Introduction to Analytic Number Theory

via Dirichlet

A Reference Guide

Daniel K. Rui - August 26, 2021

Abstract

Recently (8/23/21 – see green 2021 math notebook) I became interested in getting a semi-serious first dose of analytic number theory (I do claim to have a deep enough interest in the subject to pursue graduate studies in it, but I have not actually seriously looked over it much detail).

1 Motivation

My starting point was 3b1b's video "Pi hiding in prime regularities", as it does introduce Dirichlet characters at the end and is an excellent problem for developing interesting ideas, and serving as a springboard for more complicated work. However, I did feel that the introduction of the Dirichlet character χ was just too slick/genius (as of now; let's hope someone writes a good answer to my MSE question), so I began looking into how Dirichlet came up with them. We proceed here introducing/motivating Dirichlet's work on prime in arithmetic progressions, mainly just leaving links and thoughts about those links (and I suppose the brilliant Gauss circle problem would be left as a supplementary/bonus section/exercise in this article).

Great article (https://arxiv.org/pdf/1404.4832.pdf – a lot of it is meta-mathematical/philosophy, but Section 3 talks about the math, and Appendix (pg. 48) talks about how Dirichlet may have come up with the idea of his characters, stemming from the study of Lagrange resolvents) detailing starting motivation of studying divergence of $\sum_{p\equiv a \bmod q} \frac{1}{p}$, by starting off with Euler product proof (we'll discuss that in a later section, Section 2) that there are infinitely many primes, i.e. proving that for $s > 1$ (using Taylor expansion of $\log x$):

$$\log\left(\sum_{n=1}^{\infty} \frac{1}{n^s}\right) = \sum_p -\log\left(1 - \frac{1}{p^s}\right) = \sum_p \frac{1}{p^s} + \sum_{n=2}^{\infty} \frac{1}{n} \sum_p \frac{1}{p^{ns}} = \sum_p \frac{1}{p^s} + O(1)$$

where the constant $O(1)$ is independent of s (which recall we are taking $s > 1$) because

$$\left|\sum_{n=2}^{\infty} \frac{1}{n} \sum_p \frac{1}{p^{ns}}\right| \leq \sum_{n=2}^{\infty} \frac{1}{n} \sum_p \frac{1}{p^n} \leq \sum_{n=2}^{\infty} \frac{1}{n}\left(\int_1^{\infty} x^{-n}\, dx\right) = \sum_{n=2}^{\infty} \frac{1}{n} \cdot \frac{1}{n-1} \leq \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

Then breaking down

$$\log\left(\sum_{n=1}^{\infty} \frac{1}{n^s}\right) = \sum_p \frac{1}{p^s} + O(1) = \sum_{a=1}^{q-1}\left(\sum_{p\equiv a \bmod q} \frac{1}{p^s}\right) + O(1)$$

The hope is that we can write "similar formulas" to this line, where instead of summing over all a with weight 1, we have different weights so that via clever linear combinations we can isolate $\sum_{p\equiv a \bmod q} \frac{1}{p}$ on the RHS. These "similar formulas" would then hopefully be finite on the LHS, or at the very least not $-\infty$ at $s=1$ so that the divergence of $\log\left(\sum_{n=1}^{\infty} \frac{1}{n}\right)$ would lead to the divergence of $\sum_{p\equiv a \bmod q} \frac{1}{p}$ on the RHS. Writing this idea out, we want something like

$$\text{mysterious LHS, hopefully not } -\infty \text{ at } 1 = \sum_{a=0}^{q-1} w(a)\left(\sum_{p\equiv a \bmod q} \frac{1}{p^s}\right) + O(1)$$

where w is a "weight function" defined on $\{0,\ldots,q-1\}$, mapping into anything $(\mathbb{Z},\mathbb{R},\mathbb{C})$ with the hope that with enough of these weight functions, doing linear combinations of them will result in cancellation in all but one $a \in \{0,\ldots,q-1\}$. In other words, we want enough $w : \{0,\ldots,q-1\} \to \mathbb{C}$ so that their span (in the vector space of these such functions) contain the functions $\delta_a : \{0,\ldots,q-1\} \to \mathbb{C}$ (1 at a, 0 everywhere else); note that as these δ_a form a basis of these such functions, this is equivalent to asking that our collection of w spans this vector space.

