Talk 10.5: Dirichlet series and arithmetic functions

Additive number theory seminar

The goal of today's talk is to introduce some general multiplicative number-theoretic tools that'll be useful going forward into the next half of the class, which is focused around sieve theory. We won't define sieves today, but the problems that we'll try to use them to solve have to do with counting primes with certain additive properties, i.e. mixing additive and multiplicative number theory. We've studied some additive number theory already; before we worry about blending them, we'll just sketch some concepts from classical multiplicative number theory, including arithmetic functions, Dirichlet series, and how we can apply this sort of machinery to questions like the prime number theorem.

1 Arithmetic functions

We've seen the term "arithmetic function" before in this class: it just means a function $f : \mathbb{N} \to \mathbb{C}$, where $\mathbb{N}$ denotes the natural numbers $\{1, 2, 3, \ldots\}$ (sometimes people like to include 0 in $\mathbb{N}$, but we generally won't). In practice most of our arithmetic functions will take real values, and often (but not always) even integer values; however not much changes if we consider general complex values.

A simple example is the constant function $f(n) = 1$, which we often just denote by 1. Another is the function $\delta(n)$, which is 1 if $n = 1$ and 0 otherwise. More complicated examples include the "number of divisors" function $d(n)$, which can be written as

$$\sum_{d|n} 1$$

(e.g. $d(10) = \#\{1, 2, 5, 10\} = 4$) or the prime indicator function $1_{\mathcal{P}}(n)$ which is 1 if n is prime and 0 otherwise.

Our goal is to study these kinds of functions using certain generating series, similar to in recent talks, although we'll use a different type of generating series. However, already from these examples we can see that the behavior of this kind of function can be very wild: for example, the divisor function $d(n)$ can be as large as $\log_2(n)$ if $n = 2^k$ is a power of 2, since then every i between 0 and $k = \log_2(n)$ gives a divisor 2^i of $n = 2^k$, but it can also be as small as 2 for prime numbers (since the only divisors of primes n are 1 and n). So $d(n)$ will bounce between 2 and increasingly high values; there's no way to get an asymptotic formula for it as n becomes large, like we could for e.g. $r_{k,s}(n)$.

For these kinds of functions, we instead try to estimate the *average value*

$$\frac{1}{N} \sum_{n=1}^{N} f(n),$$

or more generally the summatory function $S_f(N) = \sum_{n=1}^{N} f(n)$. For example, even though the asymptotic behavior of $d(n)$ doesn't make sense, it turns out that for N large,

$$S_d(N) = \sum_{n=1}^{N} d(n) = N \log N + O(x).$$

(In fact, more is known: $S_d(N) = N \log N + cx + O(\sqrt{x})$ for a certain constant $c \approx 0.1544$.) So the "average value" of $d(n)$ between 1 and N is

$$\frac{1}{N} S_d(N) = \log N + O(1),$$

not quite as high as $\log_2(N)$ but growing to infinity at about the same rate.

Like for any functions, we can add, subtract, and multiply them together, as well as dividing so long as the values of the denominator aren't 0. However, there's an additional operation we can do with arithmetic functions, called Dirichlet convolution: given two arithmetic functions f and g, we define

$$(f * g)(n) = \sum_{d|n} f(d)g(n/d).$$

This definition is more symmetric than it looks: d and n/d are both divisors of n, and so if we let $e = n/d$ then $d = n/e$ and so

$$(f * g)(n) = \sum_{d|n} f(d)g(n/d) = \sum_{e|n} f(n/e)g(e) = (g * f)(n),$$

i.e. as an operation on arithmetic functions, Dirichlet convolution is commutative. One can check with a little more effort that it is also associative, and distributes across addition (i.e. $f * (g + h) = f * g + f * h$), so a high-brow way of saying this is that $(*, +)$ define a ring structure on the set of arithmetic functions.

