MATH 155 NOTES

MOOR XU
NOTES FROM A COURSE BY KANNAN SOUNDARARAJAN

Abstract. These notes were taken during Math 155 (Analytic Number Theory) taught by Kannan Soundararajan in Spring 2012 at Stanford University. They were live-TEXed during lectures in `vim` and compiled using `latexmk`. Each lecture gets its own section. The notes are not edited afterward, so there may be typos; please email corrections to `moorxu@stanford.edu`.

1. 4/3

1.1. **Introduction.** The aim of the course is to understand some problems in prime number theory. For example, we want to count the number of primes up to x by considering $\pi(x) = \sum_{p\le x} 1$. This is a problem that Gauss worked on; he made a conjecture when he was 14. Gauss conjectured that $\pi(x) \sim \int_2^x \frac{dt}{\log t} = \mathrm{li}(x) \approx \frac{x}{\log x}$. This was finally proved in around 1895, and it is known as the prime number theorem.

The proof is a very remarkable thing that builds on ideas introduced by Riemann in his paper from 1859. There, Riemann introduces some beautiful ideas in number theory. He works with the Riemann zeta function, which is $\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s}$ for $s = \sigma + it$; this converges absolutely when $\sigma > 1$. Riemann proved a beautiful functional equation $\pi^{-s/2}\Gamma(s/2)\zeta(s) = \pi^{-(1-s)/2}\Gamma(\frac{1-s}{2})\zeta(1-s)$, where Γ is Euler's gamma function.

Riemann also proved the explicit formula relating the zeros of the Riemann zeta function to the primes: $\sum_p f(p) \leftrightarrow \sum_p \tilde{f}(p)$ for some function on the primes and some transform of the function f.

This relates to one of the most famous problems in mathematics, known as the Riemann Hypothesis: All of the nontrivial zeros of $\zeta(s)$ lie on the line $\mathrm{Re}(s) = \frac{1}{2}$. This is a major unsolved problem, worth a million dollars.

One of the goals of the course is to prove the prime number theorem, and we will see some properties of the zeta function along the way. We will also see some asymptotic methods.

As another example, consider $p(n)$ is the number of partitions of an integer n. There is a remarkable theorem due to Hardy and Ramanujan that $p(n) \sim \frac{e^{c\sqrt{n}}}{d\sqrt{n}}$, where $c = \pi\sqrt{2/3}$ and $d = 4\sqrt{3}$. This led to something called the circle method.

There is also the Goldbach problem, which is unsolved, but there is known that any large odd number is the sum of three primes. The history of this conjecture is also somewhat odd. Goldbach was not a great mathematician, but he wrote to Euler, who was a great mathematician.

1.2. **Elementary ideas in prime number theory.** Let's start with some elementary ideas in prime number theory. We begin by introducing some notation.

Definition 1.1. $f = O(g)$ means that there exists C with $|f(x)| \le Cg(x)$ for all large x.

Definition 1.2. $f = o(g)$ means that for all $\varepsilon > 0$, there exists x_0 large enough such that if $x > x_0$ then $|f(x)| \le \varepsilon g(x)$.

Definition 1.3. $f \ll g$ means exactly the same thing as $f = O(g)$. $f \gg g$ means $g = O(f)$. $f \asymp g$ means that $g \gg f \gg g$.

Definition 1.4. $f \sim g$ means that $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$, which is exactly the same as saying that $f(x) = g(x) + o(g(x))$.

We have $O(f_1) + O(f_2) + \cdots = O(|f_1| + |f_2| + \cdots)$. The absolute values are important if the functions $f_1, f_2, \ldots$ change sign.

Definition 1.5. $\omega(n)$ is the number of primes that divide n. We can write $\omega(n) = \sum_{p|n} 1$.

Definition 1.6. $d(n) = \sum_{d|n} 1$ is the number of divisors of n.

These questions all try to understand some sum $A(x) = \sum_{n\le x} a(n)$ where $a(n)$ is some arithmetic function, such as $\omega(n)$, $d(n)$, or the characteristic function of the primes. Another interesting arithmetic function is the Mobius function:

Definition 1.7.

$$\mu(n) = \begin{cases} (-1)^k & m = p_1 p_2 \ldots p_k,\ p_k \text{ distinct} \\ 0 & \text{otherwise.} \end{cases}$$

This should be ± 1 somewhat randomly, and 0 occasionally.

We will also use the technique of *partial summation.* Can one pass from $A(x)$ to information about $\sum_{n\le x} a(n) f(n)$ where f is nice real-valued function?

Example 1.8. For example, can we pass from $A(x) = \sum_{n\le x} a(n)$ to $\sum_{n\le x} \frac{a(n)}{n}$?

