from complex analysis so we devote a separate section for it.

4. $L(1,\chi)$ Does Not Vanish for $\chi \neq 1$

It is natural to consider the following function:

$$\zeta_m(s) = \prod_{\chi \in \widehat{G}} L(s,\chi)$$

For a positive number p relatively prime to m, denote by $\mathrm{ord}(p)$ the order of $\bar{p}$ in the group G. There is a nice product expression for ζ_m:

Proposition 4.1. *For $Re\{s\} > 1$: $\zeta_m(s) = \prod_{\substack{p \in P \\ p \nmid m}} \left(\frac{1}{1 - p^{-s \cdot ord(p)}} \right)^{\frac{\phi(m)}{ord(p)}}$*

Proof. Note that

$$\prod_{\chi \in \widehat{G}} (1 - \chi(p)x) = (1 - x^{\mathrm{ord}(p)})^{\frac{\phi(m)}{\mathrm{ord}(p)}}$$

Indeed, let us expand $1 - x^{\mathrm{ord}(p)} = \prod_{\sigma \in U_{ord(p)}} (1 - \sigma x)$ where we denote the set of all n-th roots of unity as U_n. Since $\chi(p)$ is $\mathrm{ord}(p)$-th root of unity, it suffices to check that there are $\frac{\phi(m)}{ord(p)}$ of them attaining same value. This follows from the isomorphism of the group with its dual.

Now we are using the product formula for L-series:

$$\begin{aligned}
\zeta_m(s) &= \prod_{\chi \in \widehat{G}} L(s,\chi) \\
&= \prod_{\chi \in \widehat{G}} \prod_{p \in P} \frac{1}{1 - \chi(p)p^{-s}} \\
&= \prod_{p \in P} \prod_{\chi \in \widehat{G}} \frac{1}{1 - \chi(p)p^{-s}} \\
&= \prod_{\substack{p \in P \\ p \nmid m}} \left(\frac{1}{1 - p^{-s \cdot \mathrm{ord}(p)}} \right)^{\frac{\phi(m)}{\mathrm{ord}(p)}}
\end{aligned}$$

for $x = p^{-s}$ and we can interchange products due to absolute convergence in the domain. □

We expand the product expression above to obtain a series of the form $F(s) = \sum_{i=1}^{\infty} \frac{a_i}{i^s}$ with nonnegative real coefficients a_i. We claim the following fact for the functions of this type:

Proposition 4.2. *Suppose that $F(s)$ converges for $Re\{s\} > \rho$ where $\rho \in \mathbb{R}$. If the function extends analytically to a neighborhood to ρ then there exist an $\epsilon > 0$ such that $F(s)$ converges for all $Re\{s\} > \rho - \epsilon$.*

We will not prove this fact in this paper. Interested reader should refer to [1] or [3].

Proposition 4.3. *$L(1, \chi)$ is nonzero for $\chi \neq 1$.*

Proof. Assume $L(1, \chi) = 0$ for some character $\chi \neq 1$. All of the functions $L(s, \chi)$ are holomorphic for the region $Re\{s\} > 0$ and nontrivial character χ by proposition 3.4. Recall, that $L(s, 1)$ was extended to the domain so that it has a simple pole at 1.

So the function ζ_m is holomorphic at 1 since pole and zero cancel. Now from proposition 4.2 it follows that ζ_m converges for all s in the domain $Re\{s\} > 0$.

Now we are coming to a contradiction. Indeed, the factor in the product expression of ζ_m is

$$(1 - p^{-ord(p)s})^{-\frac{\phi(m)}{ord(p)}} = \left(\sum_{i=0}^{\infty} p^{-i \cdot ord(p)s}\right)^{\frac{\phi(m)}{ord(p)}}$$

It dominates the sum $\sum_{i=0}^{\infty} p^{-i\phi(m)s}$. Hence, if we look at $\zeta(s)$, it dominates:

$$\prod_{\substack{p \in P \\ p \nmid m}} \left(\sum_{i=0}^{\infty} p^{-i\phi(m)s}\right) = \sum_{(n,m)=1} \frac{1}{n^{\phi(m)s}}$$

If we substitute $s = \frac{1}{\phi(m)}$ we get a contradiction, since the RHS diverges. Hence, our assumption was wrong, and we are done. □

5. The Proof of the Dirichlet Theorem

By now we have all the tools required to prove the Dirichlet's theorem. In order, to give an estimation of the number of prime numbers in a set, we need the notion of *Dirichlet density*:

Definition 5.1. *For a subset A of prime numbers P, the* Dirichlet density $\delta(A)$ *of the set A is the limit of*

$$\frac{\sum_{p \in A} p^{-s}}{\sum_{p \in P} p^{-s}}$$

when s tends to 1 along the real line (provided it exists.)

