Number Theory

Jacob Denson

March 16, 2018

Table Of Contents

1 The Prime Numbers 2

2 Congruences 5
2.1 Submonoids of the Natural Numbers 5
2.2 Systems of Linear Congruences 6

3 Diophantine Approximation 8

I Analytic Number Theory 10

4 Dirichlet Series 12
4.1 Dirichlet Convolutions and Euler Products 18
4.2 Smooth Numbers 19
4.3 Gauss Sums for Quadratic Characters 21

Chapter 1

The Prime Numbers

Number theory is the study of the positive integers, those numbers you know as

$$1, 2, 3, \ldots$$

The most basic relation between these numbers is that of divisibility. An integer a is divisible by b, denoted $b \mid a$, if there is a number n for which $nb = a$. Any number n has divisors 1 and itself. Of particular interest are the primes, integers greater than 1, whose divisors consist of only itself and one. The first few examples are

$$2, 3, 5, 7, 11$$

The numbers that are left over once we remove all prime numbers are called composite. It is of great importance that one may 'compose' prime numbers to form all the composite numbers.

Theorem 1.1. *Every integer can be written as a product of prime numbers.*

Proof. If n is a prime number, then it can obviously be written as a prime. Otherwise, we may write $n = ab$, for $1 < a, b < n$. Continuing this expansion process, we may continue to expand a and b as a product of smaller numbers. Eventually these smaller numbers must be prime, for otherwise we would have an infinite decreasing chain of positive integers, of which we know the impossibility. Thus we have prime decompositions $a = p_1 p_2 \ldots p_n$, and $b = q_1 q_2 \ldots q_m$, and then $n = p_1 \ldots p_n q_1 \ldots q_m$. □

An interesting fact to notice is that if $n = ab$, then either $a \leqslant \sqrt{n}$ or $b \leqslant \sqrt{n}$. Thus every composite number is divisible by a prime number

smaller than the composite's square root. This leads to a simple procedure for finding all primes up to a certain number M. We first write down the integers

$$2, 3, 4, \ldots, M$$

and cross off all numbers divisible by 2 (all even numbers). We end up with the list

$$3, 5, 7, 9, 11, \ldots$$

Now we cross off all numbers divisible by 3. Any number which eventually ends up at the beginning of the queue must be prime, for it is not divisible by any prime smaller than it. If we continue to cross of numbers divisible by the first primes, we will find all primes. We may stop once we reach an integer bigger than $\sqrt{M}$, for if a number has not been crossed off at this point, it is not divisible by any number less than the square root of n, it must be prime. The number of operations to perform this procedure is therefore proportional to the sum of reciprocol primes

$$\sum_{p \leqslant \sqrt{M}} \frac{M}{p}$$

which is $O(M\pi(\sqrt{M}))$, where $\pi(n)$ counts the number of primes less than or equal to n. We will eventually show that $\pi(n) \sim n/\log(n)$, so that our algorithm is $O(M^{3/2}/\log(M))$. A tighter analysis can show this algorithm actually runs in $\Theta(\sqrt{M}\log\log M)$ time.

A particular decomposition of a composite number is not necessarily unique, because we can just rearrange the prime numbers

$$2 \cdot 3 = 3 \cdot 2$$

But we shall soon know that this is the only problem we can have. We shall assume all future decompositions

$$p_1^{n_1} \ldots p_m^{n_m}$$

are in standard for, with $p_1 < p_2 < \cdots < p_m$. That there is only one decomposition of each number composes exactly what is commonly known as the fundamental theorem of arithmetic, but is a bit tricky to prove formally.

Before our endeavor, however, we answer a fundamental question about the primes. Are there infinitely many of them? It is entirely possible that we have some finite set of primes. The very first proof in all of number theory shows this is not the case.

Theorem 1.2 (Euclid)**.** *There are infinitely many prime numbers.*

Proof. Let $p_1, \ldots, p_n$ be a finite collection of prime numbers, and consider the number

$$n = p_1 \ldots p_n + 1$$

Then n is not divisible by p_1, $p_2, \ldots, p_n$, because, dividing by the p_i leaves a remainder of 1. But n must be divisible by a prime, so there is some prime not among the p_i, and so no finite subset of the primes exhausts the set. □

This theorem also gives us bounds on how spread apart the prime numbers are. If $p_1, p_2, \ldots, p_n$ are all primes from 1 to n, then there is a prime between p_n and $p_1 \ldots p_n + 1$.

To start with, we essentially prove we can perform long division on $\mathbf{N}$.

Lemma 1.3. *If $n, m \in \mathbf{N}$, then we may write $m = ln + r$, where $r < n$.*

Proof. If $m < n$, the proof is trivial. Otherwise, write $m' = m - n$, apply induction, and write $m' = l'n + r$. Then $m = (l' + 1)n + r$. □

Theorem 1.4. *Every integer has a unique decomposition in standard form.*

Proof. We shall rely on a useful property, to be proved later. If p is prime, and $p \mid ab$, then $p \mid a$ or $p \mid b$. Now suppose that

$$p_1^{n_1} \ldots p_m^{n_m} = q_1^{k_1} \ldots p_l^{k_l}$$

Now $p_i \mid q_1^{k_1} \ldots q_l^{k_l}$ for each i, so $p_i \mid q_j$ for some j, hence $p_i = q_j$. Since the p_i are distinct, the q_j must also be distinct, so $m \leqslant l$. By symmetry (for we may perform the same technique with the q_i), $m = l$. For each i, we must have $n_i = k_i$, for if $n_i < k_i$, we may write

$$p_1^{n_1} \ldots p_i^0 \ldots p_m^{n_m} = p_1^{k_1} \ldots p_i^{k_i - n_i} \ldots p_m^{k_m}$$

and p_i divides the right hand side, but not the left hand side, a contradiction. □

Chapter 2

Congruences

2.1 Submonoids of the Natural Numbers

Our results about the greatest common denominator immediately have applications to subsets of the natural numbers closed under addition, semigroups. Let X denote an arbitrary subset of the natural numbers closed under addition, and let d denote the greatest common denominator of X.