More specifically, we might consider $\sum_{n\le x} \frac{1}{n}$ or $\sum_{n\le N} \log n = \log(N!)$. These sums arise from taking $a(n) = 1$, and $f(n) = \frac{1}{n}$ or $f(n) = \log n$. So then $A(x) = \lfloor x \rfloor = x - \{x\}$ where $\{x\}$ is the fractional part of x.

This is an idea of Abel, called Abel's partial summation. First, we can write this as

$$\sum_{n=1}^{N} a(n)f(n) = \sum_{n=1}^{N} f(n)\left(A(n) - A(n-1)\right) = \sum_{n=1}^{N} A(n)f(n) - \sum_{n=1}^{N} A(n-1)f(n)$$
$$= \sum_{n=1}^{N} A(n)f(n) - \sum_{n=0}^{N-1} A(n)f(n+1) = A(N)f(N) - A(0)f(1) - \sum_{n=1}^{N-1} A(n)(f(n+1) - f(n)).$$

This should look like integration by parts. It might be convenient to think about this in this way:

$$\sum_{n=1}^{N} a(n)f(n) = \int_{1^-}^{N^+} f(t)d(A(t)) = f(t)A(t)\big|_{1^-}^{N^+} - \int_{1^-}^{N^+} A(t)f'(t)\,dt.$$

To make this fully rigorous, this is the Riemann-Stieltjes integral.

Example 1.9. We will study $\sum_{n\le N}\frac{1}{n}$. Here, $A(t)=\sum_{n\le t}1=[t]=t-\{t\}$. Then

$$\begin{aligned}\sum_{n\le N}\frac{1}{n} = \int_{1^-}^{N^+}\frac{1}{t}\,d(A(t)) = \frac{A(t)}{t}\bigg|_{1^-}^{N^+} + \int_{1^-}^{N^+}[t]\frac{dt}{t^2}\\ = 1+\int_1^N\frac{dt}{t}-\int_1^N\frac{\{t\}}{t^2}\,dt = \log N+1-\int_1^N\frac{\{t\}}{t^2}\,dt.\end{aligned}$$

We want to bound the final term, which we can do in the following way:

$$\int_1^N\frac{\{t\}}{t^2}\,dt = \int_1^\infty\frac{\{t\}\,dt}{t^2}-\int_N^\infty\frac{\{t\}}{t^2}\,dt = \int_1^\infty\frac{\{t\}}{t^2}\,dt+O\left(\int_N^\infty\frac{1}{t^2}\,dt\right) = \int_1^\infty\frac{\{t\}}{t^2}\,dt+O\left(\frac{1}{N}\right).$$

Thus, we have the following result:

Proposition 1.10.

$$\sum_{n\le N}\frac{1}{n} = \log N+\left(1-\int_1^\infty\frac{\{t\}}{t^2}\,dt\right)+O\left(\frac{1}{N}\right).$$

Definition 1.11. $1-\int_1^\infty\frac{\{t\}}{t^2}\,dt=\gamma=0.577\ldots$ is known as Euler's constant.

Example 1.12. We can also apply this technique to Stirling's formula: $N!\sim\sqrt{2\pi N}(\frac{N}{e})^N$. We have

$$\sum_{n\le N}\log n = \int_{1^-}^{N^+}\log t\,d(A(t)) = N\log N-\int_{1^-}^{N^+}\frac{[t]}{t}\,dt.$$

Again, we need to approximate the last integral, which we can write as

$$\int_1^N\frac{t-\{t\}}{t}\,dt = (N-1)-\int_1^N\frac{\{t\}}{t}\,dt.$$

At this stage, we need one more idea. $\{t\}$ can be approximated by $\frac{1}{2}$. We should make this precise. That is,

$$\int_1^N\frac{\{t\}}{t}\,dt = \int_1^N\frac{\{t\}-\frac{1}{2}}{t}\,dt+\frac{1}{2}\log N.$$

Let $B(y)=\{y\}-\frac{1}{2}$; this B is for Bernoulli. Then we can write this final problematic integral as

$$\int_1^N\frac{1}{t}d\left(\int_1^t B(y)\,dy\right) = \int_1^t B(y)\,dy\bigg|_1^N + \int_1^N\frac{(\int_1^t B(y)\,dy)\,dt}{t^2}.$$

The endpoints give 0 and $B(y)$ is a periodic function with period 1, so $\int_n^{n+1}B(y)\,dy=0$. For $0<t<1$, we have $\int_0^t B(y)dy=\int_0^t(y-\frac{1}{2})\,dy=\frac{t^2}{2}-\frac{t}{2}$.

Therefore

$$\int_1^N\frac{(\int_1^t B(y)\,dy)\,dt}{t^2} = \int_1^\infty\frac{\frac{\{t\}^2}{2}-\frac{\{t\}}{2}}{t^2}\,dt+O\left(\frac{1}{N}\right).$$

Putting all of this together, we have shown that $\log N! = N\log N-N+\frac{1}{2}\log N+C+O(\frac{1}{N})$. This is Stirling's formula (up to evaluating some constants).