Also, note that in the $a=0$ case above, there is at most one prime $p \equiv 0 \bmod q \iff q$ divides p. Similarly, in the case that q is not prime and a,q share a factor other than 1 (i.e. $\gcd(a,q) \neq 1$), then there is again at most one prime $p \equiv a \bmod q$. Such a are not interesting to consider (the "at most one" prime can just be absorbed into the constant $O(1)$, not impacting anything at all), so we can simply set $w(a)=0$. In other words, we really only care about the values of w on $a \in \{0,\ldots,q-1\}$ relatively prime to q. The set of such a is otherwise known as group of units $(\mathbb{Z}/q\mathbb{Z})^\times$ (though right now only thought of as a set). Sometimes we will abbreviate $(\mathbb{Z}/q\mathbb{Z})^\times$ as $\mathbb{Z}_q^\times$. We may then think of w as either a function $w : (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}$, or as a q-periodic function $w : \mathbb{N} \to \mathbb{C}$ with zeroes at certain values; taking this latter viewpoint (i.e. the notion that w may be extended to all inputs $\mathbb{N}$), we may "bring the w inside" as follows:

$$\sum_{a\in\mathbb{Z}_q^\times}^{q-1} w(a)\left(\sum_{p\equiv a \bmod q} \frac{1}{p^s}\right) + O(1) = \sum_{a\in\mathbb{Z}_q^\times}\left(\sum_{p\equiv a \bmod q} \frac{w(p)}{p^s}\right) + O(1) = \sum_p \frac{w(p)}{p^s} + O(1)$$

This looks very similar to the above sum $\sum_p \frac{1}{p^s} + O(1) = \sum_p -\log\left(1-\frac{1}{p^s}\right)$, so is there a similar log expression for our weighted sum? Well, in that formula, we just used the Taylor expansion $-\log(1-x) = x + \frac{x^2}{2} + \frac{x^3}{3} + \ldots$ and took a sum over all primes p, meaning that if we set "x" to be instead $\frac{w(p)}{p^s}$, the sum over all primes $\sum_p \frac{w(p)}{p^s} + O(1)$ (again higher order terms, even in $|\bullet|$, shrink fast enough like above, assuming w bounded, which is totally reasonable assumption — could even bound $|w| \leq 1$ by dividing by bound B) would simply be the sum over all primes $\sum_p -\log\left(1-\frac{w(p)}{p^s}\right)$.

Can we go one step further back, to something like $\log\left(\sum_{n=1}^{\infty} \frac{1}{n^s}\right)$? Recall that the Euler product formula was noting that $\prod_p (1-p^{-s})^{-1} = \prod_p (1 + p^{-s} + (p^2)^{-s} + \ldots) = \sum_n n^{-s}$ because by the fundamental theorem of arithmetic every $n \in \mathbb{N}$ can be written as a product of primes in unique way $n = p_1^{e_1} \cdots p_k^{e_k}$, implying that every n on the RHS is attained exactly once (coefficient of n^{-s} for each $n \in \mathbb{N}$ equals 1). If we instead have $\prod_p (1 - w(p)p^{-s})^{-1} = \prod_p (1 + w(p)p^{-s} + w(p)^2(p^2)^{-s} + \ldots)$, again by the fundamental theorem of arithmetic each $n \in \mathbb{N}$ is attained on the RHS exactly once, except the coefficient of $n = p_1^{e_1} \cdots p_k^{e_k}$ is now $w(p_1)^{e_1} \cdots w(p_k)^{e_k}$.