Theorem 2.1. *X contains all but finitely many of $d\mathbf{N}$*

Proof. Dividing every element of X by d, it suffices to show that if the greatest common denominator of X is one, then X contains all but finitely many natural numbers. If we take a finite subset $x_1, \ldots, x_n \in X$ such that $\gcd(x_1, \ldots, x_n) = 1$, then there are integers $a_1, \ldots, a_n \in \mathbf{Z}$ such that $\sum a_i x_i = 1$. Consider $M = \sum |a_i| x_i$. We claim that X contains all numbers greater than or equal to M^2. Given $0 \leqslant N < M$, we can write

$$M^2 + KM + N = \sum [(M + K)|a_i| - N a_i] x_i$$

and $\sum (M+K)|a_i| - N a_i \geqslant (M+K-N)|a_i| \geqslant (M-N)|a_i| \geqslant 0$, so M^2+KM+N is a positive sum of elements of X, and therefore $M^2 + KM + N \in X$. □

Corollary 2.2. *Every submonoid of the natural numbers is finitely generated.*

Example. *The upper bound M^2 is essentially tight for the natural numbers. If we consider the set $x\mathbf{N} + (x+1)\mathbf{N}$, and if $n = ax + b(x+1) = (a+b)x + b$, where $n \equiv x - 1$ modulo x, then $b \equiv x - 1$ modulo x, and so $b \geqslant x - 1$, in which*

case we conclude $n \geqslant (x-1)(x+1)$. It follows that if N is any number chosen large enough that $N, N+1, \cdots \in x\mathbf{N} + (x+1)\mathbf{N}$, then there is $0 \leqslant k < x$ with $N + k \equiv x - 1$ modulo x, and so

$$N \geqslant (x-1)(x+1) - k \geqslant (x-1)(x+1) - x = x^2 - x - 1$$

But in this case we have $(x+1)-x = 1$, so $M = 2x+1$, and $M^2 = 4x^2+4x+1$, and so the upper bound is tight up to a constant.

2.2 Systems of Linear Congruences

The general recurrence relation $ax \equiv b \pmod{n}$ is easily solved in the general theory. If $\gcd(a,n) \mid b$, then we can write $b = m(at + nu)$, and if we define $x = mt$, then $ax \equiv b$. There are $\gcd(a,n)$ different solutions to this equation modulo n, given by

$$x \qquad x + \frac{n}{\gcd(a,n)} \qquad x + 2\frac{n}{\gcd(a,n)} \qquad \dots \qquad x + (\gcd(a,n)-1)\frac{n}{\gcd(a,n)}$$

The number of solutions is the same as the size of the kernel of the homomorphism from $\mathbf{Z}_n$ given by $x \mapsto ax$, and this contains $n/\gcd(a,n)$ elements, because this is just the order of a. In particular, if a and n are relatively prime, then the equation has a unique solution.

Now we consider the more general problem of solving a system of linear congruences. We want to find x such that

$$\begin{aligned} a_1 x &\equiv b_1 \pmod{n_1} \\ a_2 x &\equiv b_2 \pmod{n_2} \\ &\cdots \\ a_m x &\equiv b_m \pmod{n_m} \end{aligned}$$

Using the prior problem, the problem is unsolvable unless $\gcd(a_i, n_i) \mid b_i$. Then we can find separate x_i such that $a_i x_i \equiv b_i$. The problem then reduces to finding a set of c_i such that $c_i \equiv 1 \pmod{b_i}$ and $c_i \equiv 0 \pmod{b_j}$, for we can then let $x = c_1 x_1 + c_2 x_2 + \cdots + c_m x_m$. If the n_i are pairwise relatively prime (we say they are coprime), finding c_i is easy; if we set $N_i = \prod_{j \neq i} n_j$, then there is t and u such that $tn_i + uN_i = 1$. We can then set $c_i = uN_i$. Any other choice of c_i' differs by a multiple of $n_1 \dots n_m$, because we must then have $n_i \mid c_i - c_i'$ for each i, and by coprimality $n_1 \dots n_m \mid c_i - c_i'$.

Theorem 2.3. *If the n_i are coprime, then every system of linear equations has a solution, and this solution is unique modulo $n_1 \dots n_m$. In terms of ring theory, the projection map establishes an isomorphism*

$$\mathbf{Z}_{n_1 \dots n_m} \cong \mathbf{Z}_{n_1} \times \mathbf{Z}_{n_2} \cdots \times \mathbf{Z}_{n_m}$$

If the n_i are not coprime, the problem becomes more complicated.

Chapter 3

Diophantine Approximation

The rational numbers **Q**, most closely connected to the integers, form a dense subset of the set **R** of all real numbers. The field of diophantine approximation deals with how well one can approximate **R** by particular rational numbers. Because the rational numbers with a fixed denominator q divide the interval into length $1/q$ segments, there exists a rational number p/q such that $|x - p/q| \leqslant 1/2q$. The first interesting result was of Dirichlet, which was also historically one of the first uses of the pidgeonhole principle.