Let us fix a positive number m. We are interested in the set P_a of prime numbers in $\{a + nm \mid n \in \mathbb{Z}_{>0}\}$ for a relatively prime to m. In fact, we claim even stronger result than Theorem 1.1:

Theorem 5.2. *The density $\delta(P_a)$ exists and is equal to $\frac{1}{\phi(m)}$, i.e. the densities of the sets P_a and P_b are the same for any a and b relatively prime to m.*

As a corollary we will obtain a proof of Theorem 1.1 since nonzero density implies infinitude of P_a.

Proof. Our first step is to consider the following series:

Lemma 5.3. *For any character χ mod m and real $s \to 1+$,*

$$\sum_{p\in P, p\nmid m} \frac{\chi(p)}{p^s} \sim \log \frac{1}{s-1} \text{ for } \chi = 1$$

The series on the left is bounded for $\chi \neq 1$.

Proof. First note that the series above converges for $s > 1$. For $\chi = 1$, the series misses only finitely many terms of the series $\sum_{p\in P} p^{-s}$, and by proposition 3.6 we are done.

Assume now $\chi \neq 1$. To express the series we will use logarithm of L-series $\log L(s,\chi)$. We are using the product formula for $L(s,\chi)$ as in corollary 3.4:

$$\begin{aligned} \log L(s,\chi) &= \sum_{p\in P} \log \frac{1}{1-\chi(p)p^{-s}} = \sum_{p\in P}\sum_{n\geq 1} \frac{\chi(p)^n}{np^{ns}} \\ &= \sum_{p\in P} \frac{\chi(p)}{p^s} + \sum_{\substack{p\in P \\ n\geq 2}} \frac{\chi(p^n)}{np^{ns}} \end{aligned}$$

The second sum $\sum_{\substack{p\in P \\ n\geq 2}} \frac{\chi(p^n)}{np^{ns}}$ is bounded as $s \to 1+$ since it is dominated by the sum from corollary 3.6. Since $L(s,\chi)$ does not vanish in a neighborhood of $s = 1$, $\log L(s,\chi)$ is bounded in that neighborhood. Therefore, the sum $\sum_{\substack{p\in P \\ p\nmid m}} \frac{\chi(p)}{p^s}$ is bounded in the limit $s \to 1+$ as well. □

Now we are able to estimate the numerator expression of the density as follows:

Lemma 5.4.

$$\sum_{p\in P_a} p^{-s} = \frac{1}{\phi(m)} \sum_{\chi\in\widehat{G}} \overline{\chi(a)} \left(\sum_{p\in P, p\nmid m} \frac{\chi(p)}{p^s} \right)$$

Proof. Rewrite the sum on the right as follows using multiplicative property of characters:

$$\sum_{p\in P, p\nmid m} \left(\sum_{\chi\in\widehat{G}} \chi(a)^{-1}\chi(p) \right) p^{-s} = \sum_{p\in P, p\nmid m} \left(\sum_{\chi\in\widehat{G}} \chi(a^{-1}p) \right) p^{-s} = \phi(m) \sum_{p\in P_a} p^{-s}$$

Here used the formula from proposition 2.3 for sum of characters. The inner sum vanishes unless $a^{-1}p \equiv 1 \pmod m$ which is equivalent to $p \in P_a$. □

Our last step is to notice that by lemma 5.3 for real $s \to 1+$,

$$\sum_{p\in P_a} p^{-s} \sim \frac{1}{\phi(m)} \log \frac{1}{s-1}$$

since the only nonzero asymptotic contribution comes from $\chi = 1$. But we know for the denominator expression of the density $\sum_{p\in P} p^{-s} \sim \log \frac{1}{s-1}$ (corollary 3.6). Therefore, $\delta(P_a) = \frac{1}{\phi(m)}$ and we finished the proof of the theorem 5.2. □

6. Conclusion

The usage of complex analysis was essential in this proof as we can see. This demonstrates the power of complex analysis tools and complexifications of real functions in number theory. In fact, that was the step that allowed Dirichlet to prove the theorem, although it has been conjectured by mathematicians long before him. The original proof by Dirichlet in 1837 marked the beginning of rigorous analytic number theory.

References

[1] Serre, J.-P., "A Course in Arithmetic," Springer–Valley, New York, USA, 1973.
[2] Ahlfors, L., "Complex Analysis," McGraw–Hill, USA, 1966
[3] Noam Elkies "Introduction to Analytic Number Theory" course notes.
http://www.math.harvard.edu/~elkies/M259.02/index.html