Let's think about some examples: an easy one is

$$(1 * 1)(n) = \sum_{d|n} 1 \cdot 1 = d(n),$$

so $d = 1 * 1$. The function δ from above becomes more relevant here: for any arithmetic function f, we have

$$(\delta * f)(n) = \sum_{d|n} \delta(d) f(n/d) = f(n)$$

since $\delta(d) = 0$ unless $d = 1$, so $\delta * f = f$; in other words, δ is the multiplicative unit of Dirichlet convolution.

Since we have a unit, we might ask about multiplicative inverses: for example, is there a function f such that $f * 1 = \delta$? It turns out that the answer is yes: this is called the Möbius function $\mu(n)$, and it is defined as follows. Write n as a product of prime numbers, i.e. $n = p_1^{e_1} \cdots p_r^{e_r}$. If any of the e_i is greater than 1, then n is divisible by p_i^2; in this case set $\mu(n) = 0$, so $\mu(n)$ will only be nonzero if n is *squarefree*, i.e. not divisible by any squares. In this case $n = p_1 \cdots p_r$, and we define $\mu(n) = (-1)^r$. A table of the first few values of $\mu(n)$ is below.

n	1	2	3	4	5	6	7	8	9	10
$\mu(n)$	1	-1	-1	0	-1	1	-1	0	0	1

One can verify that $\mu * 1 = \delta$ at least for the first few n; proving this in general is a little tricky, but not really difficult. This is of use for a few reasons: for example, one can detect whether two numbers are relatively prime using the Möbius function and this identity. Indeed, $\gcd(a, b) = 1$ if and only if (by definition) the only number dividing both a and b is 1, so we have

$$\sum_{d|a,d|b} \mu(d) = \sum_{d|\gcd(a,b)} \mu(d)$$

equal to 1 if $\gcd(a, b) = 1$ and 0 otherwise. This will be useful for sieving!

The reason all this is of interest is that we're going to associate to each arithmetic function f a certain kind of generating series, called the Dirichlet series $D_f(s)$, with the property that $D_{f*g} = D_f D_g$; we'll discuss this more soon. The goal is that by studying D_f using analytic methods, we can get back information about S_f, and thus information "on average" about f itself. We won't get into the technicalities of how to do this too much (one can use some integrals reminiscent of the circle method), choosing instead to give a heuristic picture of how this applies to estimating some functions of interest, and in particular to counting prime numbers.

2 Dirichlet series

Let s be a complex number. For an arithmetic function f, we define a complex function

$$D_f(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}.$$

In general, this need not converge; typically it'll converge so long as the real part of s is at least some value c depending on f, sometimes called the abscissa of convergence.

The most famous example of a Dirichlet series is for the constant function 1:

$$D_1(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

This is called the Riemann zeta function, and is often written $\zeta(s)$, i.e. $\zeta = D_1$. One can also check that

$$D_\delta(s) = \sum_{n=1}^{\infty} \frac{\delta(n)}{n^s} = \frac{1}{1^s} + \frac{0}{2^s} + \frac{0}{3^s} + \cdots = 1.$$

This means that just as δ is the multiplicative unit for Dirichlet convolution, D_δ is the multiplicative unit for regular multiplication of power series, compatible with the following proposition:

Proposition. *For any two arithmetic functions f, g, we have*

$$D_{f*g} = D_f D_g.$$

In the abstract language, together with the linearity of D this means that D gives a ring homomorphism from arithmetic functions (under $(*, +)$) to complex-analytic functions defined on at least an upper half-plane (with the usual operations). It's not too hard to see that this homomorphism is injective, so if $D_f = D_g$ then $f = g$.

Proof. The proof is an exercise in expansion:

$$\begin{aligned}
D_{f*g}(s) &= \sum_{n=1}^{\infty} \frac{1}{n^s}(f * g)(n) \\
&= \sum_{n=1}^{\infty} \frac{1}{n^s} \sum_{d|n} f(d)g(n/d) \\
&= \sum_{n=1}^{\infty} \frac{1}{n^s} \sum_{d,e:de=n} f(d)g(e) \\
&= \sum_{d,e\geq 1} \frac{1}{(de)^s} f(d)g(e) \\
&= \left(\sum_{d\geq 1} \frac{f(d)}{d^s}\right)\left(\sum_{e\geq 1} \frac{g(e)}{e^s}\right) \\
&= D_f(s)D_g(s).
\end{aligned}$$

□

This means that we automatically get identities such as $D_d(s) = D_{1*1}(s) = D_1(s) \cdot D_1(s) = \zeta(s)^2$, and $D_\mu \cdot D_1 = D_\delta = 1$, i.e. $D_\mu(s) = \frac{1}{\zeta(s)}$.