Theorem 3.1 (Dirichlet's Approximation Theorem)**.** *If x is irrational, there exists infinitely many rational numbers p/q such that $|x - p/q| \leqslant 1/q^2$.*

Proof. Given a rational number a, define $\langle a \rangle = a - \lfloor a \rfloor$ to be the **fractional part** of a, which lies in the range $[0,1)$. Fix some large integer N, and consider the numbers $\langle x \rangle, \langle 2x \rangle, \ldots \langle (N+1)x \rangle$. The interval $[0,1)$ divides itself into N parts

$$\left[0, \frac{1}{N}\right), \left[\frac{1}{N}, \frac{2}{N}\right), \ldots, \left[\frac{N-1}{N}, 1\right)$$

The pidgeonhole principle then guarantees that there are two values $\langle nx \rangle$ and $\langle mx \rangle$ lying in the same of these intervals, so that $|\langle nx \rangle - \langle mx \rangle| < 1/N$. This can be reexpressed as $|(n-m)x - (\lfloor nx \rfloor - \lfloor mx \rfloor)| \leqslant 1/N$, or

$$\left| x - \frac{\lfloor nx \rfloor - \lfloor mx \rfloor}{n-m} \right| \leqslant \frac{1}{N(n-m)} \leqslant \frac{1}{(n-m)^2}$$

Denote this first approximation of x by p_1/q_1, and write $|x-p_1/q_1| = \alpha_1$. If x is irrational, then $\alpha_1 > 0$, and if we choose $N > 1/\alpha_1$, we may find p_2/q_2 with $q_2 \leqslant N$ and

$$|x - p_2/q_2| \leqslant 1/Nq_2 \leqslant 1/q_2^2$$

because $1/Nq_2 < \alpha_1$, $p_2/q_2 \neq p_1/q_1$. We may continue this process to find distinct rational numbers $p_1/q_1, p_2/q_2, \ldots$, proving the theorem. □

On the other hand, most irrational numbers cannot be approximated to a rational number p/q to a factor of $1/q^3$, which follows from this simple theorem, using the fact that the intervals covered by all the rational numbers form a very small cover of the real line. Call an irrational number x 3-approximatible if there are infinitely many rational numbers p/q with $|x - p/q| \leqslant 1/q^3$.

Theorem 3.2. *The family of 3 approximatible functions is a set of Lebesgue measure zero.*

Proof. Note that if we take a rational number p/q, whose numerator and denominator is simplified, then the bound for the 3 approximations loosens because the denominator decreases, so we can approximate x by p/q if and only if x is in the interval $I_{p/q} = [p/q - 1/q^3, p/q + 1/q^3]$, where we assume p and q are relatively prime. We prove that the 3 approximatible numbers in $[0,1]$ form a set of measure zero, for then, since $x > 1$ is approximatible if and only if $x - 1$ is approximatible, proves the general result.

If $p > q$, then $p/q - 1/q^3 \leqslant 1 + 1/q - 1/q^3 \geqslant 1$, so $I_{p/q} \cap [0,1] \subset \{1\}$, and if $p < 0$, then $p/q + 1/q^3 \leqslant 1/q^3 - 1/q \leqslant 0$, so $I_{p/q} \cap [0,1] \subset \{0\}$. Thus the only time when $I_{p/q} \cap [0,1]$ is a set of positive measure is if $0 \leqslant p \leqslant q$. But in this case then $|Ip/q| = 2/q^3$, and

$$\sum_{q=1}^{\infty} \sum_{\substack{p \leqslant q \\ \gcd(p,q)=1}} |I_{p/q}| = \sum_{q=1}^{\infty} \frac{2\varphi(q)}{q^3} \leqslant \sum_{q=1}^{\infty} \frac{2}{q^2} < \infty$$

and so the Borel-Cantelli lemma implies that the set $E = \limsup I_{p/q}$ is a set of measure zero, and this is precisely the set of 3-approximatible numbers. □

Part I

Analytic Number Theory

The main paradigm to analytic number theory is that one may obtain much about the asymptotic properties of a sequence of numbers $a_1, a_2, \ldots$ by studying the properties of certain functions obtained by infinite series in terms of the a_n. You may have already encountered the power series

$$\sum_{n=0}^{\infty} a_n z^n$$

The identity

$$\left(\sum_{n=0}^{\infty} a_n z^n\right)\left(\sum_{n=0}^{\infty} b_n z^n\right) = \sum_{n=0}^{\infty}\left(\sum_{m=0}^{n} a_m b_{n-m}\right) z^k$$

establishes a relation between products of power series and *additive convolutions* of the coefficient. This indicates that power series will yield information about sequences with good additive structure. A series you may have encountered less in other areas of analysis is the Dirichlet series

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

where s is a complex number (In the analysis of Dirichlet series, we often assume the strange notation that a general complex variable is written as $s = \sigma + it$). We expect, at least formally, for

$$\left(\sum_{n=1}^{\infty} a_n n^{-s}\right)\left(\sum_{n=1}^{\infty} b_n\right) = \sum_{n,m} a_n b_m (nm)^{-s} = \sum_{n=1}^{\infty}\left(\sum_{mk=n} a_m b_k\right) n^{-s}$$

Thus we should expect Dirichlet series to give lots of information for arithmetic functions which behave well multiplicatively under the *Dirichlet convolution*, defined for two functions $f, g : \mathbf{N} \to \mathbf{C}$ by

$$(f * g)(n) = \sum_{mk=n} f(m) g(k)$$

The study of the analytic functions that relate to the theory of numbers is known as the field of *analytic number theory*.