To be more precise, we work with $\int_0^t B(y)\,dy = \frac{\{t\}^2}{2} - \frac{\{t\}}{2}$. Subtract out the average, and apply the same method again. The average is $-\frac{1}{12}$, so our $O(\frac{1}{N})$ is actually a $-\frac{1}{12N}$, and we can continue in this manner.

This method is due to Euler, and it is known as *Euler-McLaurin summation*. The polynomials from this are called Bernoulli polynomials, and this is connected to evaluating the zeta function at even integers.

How do we compute $\zeta(2) = \frac{\pi^2}{6} = \sum_{n=1}^\infty \frac{1}{n^2}$? We don't want to directly sum it; instead, apply Euler-McLaurin summation.

We return to the primes.

Definition 1.13. The van Mangoldt function is

$$\Lambda(n) = \begin{cases} \log p & n = p^k \\ 0 & n \neq p^k. \end{cases}$$

A useful fact is that $\log n = \sum_{d|n} \Lambda(d)$, just by looking at the prime factorization of n.

Now, we can consider

$$\sum_{n\le N} \log n = \sum_{n\le N}\sum_{d|n} \Lambda(d) = \sum_{d\le N} \Lambda(d) \sum_{\substack{n\le N\\ d|n}} 1 = N\sum_{d\le N}\frac{\Lambda(d)}{d} + O\left(\sum_{d\le N}\Lambda(d)\right).$$

We'll show that this final term is $O(N)$.

2. 4/5

2.1. **Equivalent forms of counting primes.**

Last time, we introduced $\Lambda(n)$, which counts the primes with some weight. It is useful to consider $\psi(x) = \sum_{n\le x}\Lambda(n)$, and also natural to look at $\vartheta(x) = \sum_{p\le x}\log p$.

We have

$$\vartheta(x) = \int_{1^-}^x (\log t)\,d\pi(t) = \pi(t)\log t\big|_1^x - \int_1^x \frac{\pi(t)}{t}\,dt = \pi(x)\log x - \int_1^x \frac{\pi(t)}{t}\,dt$$

Recall that in the prime number theorem, we want that $\pi(x) \sim \mathrm{li}(x) = \frac{x}{\log x} + O(\frac{x}{(\log x)^2})$, so therefore we expect to see that $\vartheta(x) \sim x$. Therefore, we've shown that

$$\pi(x) = \frac{x}{\log x} + O\left(\frac{x}{(\log x)^2}\right) \implies \vartheta(x) = x + O\left(\frac{x}{\log x}\right).$$

In fact, we can be a bit more precise. Write $\pi(x) = \mathrm{li}(x) + E(x)$ for some error term E. Then

$$\vartheta(x) = \mathrm{li}(x)\log x + E(x)\log x - \int_1^x \frac{\mathrm{li}\,t}{t}\,dt - \int_1^x \frac{E(t)}{t}\,dt.$$

Then

$$\mathrm{li}(x)\log x - \int_2^x\left(\int_2^t \frac{dy}{\log y}\right)\frac{dt}{t} = \mathrm{li}(x)\log x - \int_2^x \frac{dy}{\log y}(\log x - \log y) = x + O(1).$$

What we are saying here is that $\pi(x) = \mathrm{li}(x) + E(x)$ means that $\vartheta(x) = x + E_0(x) + O(1)$ where $E_0(x) = E(x)\log x - \int_1^x \frac{E(t)}{t}\,dt$. So the nice asymptotic formula of $\pi(x)$ translates to a much nicer asymptotic formula for $\vartheta(x)$.

Conversely, given information about $\vartheta(x)$, we can get information about $\pi(x)$. Then by partial summation,

$$\pi(x) = \int_{2^-}^{x} \frac{1}{\log t} d(\vartheta(t)) = \frac{\vartheta(x)}{\log x} - \int_{2^-}^{x} \frac{\vartheta(t)}{t(\log t)^2}\, dt.$$

Hence

$$\vartheta(x) = x + E_0(x) \implies \pi(x) = \left(\frac{x}{\log x} + \int_{2^-}^{x} \frac{dt}{(\log t)^2}\right) + \frac{E_0(x)}{\log x} + \int_{2^-}^{x} \frac{E_0(t)}{t(\log t)^2}\, dt.$$

What is the first term? Integrate $\int_2^x \frac{dt}{\log t}$ by parts to see that it is just $\mathrm{li}(x)$.

The point is that the problem of finding an asymptotic formula for the number of primes is equivalent to the problem of counting the primes with weight $\log p$; one works with $\pi(x)$ and the other works with $\vartheta(x)$.