In general (i.e. for general $w : (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}$), there is not much we can say about this product — basically this product takes the prime factorization of n and "hijacks" the primes p_i into $w(p_i)$, making the product highly dependent on the exact way n factors. However, this product LOOKS very similar to one that is much much simpler; SUPPOSING we could "pull" all these prime factors and exponents into one w, we would get very neatly $w(p_1)^{e_1} \cdots w(p_k)^{e_k} = w(p_1^{e_1} \cdots p_k^{e_k}) = w(n)$ (so going from an expression that requires knowledge of how exactly n factors, to an expression of just n). In this "dream scenario", we can write that for some (very hopefully extant!) "special" w,

$$\log\left(\sum_{n=1}^{\infty} \frac{w(n)}{n^s}\right) = \sum_p -\log\left(1 - \frac{w(p)}{p^s}\right) = \sum_p \frac{w(p)}{p^s} + \sum_{n=2}^{\infty} \frac{1}{n} \sum_p \frac{w(p^n)}{p^{ns}} = \sum_p \frac{w(p)}{p^s} + O(1).$$

The property of "pulling everything in" that we really wanted w to have is simply multiplicativity, i.e. $w(mn) = w(m)w(n)$ for all $m, n \in \mathbb{Z}$. This allows us to pull the exponents e_i inside $w(p_i)^{e_i} = w(p_i^{e_i})$, and also allow us to pull the product over all prime factors inside. So, to summarize, we have found that multiplicative functions $w : (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}$ seem to be a very promising direction of study in considering our question of primes in classes modulo q. It is remarkable that a condition as simple as multiplicativity allows us to write the extremely neat Euler product type formula above. MAJOR OBSERVATION: the multiplicative functions $w : (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}$ are EXACTLY the group homomorphisms $(\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}$, where we now think of $(\mathbb{Z}/q\mathbb{Z})^\times$ as a GROUP, not just a set. TIP: it's always a good sign in math when something from another field/problem shows up!

Using the fact that w is a group homomorphism $(\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}$, we see that $w(1) = 1$ because for any $a \in (\mathbb{Z}/q\mathbb{Z})^\times$, $w(a) = w(a \cdot 1) = w(a)w(1)$. Lagrange's theorem can also be used to prove that $w(a)^{|\mathbb{Z}_q^\times|} =: w(a)^{\varphi(q)} = w(a^{\varphi(q)}) = w(1) = 1$ (as indicated by the "=:", I'm defining the Euler totient function $\varphi(q) := |\mathbb{Z}_q^\times|$). Thus, there is a cute notational coincidence because we just proved that $w(a)$ must be a root of unity in $\mathbb{C}$, which are usually denoted by some variant of the symbol "ω". However, in the literature, group homomorphisms $G \to \mathbb{C}$ are called *group characters*, and specifically the group homomorphisms $(\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}$ are called *Dirichlet characters*, and so from here on out, we will use χ instead of w. Finally, we note that the "mysterious LHS, hopefully not $-\infty$ at 1" that appeared in an above displayed equation is in fact $\log\left(\sum_{n=1}^{\infty} \frac{w(n)}{n^s}\right)$; defining the *Dirichlet L-function* $L(s, \chi) := \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$, we see the "hopefully not $-\infty$ at 1" refers to the hope/conjecture that $L(1, \chi) \neq 0$ for Dirichlet characters χ.

From this much, we already see an OUTLINE. Somehow, we must understand these Dirichlet characters well enough to see if they EXIST first of all, and if they do, see if $\delta_a : (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}$ can be written as a linear combination of such characters, or equivalently if these characters span (i.e. linearly combine into) the $\mathbb{Z}/\mathbb{R}/\mathbb{C}$-vector space of set-functions $(\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}$. This is the CHARACTER THEORY portion of the proof. We then must understand the L-functions $L(s, \chi)$ well enough to determine that $L(1, \chi)$ is non-vanishing. This is the COMPLEX ANALYSIS portion.