Chapter 4

Dirichlet Series

Having motivated the importance that Dirichlet series play in understanding multiplicative functions, we now establish their analytic properties. We begin by recalling a fundamental trick which we will use again and again, known as partial summation. Consider any two sequences of complex numbers

$$a_1, \ldots, a_N \quad b_1, \ldots, b_N$$

If we let $A(x) = \sum_{m \leqslant x} a_m$, $B(x) = \sum_{m \leqslant x} b_m$, then for any M,

$$\sum_{n=M}^{N} a_n b_n = a_N B(N) - a_M B(M-1) - \sum_{n=M}^{N-1} (a_{n+1} - a_n) B(n)$$

this is often used when we consider the pointwise product of two series, where one has an explicit summation formula. It is an analogy of the concept of integration by parts in calculus. If we recall the theory of Stieltjes integration, which defines a weighted integral

$$\int_a^b f(x)\, dg(x)$$

for any function g of bounded variation and continuous functions f on $[a, b]$. The formula satisfies an integration by parts; if the integral of f with respect to g exists, then the integral of g with respect to f exists, and

$$\int_a^b f(x)\, dg(x) = f(b)g(b) - f(a)g(a) - \int_a^b g(x)\, df(x)$$

which is the Lebesgue integral with respect to the measure μ_g satisfying $\mu_g((x,y]) = g(y) - g(x)$. The integral satisfies an integration by parts formula, which holds if f or g is continuous (which will be satisfied whenever we use the formula), such that

$$\int_a^b f(x)\, dg(x) = f(b+)g(b+) - f(a-)g(a-) - \int_a^b g(x)\, df(x)$$

If you're familiar with Riemann-Stieltjes integration, then the summation is really integration by parts, since, if we define the limiting integral

$$\int_{M-}^{N} = \lim_{t \uparrow M} \int_t^N$$

then

$$\sum_{n=M}^{N} a_n b_n = \int_{M-}^{N} a_t dB_t = a_t B_t|_{M-}^{N} - \int_{M-}^{N} B_t da_t = a_t B_t|_{M-}^{N} - \sum_{n=M}^{N-1} B_n(a_{n+1} - a_n)$$

Often this formula is useful when a_t is differentiable in t, in which case $da_t = a_t' dt$, and so the right hand side has a normal integral, which is normally easier to estimate than a discrete series.

Theorem 4.1. *If $\alpha(s) = \sum a_n n^{-s}$ converges pointwise as $s = s_0$, and $H > 0$, then $\alpha(s)$ is uniformly convergent in the section*

$$S = \{s : \sigma \geqslant \sigma_0, |t - t_0| \leqslant H(\sigma - \sigma_0)\}$$

Taking $H \to \infty$, we see α converges locally uniformly for all s with $\sigma > s_0$ to an analytic function.

Proof. Let

$$R(u) = \sum_{n>u} a_n n^{-s_0}$$

be the remainder of higher order terms of $\alpha(s_0)$. Note that $a_n = (R(n-1) - R(n))n^{s_0}$, which means that

$$\begin{aligned}\sum_{n=M+1}^{N} a_n n^{-s} &= -\int_M^N x^{s_0-s} dR(x) = \int_M^N R(x) dx^{s_0-s} + M^{s_0-s}R(M) - N^{s_0-s}R(N) \\ &= (s_0 - s)\int_M^N R(x) x^{s_0-s-1} du + M^{s_0-s}R(M) - N^{s_0-s}R(N)\end{aligned}$$

If $R(u) \leqslant \varepsilon$ for all $u \geqslant M$, then this implies

$$\left| \sum_{n=M+1}^{N} a_n n^{-s} \right| \leqslant \varepsilon |s_0 - s| \int_M^N x^{\sigma_0 - \sigma - 1} + \varepsilon M^{s_0 - s} - N^{s_0 - s} R(N) = \varepsilon \left(\frac{|s_0 - s|}{\sigma - \sigma_0} + 2 \right)$$

Since $|s_0 - s| \leqslant |t_0 - t| + |\sigma_0 - \sigma|$, these terms are uniformly small if $|t_0 - t| \leqslant H|\sigma_0 - \sigma|$ for a fixed H. □

An immediate corollary is that the points where α converges lie in a half plane, which we denote by $\sigma = \sigma_c$, where $\sigma_c \in \mathbf{R} \cup \{\pm\infty\}$ is known as the *abscissa of convergence*. Another corollary is that we know α is complex analytic for $\sigma > \sigma_c$, and therefore the α' is given by the Dirichlet series

$$\alpha'(s) = -\sum_{n=1}^{\infty} \log n \; a_n n^{-s}$$

which is locally uniformly convergent for $\sigma > \sigma_c$, since if $\sum_{n=1}^{\infty} a_n n^{-\sigma_0}$ converges, then for $\varepsilon > 0$,

$$\sum_{n=1}^{\infty} a_n \log n \; n^{-s-\varepsilon} = \sum_{n=1}^{\infty} a_n n^{-s} \frac{\log n}{n^{\varepsilon}}$$

converges locally uniformly in s and ε by a variant of Dirichlet's test, using the partial summation formula, since $n^{-1} \log n$ is decreasing for $n \geqslant 2$.

Theorem 4.2. *Let $A(x) = \sum_{n \leqslant x} a_n$. If $\sigma_c < 0$, then $A(x)$ is bounded, and*

$$\alpha(s) = \sum_{n=1}^{\infty} a_n n^{-s} = s \int_1^{\infty} A(x) x^{-s-1} \; dx$$

for $\sigma > 0$. If $\sigma_c \geqslant 0$, then

$$\limsup_{x \to \infty} \frac{\log |A(x)|}{\log x} = \sigma_c$$

and the integral formula holds for $\sigma > \sigma_c$.

