

(September 24, 2013)

Primes in arithmetic progressions

Paul Garrett garrett@math.umn.edu <http://www.math.umn.edu/~garrett/>

[This document is http://www.math.umn.edu/~garrett/m/mfms/notes_2013-14/05_Dirichlet.pdf]

1. Dirichlet's theorem
2. Dual groups of abelian groups
3. Appendix: analytic continuations

Dirichlet's 1837 theorem combines Euler's argument for the infinitude of primes with *harmonic analysis on finite abelian groups*, and subtler things, to show that there are infinitely many primes $p = a \pmod N$ for fixed a invertible modulo fixed N .

The most intelligible proof uses a bit of analysis, in addition to then-novel algebraic ideas. The analytic idea already arose with Euler's proof of the infinitude of primes, recalled below. Dirichlet's new algebraic ideas isolated primes in different congruence classes modulo N .

Specifically, Dirichlet introduced the *dual group*, or *group of characters*, of a finite abelian group. This was an impetus to the development of the abstract notion of *group*, and of *group representations*, by Schur and Frobenius.

The subtle element is *non-vanishing of L-functions* $L(s, \chi)$ at $s = 1$. For expediency, a first proof of this nonvanishing is given in a Supplement. There are two better lines of argument for non-vanishing $L(1, \chi) \neq 0$, both giving *reasons* for L -functions' non-vanishing. The simpler of the two was used by Dirichlet, expressing products of Dirichlet L -functions as *zeta functions of number fields*. The less simple argument is about 50 years old, and uses *Eisenstein series*. Both better viewpoints will be explained subsequently.

1. Dirichlet's theorem

In addition to Euler's observation that the analytic behavior^[1] of $\zeta(s)$ at $s = 1$ implied the existence of infinitely-many primes, Dirichlet found an algebraic device to focus attention on single congruence classes modulo N .

This section gives the central argument, and in doing so *uncovers* several issues taken up subsequently.

[1.0.1] Theorem: (*Dirichlet*) Given an integer $N > 1$ and an integer a such that $\gcd(a, N) = 1$, there are infinitely many primes p with

$$p = a \pmod N$$

[1.0.2] Remark: The *gcd* condition is *necessary*, since $\gcd(a, N) > 1$ implies there is at most a single prime p meeting the condition $p = a \pmod n$, since any such p would be divisible by the *gcd*. The point is that this *obvious necessary* condition is also *sufficient*.

[1.0.3] Remark: For $a = 1$, there is a simple, purely algebraic argument using cyclotomic polynomials, resembling the Euclidean argument. For general a the intelligible argument involves a little analysis.

Proof: A *character* modulo N is a group homomorphism

$$\chi : (\mathbb{Z}/N)^\times \longrightarrow \mathbb{C}^\times$$

^[1] Euler's proof uses only simple properties of $\zeta(s)$, and only of $\zeta(s)$ as a function of a *real*, rather than *complex*, variable. Given the status of complex number and complex analysis in Euler's time, this is not surprising. It is slightly more surprising that Dirichlet's original argument also was a real-variable argument. Still, until Riemann's 1859 memoir there was little reason to believe that the behavior of $\zeta(s)$ off the real line played a critical role.

Given such a character, *extend it by 0* to all of \mathbb{Z}/N , by defining $\chi(a) = 0$ for a not invertible modulo N . Then compose χ with the reduction-mod- N map $\mathbb{Z} \rightarrow \mathbb{Z}/N$ and consider χ as a function on \mathbb{Z} . Even when extended by 0 the function χ is still *multiplicative* in the sense that

$$\chi(mn) = \chi(m) \cdot \chi(n)$$

whether or not either of the values is 0. The pulled-back-to- \mathbb{Z} version of χ , with the extension by 0, is a *Dirichlet character*. The *trivial* Dirichlet character χ_o modulo N is the character which takes only the value 1 (and 0).

Recall the standard *cancellation trick*, that applies more generally to arbitrary finite groups:

$$\sum_{a \bmod N} \chi(a) = \begin{cases} \varphi(N) & (\text{for } \chi = \chi_o) \\ 0 & (\text{otherwise}) \end{cases}$$

where φ is Euler's *totient* function. Dirichlet's *dual* trick is to sum over characters $\chi \bmod N$ evaluated at fixed a in $(\mathbb{Z}/N)^\times$: we *claim* that

$$\sum_{\chi} \chi(a) = \begin{cases} \varphi(N) & (\text{for } a = 1 \bmod N) \\ 0 & (\text{otherwise}) \end{cases}$$

We will prove this in the next section.