One can also determine the Dirichlet series of other functions in terms of these building blocks using some alternate methods. For example, differentiating $\zeta(s)$ gives

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

so if log denotes the arithmetic function $n \mapsto \log n$ then $D_{\log}(s) = -\zeta'(s)$. If we write $\mathrm{id}(n) = n$, or more generally $\mathrm{id}_k(n) = n^k$, then

$$D_{\mathrm{id}_k}(s) = \sum_{n=1}^{\infty} \frac{n^k}{n^s} = \sum_{n=1}^{\infty} \frac{1}{n^{s-k}} = \zeta(s-k).$$

We can combine these in various ways: for example, consider the sum-of-divisors function

$$\sigma(n) = \sum_{d|n} d,$$

or more generally the sum of powers of divisors

$$\sigma_k(n) = \sum_{d|n} d^k.$$

We can observe that $\sigma_k = \mathrm{id}_k * 1$ (with $\sigma = \mathrm{id} * 1$), so without doing any further work we find

$$D_{\sigma_k}(s) = D_{\mathrm{id}_k}(s) D_1(s) = \zeta(s-k)\zeta(s).$$

Before proceeding further, let's make some analytic observations about $\zeta(s)$ which we can use to comment on the other arithmetic functions at play here. One can check via the integral test that $\zeta(s)$ converges when the real part of s is greater than 1 and diverges if it's ≤ 1. In fact, via some magic one can show that for $\mathrm{Re}(s) > 1$, one has

$$\zeta(s) = \frac{s}{s-1} - s \int_1^\infty \frac{\{x\}}{x^{s+1}}\, dx$$

where $\{x\}$ is the fractional part of x. (Indeed, the right-hand side is

$$s \int_1^\infty \frac{1}{x^s}\, dx - s \int_1^\infty \frac{\{x\}}{x^{s+1}}\, dx = s \int_1^\infty \frac{x - \{x\}}{x^{s+1}}\, dx s \sum_{n=1}^\infty \int_n^{n+1} \frac{\lfloor x \rfloor}{x^{s+1}}\, dx$$
$$= s \sum_{n=1}^\infty n \int_n^{n+1} \frac{1}{x^{s+1}} = s \sum_{n=1}^\infty n \left(\frac{1}{sn^s} - \frac{1}{s(n+1)^s} \right) = \sum_{n=1}^\infty \frac{n}{n^s} - \sum_{n=1}^\infty \frac{n}{(n+1)^s}$$
$$= \sum_{n=1}^\infty \frac{n}{n^s} - \sum_{n=2}^\infty \frac{n-1}{n^s} = \sum_{n=1}^\infty \frac{1}{n^s} = \zeta(s).)$$

The miracle is that the integral here is actually convergent for all $\mathrm{Re}(s) > 0$, so this gives an extension of $\zeta(s)$ to this region! In particular, near $s = 1$ this tells us that $\zeta(s) = \frac{1}{s-1} + O(1)$, so ζ has a simple pole at $s = 1$.

Here's an observation: since $\zeta(s)$ has a pole at $s = 1$, we found above that $D_{\mathrm{id}_k}(s) = \zeta(s-k)$, so D_{id_k} has a pole at $s = k+1$. On the other hand of course $\mathrm{id}_k(n) = n^k$, so we can compute its summatory function explicitly, or approximate

$$S_{\mathrm{id}_k}(N) = \sum_{n=1}^N n^k \approx \int_0^N x^k\, dx = \frac{N^{k+1}}{k+1}.$$

This suggests the following principle: if D_f has a simple pole at c, then we expect that $S_f(N)$ is of order N^c. Indeed, if so then "on average" $f(n) \approx \frac{1}{n} S_f(n) \approx n^{c-1} = \mathrm{id}_{c-1}(n)$, so $D_f(s) \approx D_{\sigma_{c-1}}(s) = \zeta(s-c+1)$ which has a pole at $s = c$.