Proof. Using Riemann Stieltjes integration, we find

$$\sum_{n=1}^{N} a_n n^{-s} = \int_{1^-}^{N} x^{-s} dA(x) = N^{-s} A(N) + s \int_1^N A(x) x^{-s-1} dx$$

Let γ denote the limsup in the theorem. Then for $u > \gamma$, $A(x) = O(x^u)$, and so if $\sigma > u$, the integral converges absolutely as $N \to \infty$. Since $\sigma > 0$, $N^{-s}A(N) = O(N^{u-\sigma})$ tends to zero, we conclude that $\alpha(s)$ converges pointwise and is given by the formula in the theorem. If $\sigma_c < 0$, then $\alpha(0)$ is well-defined, hence $A(x)$ is bounded, and so $\gamma = 0$, implying the equation above holds for all $\sigma > 0$. We know $\gamma \geqslant \sigma_c$, and the only remaining part of the theorem is to show that if $\sigma_c \geqslant 0$, then $\gamma \leqslant \sigma_c$. So suppose $\sigma_0 > \sigma_c$. Then, using the calculations with the remainder term R in the last theorem for $s = 0$ and $M = 0$, we conclude

$$A(N) = -R(N)N^{\sigma_0} + \sigma_0 \int_0^N R(u)u^{\sigma_0 - 1}$$

Since the remainder term is bounded, this implies $A(N) = O(N^{\sigma_0})$, hence $\gamma \leqslant \sigma_0$. Since σ_0 was arbitrary, we conclude that $\gamma \leqslant \sigma_c$. □

Remark. This is the first example of a case where the analytic properties of some function give us information about a given sequence of numbers. In this case, the convergence properties of the Dirichlet series give us

Thus we see that understanding the analytic properties of a Dirichlet series begins to give us information about the original series, in this case, about the growth rate of the partial sums.

Despite both having the property of analyticity, Dirichlet series and power series have differing analytic properties. For instance, power series always converge absolutely within their domain of definition, and converge rapidly to their value, whereas Dirichlet series are more resistant to convergence. One result of this is that a Dirichlet series may not converge absolutely within it's abcissa of convergence.

Example. *The series*

$$\sum_{n=1}^{\infty} (-1)^{n-1} n^{-s}$$

converges for all real valued $s > 0$, by Leibnitz's convergence test. But our discussion implies the series converges for all $s > 0$. On the other hand, taking absolute values of the terms, the series

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

cannot converge for $\sigma < 1$, since the harmonic series $\sum_{n=1}^{\infty} 1/n$ diverges.

For any Dirichlet series, taking absolute values of terms gives another Dirichlet series, hence another abcissa of convergence, and we let σ_a denote the abcissa on the right of which the series converges absolutely. Thus for the series above, we know that $\sigma_a = 1$, yet $\sigma_c \leqslant 0$, so there can be a 'gap' between absolute convergence and pointwise convergence. This is as big a gap as we can get, however.

Theorem 4.3. $\sigma_c \leqslant \sigma_a \leqslant \sigma_c + 1$.

Proof. The first inequality is trivial. To prove the second inequality, we note that if $\sum a_n n^{-\sigma}$ converges, then $|a_n| = O(n^\sigma)$, hence

$$\sum |a_n| n^{-\sigma-1-\varepsilon} = O\left(\sum n^{-1-\varepsilon}\right) < \infty$$

□

Often, we will want to bound a term in terms of the absolute value $|t|$ of the complex part. However, when we want this bound to work uniformly for small $|t|$ and for large $|t|$, we have to shift t by some constant, for otherwise $|t| \to 0$. For arcane reasons, the standard is to bound the term above by $\tau := |t| + 4$.

Theorem 4.4. *Suppose $\alpha(s) = \sum a_n n^{-s}$ has abcissa of convergence σ_c. Then*

$$\alpha(s) \ll \tau^{1-\delta+\varepsilon}$$

uniformly for $\sigma \geqslant \sigma_c + \delta$, if $0 < \varepsilon < \delta < 1.$, where the constant depends on the coefficients a_n, δ, and ε

Proof. Let s be a number with $\sigma \geqslant \sigma_c + \delta$. With $s_0 = \sigma_c + \varepsilon$, applying the remainder term calculations from our first theorem, and then letting $N \to \infty$, we conclude

$$\alpha(s) = \sum_{n=1}^{M} a_n n^{-s} + R(M) M^{\sigma_c+\varepsilon-s} + (\sigma_c + \varepsilon - s) \int_M^\infty R(u) u^{\sigma_c+\varepsilon-s-1}$$

The series converges, so $a_n = O(n^{\sigma_c+\varepsilon})$, and $R(u) = o(1)$, thus

$$\alpha(s) = O\left(\sum_{n=1}^{M} n^{\varepsilon-\delta} + M^{\sigma-\delta} + \frac{|\delta_c + \varepsilon - s|}{\delta - \delta_c - \varepsilon} M^{\sigma_c+\varepsilon-\delta}\right)$$

and

$$\sum_{n=1}^{M} n^{\varepsilon-\delta} < \int_0^M u^{\varepsilon-\delta} du = \frac{M^{1+\varepsilon-\delta}}{1+\varepsilon-\delta} = O\left(M^{1+\varepsilon-\delta}\right)$$

Taking setting $M = \tau + O(1)$ gives the theorem. □

We can now use this fact to show that Dirichlet expansions of a function are unique. Note, however, that not all analytic functions have Dirichlet expansions. This is a very strong property of certain analytic functions.