Granting this, for b invertible modulo N ,

$$\sum_{\chi} \chi(a)\chi(b)^{-1} = \sum_{\chi} \chi(ab^{-1}) = \begin{cases} \varphi(N) & (\text{for } a = b \bmod N) \\ 0 & (\text{otherwise}) \end{cases}$$

Given a Dirichlet character χ modulo N , the corresponding *Dirichlet L-function* is

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

By the multiplicative property $\chi(mn) = \chi(m)\chi(n)$, each such L -function has an *Euler product* expansion

$$L(s, \chi) = \prod_{p \text{ prime}, p \nmid N} \frac{1}{1 - \chi(p)p^{-s}}$$

proven as for $\zeta(s)$, by expanding geometric series. Take a logarithmic derivative, as with zeta:

$$\frac{d}{ds} \log L(s, \chi) = \sum_{p \nmid N} \sum_{m \geq 1} \frac{\chi(p)^m \log p}{p^{ms}} = \sum_{p \nmid N} \frac{\chi(p) \log p}{p^s} + \sum_{p \nmid N, m \geq 2} \frac{\chi(p)^m \log p}{p^{ms}}$$

The second sum on the right will turn out to be subordinate to the first, so we aim our attention at the first sum, where $m = 1$.

To pick out the primes p with $p = a \bmod N$, use Dirichlet's sum-over- χ trick to obtain

$$\sum_{\chi \bmod N} \chi^{-1}(a) \cdot \frac{\chi(p) \log p}{p^s} = \begin{cases} \varphi(N) \cdot \frac{\log p}{p^s} & (\text{for } p = a \bmod N) \\ 0 & (\text{otherwise}) \end{cases}$$

Thus,

$$\begin{aligned} \sum_{\chi \bmod N} \chi^{-1}(a) \frac{d}{ds} \log L(s, \chi) &= \sum_{\chi \bmod N} \chi^{-1}(a) \sum_{p \nmid N, m \geq 1} \frac{\chi(p)^m \log p}{p^{ms}} \\ &= \varphi(N) \sum_{p=a \bmod N} \frac{\log p}{p^s} + \sum_{\chi \bmod N} \chi^{-1}(a) \sum_{p \nmid N, m \geq 2} \frac{\chi(p)^m \log p}{p^{ms}} \end{aligned}$$

We do not care about cancellation in the second sum. All that we need is its absolute convergence for $\operatorname{Re}(s) > \frac{1}{2}$, needing no subtle information about primes. Dominate the sum over primes by the corresponding sum over integers ≥ 2 . Namely,

$$\sum_{p \nmid N, m \geq 2} \left| \frac{\chi(p)^m \log p}{p^{ms}} \right| \leq \sum_{n \geq 2, m \geq 2} \frac{\log n}{n^{m\sigma}} = \sum_{n \geq 2} \frac{(\log n)/n^{2\sigma}}{1 - n^{-\sigma}} \leq \frac{1}{1 - 2^{-\sigma}} \sum_{n \geq 2} \frac{\log n}{n^{2\sigma}}$$

where $\sigma = \operatorname{Re}(s)$. This converges for $\operatorname{Re}(s) > \frac{1}{2}$. That is, for $s \rightarrow 1^+$,

$$\sum_{\chi \bmod N} \chi^{-1}(a) \frac{d}{ds} \log L(s, \chi) = \varphi(N) \sum_{p=a \bmod N} \frac{\log p}{p^s} + (\text{something continuous at } s = 1)$$

We have isolated the primes $p = a \bmod N$. Thus, as Dirichlet saw, to prove the infinitude of primes $p = a \bmod N$ it would suffice to show that the left-hand side of the last inequality blows up at $s = 1$. In particular, for the *trivial* character $\chi_o \bmod N$, with values

$$\chi(b) = \begin{cases} 1 & (\text{for } \gcd(b, N) = 1) \\ 0 & (\text{for } \gcd(b, N) > 1) \end{cases}$$

the associated L -function is essentially the zeta function, namely

$$L(s, \chi_o) = \zeta(s) \cdot \prod_{p \mid N} \left(1 - \frac{1}{p^s}\right)$$

Since none of those finitely-many factors for primes dividing N is 0 at $s = 1$, $L(s, \chi_o)$ still blows up at $s = 1$, like a non-zero constant multiple of $1/(s - 1)$.