2 Euler Product

Before we do all that, I want to say a few words on the Euler product. *On formulas that yield information about additive decomposition*. Easiest/earliest entry point is related to the identity $(a^2 - b^2) = (a-b)(a+b)$, more specifically $x^2-1 = (x-1)(x+1)$ and the geometric series. Upon seeing the formula $(a^2-b^2) = (a-b)(a+b) \implies x^2-1 = (x-1)(x+1)$, a high-schooler (such as myself; I do remember doing this) may see that with powers of 2 in exponent this formula can be iterated to yield $x^{2^n} - 1 = (x-1)(1+x)(1+x^2)(1+x^4)(1+x^8)\cdots(1+x^{2^{n-1}})$. One may recall from geometric series (literally the first series a high-schooler would learn) that $\frac{x^n-1}{x-1} = 1+x+x^2+\ldots+x^{n-1}$. Combining the two ideas/formulas/expressions we have

$$1+x+x^2+\ldots+x^{2^n-1} = \frac{x^{2^n}-1}{x-1} = (1+x)(1+x^2)(1+x^4)(1+x^8)\cdots(1+x^{2^{n-1}})$$

Already note of interest, we have a pure sum on the LHS and product on RHS. Also, every exponent $\{1,\ldots,2^n-1\}$ appears on the LHS with coefficient exactly 1. I think I can be confident in saying that even without understanding particularly deep math, this is very eye-catching at first glance. But why is this equation true (conceptual proof instead of algebraic manipulation)? When we multiply out the RHS, we see that this equation corresponds to the fact that every number $\{1,\ldots,2^n-1\}$ can be written in a unique way as 1/0-weighted sum of powers of 2 $\{1,2,\ldots,2^{n-1}\}$. In other words, I have n choices (i.e. digit places) where at choice i I choose between $+2^i$ or $+0$ (corresponding to n factors $(1+x)(1+x^2)(1+x^4)(1+x^8)\cdots(1+x^{2^{n-1}})$ where each factor has choice between x^{2^i} or x^0).

So exponents of this "dummy variable" x encode information about additive decomposition based on some "basis elements" $\{0 \text{ or } 1, 0 \text{ or } 2, 0 \text{ or } 4, \ldots\}$ (the "or" is exclusive, which we henceforth denote "xor" for "exclusive or"). The reason why I call x a "dummy variable" is because it is actually not doing anything/representing anything, it is just sort of serving as a "clothes line" on which the "clips" (the exponents of x) are interesting. Three more examples: for base-3 decomposition, we have "basis elements" $\{0 \text{ xor } 1 \text{ xor } 2, 0 \text{ xor } 3 \text{ xor } 2\cdot 3, 0 \text{ xor } 3^2 \text{ xor } 2\cdot 3^2, \ldots\}$, in other words $(1+x+x^2)(1+x^3+x^6)(1+x^9+x^{18})\cdots = 1+x+x^2+x^3+\ldots$ and so on (again coefficient one means unique decomposition in base-3).

Next two examples very similar: consider product $(1+x)(1+x^2)(1+x^3)\cdots$ corresponding to basis elements $\{0 \text{ xor } 1, 0 \text{ xor } 2, 0 \text{ xor } 3, \ldots\}$ results in a polynomial/formal power series (i.e. infinite clothes line) where coefficient of x^n is the number of ways to write n as sum of distinct numbers. For coefficient of x^n equal to number of ways to partition n into positive integers, we want for instance the option to have multiple 1's; however, if 1 appears in multiple "basis elements", say in three "basis elements", there are $\binom{3}{2}$ ways of choosing two 1's. So then all possible number of 1's must appear in one "basis element", and indeed as each partition can have either zero 1's XOR one 1's XOR two 1's XOR etc. we see that the basis $\{0 \text{ xor } 1 \text{ xor } 1+1 \text{ xor } \ldots, 0 \text{ xor } 2 \text{ xor } 2+2 \text{ xor } \ldots, \ldots\}$ corresponding to $(1+x+x^2+\ldots)(1+x^2+x^4+\ldots)(1+x^3+x^6+\ldots)\cdots$ and so on. In other words we have encoded the partition number $p(n)$ in the formula above, which we can write succinctly (think of it just as notation representing above product) as $\sum_{n=1}^{\infty} p(n)x^n = \prod_{n=1}^{\infty} \frac{1}{1-x^n}$.