This is a very heuristic conclusion, but it turns out it can be made rigorous, so we'll run with it for now. In particular, the fact that ζ has a simple pole at 1 reflects that

$$S_1(N) = \sum_{n=1}^N 1 = N$$

is "approximately" $N^1 = N$, so similar behavior in other functions implies a similar estimate.

If there are multiple poles, the pole with the largest real part gives the dominant term in the approximation, so really we're just concerned with finding the pole with the largest real part. This is lucky for us because for all we know $\zeta(s)$ may have various other poles,

but from the expression above we can check that they all have real part at most 0 so don't contribute meaningfully to our estimates. (In fact it turns out ζ has no other poles.)

Applying our heuristic to e.g. $\sigma_k(n)$, we see that for $k > 0$, $D_{\sigma_k} = \zeta(s-k)\zeta(s)$ has largest pole at $s = k+1$ and so heuristically $S_{\sigma_k}(N) \approx \frac{N^{k+1}}{k+1}$ up to some multiplicative constant. This constant should be $\zeta(k+1)$, which for $k > 0$ is finite. For example, it turns out that $\zeta(2) = \frac{\pi^2}{6}$, so we expect that $S_\sigma(N) \approx \frac{\pi^2}{6} \cdot \frac{1}{2}N^2 = \frac{\pi^2}{12}N^2$, i.e. the "average value" of $\sigma(n)$ is $\frac{\pi^2}{12}n$.

3 Euler products

There exists another formula for the Riemann zeta function, called an Euler product:

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

where the product is over all prime numbers. This also converges for $\mathrm{Re}(s) > 1$. To see this claim, note that

$$(1 - p^{-s})^{-1} = \frac{1}{1 - p^{-s}} = 1 + p^{-s} + p^{-2s} + p^{-3s} + \cdots,$$

so the product is

$$(1 + 2^{-s} + 2^{-2s} + \cdots)(1 + 3^{-s} + 3^{-2s} + \cdots)\cdots.$$

If we try to expand out the product, the first term is 1; then we get a term of p^{-s} for every prime p; then terms $p^{-s}q^{-s}$ for pairs of primes p and q, as well as p^{-2s}; and so on. All in all, we get one term $p_1^{-e_1 s} \cdots p_r^{-e_r s}$ for every collection of primes with multiplicity $p_1^{e_s} \cdots p_r^{e_s}$. But (by the fundamental theorem of arithmetic) such a collection is exactly the same thing as a natural number n; so the expansion of the product is exactly

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

Thus in a way this product formula encodes the fundamental theorem of arithmetic.

(This sort of expansion may help to explain why Dirichlet series are the right version of generating series for multiplicative, rather than additive, number theory: when doing this sort of expansion, we find that the number of occurrences of n^{-s} is the number of ways that n can be written as a *product* of factors of a given kind, rather than a sum. Here we're looking at the number of ways n can be written as a product of prime powers, up to order, which is of course 1 by the fundamental theorem of arithmetic.)

One can find similar Euler product expansions for other Dirichlet series D_f so long as f is completely multiplicative, i.e. satisfies the property $f(ab) = f(a)f(b)$ for all a, b; but we won't need this today. Instead, looking at this product, we could turn it into a sum by taking the logarithm,

$$\log \zeta(s) = -\sum_p \log(1 - p^{-s});$$

we'd like to try and understand this as a Dirichlet series for some arithmetic function having to do with detecting primes.