By contrast, we will show below that for *non-trivial* character $\chi \bmod N$, $\lim_{s \rightarrow 1^+} L(s, \chi)$ is *finite*, and

$$\lim_{s \rightarrow 1^+} L(s, \chi) \neq 0$$

Thus, for non-trivial character, the logarithmic derivative is finite and non-zero at $s = 1$. Putting this all together, we will have

$$\lim_{s \rightarrow 1^+} \sum_{\chi \bmod N} \chi(a) \frac{d}{ds} \log L(s, \chi) = +\infty$$

Then necessarily

$$\lim_{s \rightarrow 1^+} \sum_{p=a \bmod N} \frac{\log p}{p^s} = +\infty$$

and there must be infinitely many primes $p = a \bmod N$. ///

[1.1] What remains to be done? The non-vanishing of the non-trivial L -functions at 1 is the crucial technical point left unfinished, and a place-holder proof appears in a Supplement. Dirichlet's dual cancellation trick is proven in the next section, as a consequence of Fourier analysis on finite abelian groups, the latter

treated in a supplement as a corollary of finite-dimensional spectral theory. We check below that the L -functions $L(s, \chi)$ have analytic continuations to regions including $s = 1$.

2. Dual groups of abelian groups

Dirichlet's use of group characters to isolate primes in a specified congruence class modulo N was a big innovation in 1837. These ideas were predecessors of the group theory work of Frobenius and Schur 50 years later, and one of the ancestors of *representation theory* of groups.

The *dual group* or *group of characters* \widehat{G} of a finite abelian group G is by definition

$$\widehat{G} = \{\text{group homomorphisms } \chi : G \rightarrow \mathbb{C}^\times\}$$

This \widehat{G} is itself an abelian group under the operation on characters defined for $g \in G$ by

$$(\chi_1 \cdot \chi_2)(g) = \chi_1(g) \cdot \chi_2(g)$$

There is an inner product $\langle \cdot, \cdot \rangle$ on complex-valued functions on G , given by

$$\langle f, F \rangle = \sum_{g \in G} f(g) \cdot \overline{F}(g) \quad (\text{for } f, g \text{ complex-valued on } G)$$

Let $L^2(G)$ refer to the space of complex-valued functions on G with this inner product.

Recall^[2] the following basic result on Fourier expansions on finite abelian groups:

[2.0.1] Theorem: For a finite abelian group G with dual group \widehat{G} , any complex-valued function f on G has a Fourier expansion

$$f(g) = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \widehat{f}(\chi) \chi(g) \quad (\text{for all } g \in G)$$

where the Fourier coefficients $\widehat{f}(\chi)$ are

$$\widehat{f}(\chi) = \sum_{g \in G} f(g) \overline{\chi}(g) = \langle f, \chi \rangle$$

The characters are an *orthogonal basis* for $L^2(G)$. In particular, Fourier coefficients are unique. ///

[2.0.2] Corollary: Let G be a finite abelian group. For $g \neq e$ in G , there is a character $\chi \in \widehat{G}$ such that $\chi(g) \neq 1$.^[3]

Proof: Suppose that $\chi(g) = 1$ for all $\chi \in \widehat{G}$. That is, $\chi(g) = \chi(e)$ for all χ . Then, for any coefficients c_χ ,

$$\sum_{\chi} c_\chi \chi(e) = \sum_{\chi} c_\chi \chi(g)$$

[2] Proven in a Supplement. This is really about *commuting unitary operators* on finite-dimensional complex vector spaces, and the main point is the spectral theorem for unitary operators, and the *simultaneous* diagonalization of commuting diagonal operators.

[3] This idea that characters can distinguish group elements from each other is just the tip of an iceberg.

Since every function on the group has such a Fourier expansion, this says that every function on G has the same value at g as at e . Thus, $g = e$. ///

[2.0.3] Corollary: For a finite abelian group G ,

$$|G| = |\widehat{G}|$$

Proof: The characters form an orthogonal basis for $L^2(G)$, so the number of characters is the dimension of $L^2(G)$, which is $|G|$. ///

[2.0.4] Remark: In fact, using the structure theorem for finite abelian groups, one can show that G and its dual are *isomorphic*, but this isomorphism is *not canonical*.