Because multiplying x^m, x^n adds exponent, above techniques allow us to understand some combinatorics of additive decompositions in terms of coefficient of resulting formal power series. For multiplicative decomposition, exponent not useful, so we forget about the x entirely. But the idea/principle remains. We know that $n \in \mathbb{N}$ can be factored uniquely into product of primes, so in the spirit of our say base-2 additive decomposition above, we have for $\mathbb{N}$ a *multiplicative* decomposition based on some "basis elements" $\{1 \text{ xor } 2 \text{ xor } 2^2 \text{ xor } \ldots, 1 \text{ xor } 3 \text{ xor } 3^2 \text{ xor } \ldots, 1 \text{ xor } 5 \text{ xor } 5^2 \text{ xor } \ldots, \ldots\}$ corresponding to $(1+2+2^2+2^3+\ldots)(1+3+3^2+\ldots)(1+5+5^2+\ldots)(1+7+\ldots)\cdots = 1+2+3+4+\ldots$ and so on.

This makes sense from a conceptual point of view — for any $n \in \mathbb{N}$, going up on the LHS to a large enough but finite number of "basis elements" with large but finite "length" for each "basis element" (so like length l meaning $[1 + 2 + 2^2 + \ldots + 2^l]$), we see that n appears; and indeed no matter how many basis elements you take and how long you take each basis element n appears ONLY ONCE. At face value though both sides are ∞. No worries, we can still use this EXACT principle to encode this same fundamental theorem of arithmetic fact but with convergence on both sides; simply consider

$$\prod_p \frac{1}{1 - \left[\frac{1}{p^2}\right]} = \prod_p \left(1 + \left[\frac{1}{p^2}\right] + \left[\frac{1}{p^2}\right]^2 + \ldots\right) = \sum_{n=1}^{\infty} \frac{1}{n^2}$$

Indeed, this works for any exponent n^{-s} for $s > 1$. There's probably more to say regarding convergence of sum and product, and rigorously proving this equality, but really this document is more of a guide than an actual account of the detailed proof.

3 Preliminary Concrete Examples

https://abel.math.harvard.edu/~elkies/M259.02/dirichlet.pdf talks a bit about concrete examples of $q = 4, 8$.

4 Character Theory (References)

Almost every source I looked at had a section on this. It's covered in first halves of https://math.mit.edu/classes/18.785/2015fa/LectureNotes17.pdf (18.785 Number theory I Lecture #17 Fall 2015 11/10/2015), https://fse.studenttheses.ub.rug.nl/17834/1/bMATH_2018_MintjesMW.pdf ("The Proof of Dirichlet's Theorem on Arithmetic Progressions and its Variations", M. W. Mintjes Bachelor's Project Mathematics, University of Groningen — beautifully typeset!). Basically a form of Fourier analysis, as the characters are orthogonal basis of $L^2(G)$ — see also https://www-users.cse.umn.edu/~garrett/m/mfms/notes_c/dirichlet.pdf ("Primes in arithmetic progressions" (April 12, 2011) Paul Garrett). For textbook, I'm sure Dummit & Foote has section on character theory; really any algebra/representation theory textbook.

5 Complex Analysis (References)

https://mathoverflow.net/a/26096/112504 provides sketch: assume $L(1,\chi) = 0$; then because $\zeta(s)$ has pole at $s = 1$, product $\zeta(s)L(s,\chi)$ should extend nicely past $s = 1$. Coefficients of Dirichlet series for this product are $c(n) = \sum_{d|n} \chi(d)$. Converge absolutely on $\mathrm{Re}(s) > 1$, but because $\zeta(s)L(s,\chi)$ extend analytically to $\mathrm{Re}(s) > 0$, the Dirichlet series converge conditionally for $\mathrm{Re}(s) > 0$ (LANDAU THEOREM). However, $c(n^2) > 0$, and $\sum_{n=1}^{\infty} c(n)n^{-1/2}$, leading to contradiction.

Requires knowledge that $\zeta(s), L(s,\chi)$ extend meromorphically/analytically to $\mathrm{Re}(s) > 0$. Refer to Terry Tao's Notes 2 of his 254A 2014 blog post; ctrl-F "meromorphic extension" and "holomorphic on this region". Basically the method is to prove the local uniform convergence of partial sums. Terry's whole Notes 2 is VERY long and chock full of information: https://terrytao.wordpress.com/2014/12/09/254a-notes-2-complex-analytic-multiplicative-number-theory/.