Theorem 4.5. *If $\sum a_n n^{-s} = \sum b_n n^{-s}$ in some half plane, then $a_n = b_n$ for all n.*

Proof. It suffices to prove this is true if $\sum c_n n^{-s}$ vanishes in the half plane $\sigma > \sigma_0$. Let c_N be the first non-zero coefficient of the series. Then

$$c_N = -\sum_{n>N} c_n (N/n)^{\sigma}$$

This series is absolutely convergent for $\sigma > \sigma_0 + 1$. But the terms on the right hand side converge pointwise to zero as $\sigma \to \infty$, which means $c_N = 0$, a contradiction which shows the uniqueness of the expansion. □

Another point on which the two series differ is that, whereas on the boundary of convergence a power series always has a singularity, a Dirichlet series need not have a singularity at the boundary. It does not take much to prove that the Dirichlet series

$$\sum_{n=1}^{\infty} (-1)^{n-1} n^{-s}$$

corresponds to the holomorphic function $(1 - 2^{1-s})\zeta(s)$, which is entire (the proof of this involves Poisson summation, and we delay this to a later point). However, if the terms of a Dirichlet series are positive, we can guarantee the existence of a singularity at the boundary.

Theorem 4.6 (Landau)**.** *Let $\alpha(s) = \sum a_n n^{-s}$ be a Dirichlet series with finite abcissa of convergence σ_c. If $a_n \geqslant 0$ for all n, then the point σ_c is the singularity of the function $\alpha(s)$.*

Proof. Without loss of generality, we may assume $\sigma_c = 0$. Suppose $\alpha(s)$ is analytic at $s = 0$, and therefore extendable to a holomorphic function in a disk D of radius δ around the origin. We write

$$\alpha(s) = \sum_{k=0}^{\infty} c_k (s-1)^k$$

as a power series at $s = 1$, which converges in a disk at least of radius $\sqrt{1+\delta^2}$, since the nearest non-holmorphic point is $i\delta$. The coefficients of the power series expansion are

$$c_k = \frac{\alpha^{(k)}(1)}{k!} = \frac{1}{k!} \sum_{n=1}^{\infty} \frac{a_n (-\log n)^k}{n}$$

Thus

$$\alpha(s) = \sum_{k=0}^{\infty} \frac{(1-s)^k}{k!} \sum_{n=1}^{\infty} a_n (\log n)^k n^{-1}$$

If $s < 1$, all terms are non-negative, we may interchange summation and write

$$\alpha(s) = \sum_{n=1}^{\infty} \frac{a_n}{n} \sum_{k=0}^{\infty} \frac{(1-s)^k}{k!} (\log n)^k = \sum_{n=1}^{\infty} \frac{a_n}{n} \exp((1-s)\log n) = \sum_{n=1}^{\infty} a_n n^{-s}$$

and the left hand sum is finite for $s < \sqrt{1+\delta^2} - 1$, a contradiction completing the proof. □

4.1 Dirichlet Convolutions and Euler Products

The most basic and fundamental Dirichlet series is the Riemann zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

We know it converges absolutely for $\sigma > 1$, and has a signularity at $\sigma = 1$. Euler established a fundamental connection between this function and the primes, establishing that for $\sigma > 1$,

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

where the product is taken over *prime numbers*. One can see it as an analytic variant of the fundamental theorem of arithmetic. In this section we will prove this representation, as well as other representations for more general Dirichlet series.

We have brashly reasoned that Dirichlet series work well to understand multiplicative functions. Here we establish a first relation between multiplication of Dirichlet series and the multiplicativity of the series' coefficients.

Theorem 4.7. *Let $\alpha(s) = \sum a(n)n^{-s}$ and $\beta(s) = \sum b(n)n^{-s}$ be two Dirichlet series, and set $\gamma(s) = \sum (a*b)(n)n^{-s}$, where $a*b$ is the Dirichlet convolution of the two series. Then whenever α and β are both absolutely convergent, so too is γ, and $\gamma(s) = \alpha(s)\beta(s)$.*

Theorem 4.8. *If $\alpha(s) = \sum a(n)n^{-s} < \infty$, where a is a multiplicative function, then*

$$\alpha(s) = \prod_p \left(1 + a(p)p^{-s} + a(p^2)p^{-2s} + \dots\right)$$

4.2 Smooth Numbers

Let $\psi(x,y)$ denote the number of y-smooth integers up to x, i.e.

$$\psi(x,y) = \#\{1 \leqslant n \leqslant x : p \mid n \Rightarrow p \leqslant y\}$$

If $y > x$, then $\psi(x,y) = \lfloor x \rfloor = x + O(1)$. If $\sqrt{x} \leqslant y \leqslant x$, then any integer $n \leqslant x$ has at most one prime factor in $(\sqrt{x}, x]$, and so

$$\psi(x,y) = \lfloor x \rfloor - \sum_{y < p \leqslant x} \sum_{n \leqslant x} 1 = x + o(1) + \sum_{y \leqslant p \leqslant x} \lfloor \frac{x}{p} \rfloor^{p \mid n} = x - x \sum_{y \leqslant p \leqslant x} \frac{1}{p} + O(\pi(x))$$

Thus $\psi(x,y) = x(1 - \log(\log x / \log y)) + O(x/\log x)$. If $u = \log x / \log y$, then $\psi(x,y) = x(1 - \log u) + O(x/\log x)$.

Theorem 4.9 (Dickman, 1930)**.** *Define $\rho : [0,\infty) \to \mathbf{R}$ to be the unique continuous function satisfying $\rho(u) = 1$ for $0 \leqslant u \leqslant 1$, and $u\rho(u) = -\rho(u-1)$ for $u > 1$. Then for any $U > 0$,*

$$\psi(x, x^{1/u}) = \rho(u)x + O(x/\log x)$$

Proof. Let's consider the properties of ρ. Since $u\rho'(u) = -\rho(u-1)$, we find

$$\rho(v) = \rho(u) - \int_u^v \rho(t-1)\,dt$$

THus $(u\rho(u))' = \rho(u) - \rho(u-1)$, we find

$$u\rho(u) = \int_{u-1}^{u} \rho(t)\,dt$$

For the definition, it is obvious that $\rho(u)$ is positive and nonincreasing for $u > 0$ (consider the smallest value where ρ is zero). ρ decays incredibly fast, but we will not prove this ($\rho(u) \equiv 1/u^u$). (We ran out of time so we didn't prove this).