[2.0.5] Corollary: (*Dual version of cancellation trick*) For g in a finite abelian group,

$$\sum_{\chi \in \widehat{G}} \chi(g) = \begin{cases} |\widehat{G}| & (\text{for } g = e) \\ 0 & (\text{otherwise}) \end{cases}$$

Proof: If $g = e$, then the sum counts the characters in \widehat{G} . On the other hand, given $g \neq e$ in G , let χ_1 be in \widehat{G} such that $\chi_1(g) \neq 1$, from a previous corollary. The map on \widehat{G}

$$\chi \rightarrow \chi_1 \cdot \chi$$

is a bijection of \widehat{G} to itself, so

$$\sum_{\chi \in \widehat{G}} \chi(g) = \sum_{\chi \in \widehat{G}} (\chi \cdot \chi_1)(g) = \chi_1(g) \cdot \sum_{\chi \in \widehat{G}} \chi(g)$$

which gives

$$(1 - \chi_1(g)) \cdot \sum_{\chi \in \widehat{G}} \chi(g) = 0$$

Since $1 - \chi_1(g) \neq 0$, it must be that the sum is 0. ///

3. Appendix: analytic continuations

Dirichlet's original argument did not emphasize holomorphic functions, but by now we know that discussion of vanishing and blowing-up of functions is most clearly and simply accomplished if the functions are *meromorphic* when viewed as functions of a complex variable.

For the purposes of Dirichlet's theorem, it suffices to meromorphically continue^[4] the L -functions just a little, to $\text{Re}(s) > 0$. This limited analytic continuation allows a simpler argument than analytic continuation to the entire plane.

[4] An extension of a holomorphic function to a larger region, on which it may have some poles, is called a *meromorphic continuation*. There is *no* general methodology for proving that functions have meromorphic continuations, due in part to the fact that, generically, functions *do not* have continuations beyond some natural region where they're defined by a convergent series or integral. Indeed, to be able to prove a meromorphic continuation result for a given function is tantamount to proving that it has some deeper significance.

[3.0.1] Claim: The Dirichlet L -functions

$$L(s, \chi) = \sum_n \frac{\chi(n)}{n^s} = \prod_p \frac{1}{1 - \chi(p)p^{-s}}$$

have meromorphic continuations to $\operatorname{Re}(s) > 0$. For χ non-trivial, $L(s, \chi)$ is *holomorphic* on that half-plane. For χ trivial, $L(s, \chi_o)$ has a *simple* pole at $s = 1$ and is holomorphic otherwise.

Proof: First, to treat the trivial character $\chi_o \bmod N$, recall, as already observed, that the corresponding L -function differs in an elementary way from $\zeta(s)$, namely

$$L(s, \chi_o) = \zeta(s) \cdot \prod_{p|N} \left(1 - \frac{1}{p^s}\right)$$

Thus, analytically continue $\zeta(s)$ instead of $L(s, \chi_o)$. As earlier, to analytically continue $\zeta(s)$ to $\operatorname{Re}(s) > 0$ in an elementary way, observe that the sum for $\zeta(s)$ is fairly well approximated by

$$\zeta(s) - \frac{1}{s-1} = \sum_{n=1}^{\infty} \frac{1}{n^s} - \int_1^{\infty} \frac{dx}{x^s} = \sum_{n=1}^{\infty} \left[\frac{1}{n^s} - \frac{\left(\frac{1}{n^{s-1}} - \frac{1}{(n+1)^{s-1}}\right)}{1-s} \right]$$

Since

$$\frac{\left(\frac{1}{n^{s-1}} - \frac{1}{(n+1)^{s-1}}\right)}{1-s} = \frac{1}{n^s} + O\left(\frac{1}{n^{s+1}}\right)$$

with a uniform O -term, we obtain

$$\zeta(s) - \frac{1}{s-1} = \sum_n O\left(\frac{1}{n^{s+1}}\right) = \text{holomorphic for } \operatorname{Re}(s) > 0$$

The obvious analytic continuation of $1/(s-1)$ allows analytic continuation of $\zeta(s)$.

A similar relatively elementary analytic continuation argument for *non-trivial* characters uses *partial summation*. That is, let $\{a_n\}$ and $\{b_n\}$ be sequences of complex numbers such that the partial sums $A_n = \sum_{i=1}^n a_i$ are *bounded*, and $b_n \rightarrow 0$. Then it is useful to rearrange (taking $A_0 = 0$ for notational convenience)

$$\sum_{n=1}^{\infty} a_n b_n = \sum_{n=1}^{\infty} (A_n - A_{n-1}) b_n = \sum_{n=0}^{\infty} A_n b_n - \sum_{n=0}^{\infty} A_n b_{n+1} = \sum_{n=0}^{\infty} A_n (b_n - b_{n+1})$$

Taking $a_n = \chi(n)$ and $b_n = 1/n^s$ gives

$$L(s, \chi) = \sum_{n=0}^{\infty} \left(\sum_{\ell=1}^n \chi(\ell) \right) \left(\frac{1}{n^s} - \frac{1}{(n+1)^s} \right)$$

The difference $1/n^s - 1/(n+1)^s$ is s/n^{s+1} up to higher-order terms, so this expression gives a holomorphic function for $\operatorname{Re}(s) > 0$. ///