OVERVIEW/REVIEW OF MANY DIFFERENT PROOFS: https://fse.studenttheses.ub.rug.nl/17834/1/bMATH_2018_MintjesMW.pdf ("The Proof of Dirichlet's Theorem on Arithmetic Progressions and its Variations", M. W. Mintjes Bachelor's Project Mathematics, University of Groningen) referenced above in character theory section, has SECTION 4 devoted to proving the complex analysis portion in a different way then what author did in SECTION 3 (SECTION 4 proof similar to above sketched proof from MathOverflow). SECTION 5 reviews different proofs, pointing out strenghts/weaknesses.

More Terry: https://mathoverflow.net/a/29435/112504 and then blog post https://terrytao.wordpress.com/2009/09/24/the-prime-number-theorem-in-arithmetic-progressions-and-dueling-conspiracies/. More on analytic continuation of $\zeta(s)$: https://mathoverflow.net/questions/58004/how-does-one-motivate-the-analytic-continuation-of-the-riemann-zeta-function. For REALLY heavy duty machinery complex analysis, we have Terry's ENORMOUS blog post on functional equations + gamma, beta, digamma functions + reflection/duplication/multiplication + Poisson summation + Fourier inversion https://terrytao.wordpress.com/2014/12/15/254a-supplement-3-the-gamma-function-and-the-functional-equation-optional/.

https://math.uchicago.edu/~may/REU2012/REUPapers/LiAng.pdf ("DIRICHLET'S THEOREM ABOUT PRIMES IN ARITHMETIC PROGRESSIONS" Ang Li) contains proof of this complex analysis portion using a lot of really hands on bounding. May be nice approach if one doesn't want to introduce so much complex analysis theory.

Another approach of interest (though probably more complicated/messy) is Mathologer's "super sum" approach in his masterclass video "Numberphile v. Math: the truth about 1+2+3+...=-1/12".

6 Algebraic Number Theory

Part of the reason why I love the 3b1b video "Pi hiding in prime regularities" that I started off with in Section 1 is because all that work with factorization in $\mathbb{Z}[i]$ is the first step into algebraic number theory, and connections with quadratic fields, and Dirichlet's class number formula. I don't know anything about any of that now (maybe good reference is `https://people.reed.edu/~jerry/361/lectures/iqclassno.pdf` "THE DIRICHLET CLASS NUMBER FORMULA FOR IMAGINARY QUADRATIC FIELDS" Jerry Shurman Reed College), but I do know enough that I can see what a brilliant problem this Gauss circle problem is. All the more crucial that I find a good answer to my MSE question "How would one motivate/know to introduce the Dirichlet character in the formula for the number of lattice points on a circle of radius $\sqrt{N}$".

And about imaginary quadratic fields, I recently heard about a professor at CU, Katherine Stange (`https://math.katestange.net/` — brilliant website, well designed, nice to look at, full of resources for students, visualizations, obviously the work of someone who loves teaching) who has some work on visualizing them: `http://math.colorado.edu/~kstange/papers/Stange-short-exp.pdf` ("Visualizing imaginary quadratic fields") with a tiny bit of explanation in the form of a Reddit thread `https://www.reddit.com/r/math/comments/2xs4t7/visualising_complex_quadratic_number_fields/`; see also her fully-fledged paper on the topic `https://arxiv.org/pdf/1410.0417.pdf` ("Visualising the arithmetic of imaginary quadratic fields").

7 Collected Links

If one would like to have most of these above mentioned links in one list, see `https://www.one-tab.com/page/uLC4YXuWTOWZxMWfER2_qg`.

As concluding words, I mention Andrew Granville's excellent survey/overlook of basically the entire subject of analytic number theory (asymptotics, large/small prime gaps, sieve methods, circle method, Selberg/Langlands class of L-functions, etc.) `https://dms.umontreal.ca/~andrew/PDF/PrinceComp.pdf`.