In fact, to make the algebra easier it's better to differentiate:

$$\frac{d}{ds}\log\zeta(s) = -\sum_p \frac{d}{ds}\log(1-p^{-s}) = -\sum_p \frac{\log p}{1-p^{-s}} = -\sum_p \sum_{j=0}^{\infty} p^{-js}\log p.$$

This suggests the following notation: if n is a prime power $n = p^j$, let $\Lambda(n) = \log p = \frac{1}{j}\log n$; and if n is not a prime power, $\Lambda(n) = 0$ (Λ detects prime powers, inversely weighted by the power; primes count for the most, then squares of primes, and so on). This is called the von Mangoldt function. Plugging into the above, we get

$$-\frac{d}{ds}\log\zeta(s) = \sum_{n=1}^{\infty}\frac{\Lambda(n)}{n^s} = D_\Lambda(s).$$

On the other hand, $-\frac{d}{ds}\log\zeta(s) = -\frac{\zeta'(s)}{\zeta(s)} = -\zeta'(s)\cdot\frac{1}{\zeta(s)}$, so since $D_{\log} = -\zeta'$ and $D_\mu = \frac{1}{\zeta}$ we find $\Lambda = \log * \mu$, or equivalently $\log = \Lambda * 1$, i.e.

$$\log n = \sum_{d|n}\Lambda(d).$$

One can also interpret this as a version of the fundamental theorem of arithmetic, details left to the reader.

4 The prime number theorem and associated questions

Equipped with these identities, we can now try to study prime numbers. In particular, one can show by fairly elementary methods that the prime counting function $\pi(N) = S_{1_{\mathcal{P}}}(N)$ is closely related to S_Λ; indeed we have

$$\pi(N) \approx \frac{S_\Lambda(N)}{\log N}.$$

(The error term can be made explicit, but is not of concern to us.) So we'd like to study S_Λ. For this, we employ our heuristic from above: it suffices to know where the largest poles of $D_\Lambda = -\frac{\zeta'}{\zeta}$ are.

We know that $\zeta(s) = \frac{1}{s-1} + O(1)$ near $s = 1$, so $-\zeta'(s) = \frac{1}{(s-1)^2}$ and so $-\frac{\zeta'(s)}{\zeta(s)} \approx \frac{1}{s-1}$ near $s = 1$; this is the same as the original behavior of ζ, and so we expect that $S_\Lambda(N) \approx N$, which would give

$$\pi(N) \approx \frac{N}{\log N},$$

which indeed is the prime number theorem. However, there is an issue: in addition to the pole coming from the poles of ζ' and ζ, if $\zeta(s) = 0$ then we also get a pole of $-\frac{\zeta'}{\zeta}$!

To fix this, we need to study the zeros of $\zeta(s)$. From the convergence of the product formula we know there are none with real part > 1. Via a clever argument, one can show

that there are also none with real part equal to 1; this suffices to prove the prime number theorem, since it means the largest contribution comes from the pole at $s = 1$. If we knew for sure that there were no zeros with real part greater than c, then we would know that D_Λ had no poles with real part greater than c and could derive a bound on the error in the prime number theorem (or a more precise version of it), which says that the error is $O(N^{c+\epsilon})$. The "best possible" version of this is if there were no zeros with real part greater than $\frac{1}{2}$ (since we know there are zeros with this real part); this is the Riemann hypothesis, and is wildly open.

We can also interpret this in terms of bounding S_μ. Recall that $D_\mu = \frac{1}{\zeta}$, so it has poles at the zeros of ζ. Thus under the same condition above, we expect

$$S_\mu(N) = O(N^{c+\epsilon}).$$

The analytic interpretation of the prime number theorem is that ζ has no zeros with real part 1; this (together with some more careful analysis) means that S_μ is "sub-linear," i.e.

$$\lim_{N\to\infty} \frac{1}{N} S_\mu(N) = 0,$$

i.e. there is meaningful cancellation in S_μ; this is analogous to some of our results on exponential sums.

If μ were truly "random," we'd expect from the theory of random walks that

$$S_\mu(N) = O(N^{\frac{1}{2}+\epsilon}).$$

In fact this is exactly what is predicted by the Riemann hypothesis; so one can think of this sort of strong conjecture about the distribution of the prime numbers as saying that the Möbius function is pseudorandom.