Now we prove the theorem when u is an integer by inductinos. The theorem is easy to see for $u = 1$. The key idea to form the induction based on the formula

$$\psi(x,y) = 1 + \sum_{p \leqslant y} \#\{nleqx : P(n) = p\}$$

where $P(n)$ is the largest positive factor of n. Now

$$\#\{n \leqslant x : P(n) = p)\} = \psi(x/p, p)$$

□

Now set $\Phi(x,y) = \#\{n \leqslant x : p \mid n \Rightarrow p \geqslant y\}$. If $y > x$, $n = 1$ is counted, so $\Phi(x,y) = 1$. For $\sqrt{x} \leqslant y \leqslant x$, the # counted by $\Phi(x,y)$ are 1 and all primes in $(y,x]$, so

$$\Phi(x,y) = \pi(x) - \pi(y) + O(1) = \frac{x}{\log x} - \frac{y}{\log y} + O\left(\frac{x}{(\log x)^2}\right)$$

for smaller y, heuristically, the events "not divisible by p", for small primes p are independent with probability $1 - p^{-1}$, hence

$$\Phi(x,y) \cong \prod_{p \leqslant y} (1 - p^{-1}) \cong e^{-\gamma} \frac{x}{\log x}$$

Theorem 4.10 (Buchstab)**.** *Let $u = \log x / \log y$, and $w(u)$ the function satisfying $w(u) = u^{-1}$ for $1 \leqslant u \leqslant 2$, and*

Proof. s □

4.3 Gauss Sums for Quadratic Characters

For any prime $p \geqslant 3$, the Legendre symbol $\left(\frac{n}{p}\right)_L$ is a primitive quadratic Dirichlet character modulo p, defined by

$$\left(\frac{n}{p}\right) = \begin{cases} +1 & : n \text{ is a quadratic residue mod } p \text{ and } n \not\equiv 0 \bmod p \\ -1 & : n \text{ is a qudratic non-residue modulo } p \\ 0 & : n \equiv 0 \end{cases}$$

Thus we may calculate the Gauss sum

$$\begin{aligned} \tau_p = \tau\left(\frac{\cdot}{p}\right)_L &= \sum_{n=1}^{p} \left(\frac{n}{p}\right)_L e(n/p) \\ &= \sum_{n=1}^{p} \left(1 + \left(\frac{n}{p}\right)_L\right) e(n/p) \end{aligned}$$

since $\sum e(n/p) = 0$. Thus

$$\begin{aligned} \tau_p &= \sum_{n=1}^{p} \#\{m \pmod p : b^2 \equiv a \pmod p\} e(n/p) \\ &= \sum_{m=1}^{p} e(m^2/p) = \sum_{m=1}^{p} e^{2\pi i m^2/p} \end{aligned}$$

Recall that $\tau(\overline{\chi}) = \chi(-1)\overline{\tau(\chi)}$ so

$$\tau_p = \tau\left(\left(\frac{\cdot}{p}\right)_L\right) = \tau\left(\overline{\left(\frac{\cdot}{p}\right)_L}\right) = \left(\frac{-1}{p}\right)_L \tau_p$$

Multiply by τ_p to conclude

$$\tau_p^2 = \left(\frac{-1}{p}\right)_L \overline{\tau_p}\tau_p = \left(\frac{-1}{p}\right)_L |\tau_p|^2 = \left(\frac{-1}{p}\right)_L p$$

hence

$$\tau_p = \pm\sqrt{\left(\frac{-1}{p}\right)_L p} = \begin{cases} \pm\sqrt{p} & : p \equiv 1 \pmod 4 \\ \pm i\sqrt{p} & : p \equiv 3 \pmod 4 \end{cases}$$

An integer d is a **fundamental** or **quadratic discriminant** if either $d \equiv 1 \pmod 4$ and d is squarefree, or $d = 4m$ where $m \equiv 2$ or 3 modulo 4 and m is squarefree.

We define the **Kronecker symbol** $\left(\frac{d}{n}\right)_K$ for a fundamental discriminant d and a nonzero integer n as follows. If p is prime, then $\left(\frac{d}{p}\right)_K = \left(\frac{d}{p}\right)_L$. The symbol is totally multiplicative in n, and $\left(\frac{d}{-1}\right)_L = \operatorname{sgn}(d)$. If d is a fundamental discriminant, then $\chi_d(n) = \left(\frac{d}{n}\right)_K$ is a primitive quadratic characters modulo $|d|$. Every primitive quadratic Dirichlet character is equal to χ_d for some fundamental discriminant d, and if d_1 and d_2 are coprime fundamental discriminant, then

$$\left(\frac{d_1}{d_2}\right)_K \left(\frac{d_2}{d_1}\right)_K = \begin{cases} 1 & d_1 > 0 \text{ or } d_2 > 0 \\ -1 & d_1 < 0 \text{ and } d_2 < 0 \end{cases}$$

as an example

$$\left(\frac{n}{p}\right)_L = \begin{cases} \chi_p(n) & : p \equiv 1 \pmod 4 \\ \chi_{-p}(n) & : p \equiv 3 \pmod 4 \end{cases}$$

Recall that $L(1, \chi) \neq 0$. The proof depended on whether χ was complex or real quadratic. When $\chi = \chi_d$ is quadratic and primitive. We know from Dirichlet's theorem that the exact value of $L(1, \chi_d)$ is related to the arithmetic of the quadratic number field $\mathbf{Q}(\sqrt{d})$. If $d < 0$, then $L(1, \chi_d) = 2\pi h(d)/w_d\sqrt{|d|}$, where $h(d)$ is the class number of $\mathbf{Q}(\sqrt{d})$, and w_d is the number of roots of unity in $\mathbf{Q}(\sqrt{d})$.

If (x_0, y_0) is the minimal positive solution to Pell's equation $x_0^2 - dy_0^2 = 4$, and we let $\eta_d = (1/2)(x_0 + y_0\sqrt{d})$, then $L(1, \chi_d) = h(d) \log \eta_d / \sqrt{d}$.

Theorem 4.11 (Polya-Vinogradov)**.** *Let χ modulo q be a nonprincipal character. Then*

$$\sum_{M < n \leq M+N} \chi(n) \ll \sqrt{q} \log q$$

We can't improve this in general, since $\sum \chi(n) = \Omega(\sqrt{q})$, but we can prove there is cancellation, which gives $\sum_{M<n\leq M+N} \chi(n) = o(N)$ for N as small as roughly $q^{1/4+\varepsilon}$.

Proof. We can assume $N < q$. If we assume χ is primitive, we know

$$\sum_{a=1}^{q} \overline{\chi}(a) e(an/q) = \chi(n)\tau(\overline{\chi})$$

But also $|\tau(\overline{\chi})| = \sqrt{q} \neq 0$, so

$$\chi(n) = \frac{1}{\tau(\overline{\chi})} \sum_{a=1}^{q} \overline{\chi}(a) e(an/q)$$

Summing over n and switching the order of summation, we find

$$\sum_{M<n\leqslant M+N} \chi(n) = \sum_{M<n\leqslant M+N} \frac{1}{\tau(\overline{\chi})} \sum_{a=1}^{q} \overline{\chi}(a) e(an/q) = \frac{1}{\tau(\overline{\chi})} \sum_{a=1}^{q} \overline{\chi}(a) \sum_{M<n\leqslant M+N} e(an/q)$$

when $a \neq 0$, the inner sum is a geometric series equaling

$$\frac{e(\text{something}) - e(\text{something})}{1 - e(a/q)} \ll d(a/q, \mathbf{Z})^{-1}$$

we set

$$\sum_{M<n\leqslant M+N} \chi(n) \ll \frac{1}{\sqrt{q}} \left(N + \sum_{a=1}^{q-1} q/a \right) \ll \frac{1}{\sqrt{q}} (q + q \log q)$$

This completes the proof when χ is primitive. In general, if χ is induced by a primitive character χ^* modulo q^*, and if $r = \prod_{p \mid q, p \nmid q^*} p$, then

$$\begin{aligned}
\sum_{N<n\leqslant M+N} \chi(n) &= \sum_{\substack{M<n\leqslant M+N \\ (r,n)=1}} \chi^*(n) = \sum_{M<n\leqslant M+N} \chi^*(n) \sum_{d \mid (r,n)} \mu(d) \\
&= \sum_{d \mid r} \mu(d) \sum_{\substack{M<n\leqslant M+N \\ d \mid n}} \chi^*(n) = \sum_{d \mid r} \mu(d) \sum_{M/d<m\leqslant (M+N)/d} \chi^*(d)\chi^*(m) \\
&\ll \sum_{d \mid r} \mu^2(d) \sqrt{q^*} \log q^* \\
&\ll 2^{\omega(r)} \sqrt{q^*} \log q
\end{aligned}$$

It's easy to show $2^{\omega(r)} \ll \sqrt{r} \leqslant \sqrt{q/q^*}$. □

An application of this theorem is that if n_x is the smallest integer with $\chi(n_x) \neq 1$, then, by considering $\sum_{1 \leqslant n < n_x} \chi(n)$, we set $n_x \ll \sqrt{q} \log q$. Using the fact that $\chi(n) = 1$ for alll n_x friable n, this gives

$$n_x \ll_\varepsilon q^{1/2\sqrt{e}+\varepsilon}$$

Another application is that if p is prime, there are $\phi(p-1)$ primitive roots modulo p in the interval $[1, p-1]$. Since we have no reason for the roots to occur in a particular ordering, we might expect that the distribution of these roots is random. That is, in an interval $0 \leqslant M < M+N < p$, to have about $\phi(p-1)/(p-1)N$ primitive roots. We can show

$$\#\{\text{primitive roots modulo } p \text{ in } (M, M+N]\} = \frac{\phi(p-1)}{p-1} N + O\left(\phi(p-1)\sqrt{p}\log p\right)$$

Almost all progress on estimates for incompletes character sums comes from Polya's identity. Let

$$f_\chi(\alpha) = \sum_{0 < n \leqslant \alpha q} \chi(m)$$

if $\chi \neq \chi_0$. $f_x(\alpha)$ is periodic of period 1, we can compute it's Fourier expansion

$$f_\chi(\alpha) = -q^{-1} \sum_{n=1}^{q} n\chi(n) + \frac{\tau(\chi)}{2\pi i} \sum_{k=1}^{\infty} \frac{\overline{\chi}(k)}{k} \left(e(\alpha k) + \chi(-1) e(-\alpha k)\right)$$

We could have obtained Polya Vinogradov by estimating this sum. Now if χ is primitive modulo q, then

$$\sum_{0 < n \leqslant q/2} \chi(n) = \begin{cases} 0 & : \text{if } \chi \text{ is even} \\ (2 - \chi(2)) \frac{\tau(\chi)}{\pi i} L(1, \overline{\chi}) & : \chi \text{ odd} \end{cases}$$

In particular, if $d < 0$ is a fundamental discriminant, then

$$\sum_{0 < n \leqslant q/2} \left(\frac{d}{n}\right)_K > 0$$