Remark. The Riemann hypothesis is equivalent to $\pi(x) = \mathrm{li}(x) + O(x^{1/2+\varepsilon})$, which is exactly the same strength as $\vartheta(x) = x + O(x^{1/2+\varepsilon})$.

To summarize, the prime number theorem is equivalent to the fact that $\vartheta(x) \sim x$ (and the relationship is a bit more precise than this). We claim that this is also equivalent to $\psi(x) \sim x$. Why is this? We have

$$\psi(x) = \sum_{n \le x} \Lambda(n) = \vartheta(x) + \vartheta(x^{1/2}) + \vartheta(x^{1/3}) + \cdots .$$

We can trivially bound $\theta(x^{1/2}) = O(\sqrt{x}\log x)$, and do this for each of the roots. So therefore we in fact even have that the Riemann hypothesis is equivalent to $\psi(x) = x + O(x^{1/2} + \varepsilon)$.

We've reduced our formulas to a number of equivalent forms. Why is one better than another? Later in the course, we'll see the explicit formula, which says that $\psi(x) = x - \sum_\rho \frac{x^\rho}{\rho}$ where the sum is over the zeros of $\zeta(s)$. In this way, $\psi(x)$ is a bit more natural than the other forms.

2.2. **Chebyshev bounds.** We will show that $(c + o(1))x \le \psi(x) \le (C + o(1))x$ for some constants c and C.

Take a large number N and look at the middle binomial coefficient $\binom{2N}{N}$, for which we have good upper and lower bounds. In particular,

$$\frac{4^N}{2N+1} \le \binom{2N}{N} \le 4^N.$$

Look at the prime factorization of $\binom{2N}{N}$.

Given a prime p, what is the largest power of p dividing $n!$? The answer is $[\frac{n}{p}] + [\frac{n}{p^2}] + \cdots$. So therefore

$$\binom{2N}{N} = \prod_{p \le 2N} p^{[\frac{2N}{p}] + [\frac{2n}{p^2}] + \cdots - 2[\frac{N}{p}] - 2[\frac{N}{p^2}] - \cdots}.$$

Recall that $\log n = \sum_{d|n} \Lambda(d)$ (check this) and $\log N! = \sum_{n \le N} \sum_{d|n} \Lambda(d) = \sum_{d \le N} \Lambda(d)[\frac{N}{d}]$.

In either way, we see that

$$\log \binom{2N}{N} = \sum_{d \le 2N} \Lambda(d) \left(\left[\frac{2N}{d}\right] - 2\left[\frac{N}{d}\right]\right).$$

Note that

$$[2x] - 2[x] = \begin{cases} 0 & \{x\} < \frac{1}{2} \\ 1 & \{x\} \geq \frac{1}{2}. \end{cases}$$

Therefore, we have

$$\log \binom{2N}{N} \leq \sum_{d \leq 2N} \Lambda(d) = \psi(2N),$$

and hence $\psi(2N) \geq N \log 4 - \log(2N+1)$. Then $\psi(2N) \geq 2N(\log 2 + o(1))$. So this gives $\psi(x) \geq x(\log 2 + o(1))$.

We also know that

$$\log \binom{2N}{N} \geq \sum_{N+1 \leq d \leq 2N} \Lambda(d) = \psi(2N) - \psi(N),$$

so we can conclude that $\psi(2N) - \psi(N) \leq N \log 4$. This yields $\psi(2x) - \psi(x) \leq x(\log 4 + o(1))$. This can give us a bound for $\psi(x)$ by dividing into intervals of dyadic blocks: $\psi(x) - \psi(x/2) \leq \frac{x}{2}(\log 4 + o(1))$, and so on. This gives $\psi(x) \leq x(\log 4 + o(1))$.

So what we've proved is the following theorem:

Theorem 2.1 (Chebyshev)**.**

$$A = \limsup_{x \to \infty} \frac{\psi(x)}{x} \leq \log 4$$
$$a = \liminf_{x \to \infty} \frac{\psi(x)}{x} \geq \log 2.$$

Also, looking at what we did above, we should also have that $\log 4 \geq \limsup \frac{\pi(x)}{x/\log x}$ and $\liminf \frac{\pi(x)}{x/\log x} \geq \log 2$.

There's one more beautiful thing that Chebyshev proved.

Theorem 2.2. *If $A = a$, i.e. $\lim_{x \to \infty} \frac{\psi(x)}{x}$ exists, then $a = A = 1$.*

There's Bertrand's postulate, which says that there is a prime between x and $2x$:

Theorem 2.3 (Bertrand's Postulate)**.** $\pi(2x) - \pi(x) > 0$.

These bounds come very close to proving Bertrand's postulate. We can try to tweak the bounds to do this, and in fact, this can be done without too much difficulty. We can use

$$\left[\frac{2N}{d}\right] - 2\left[\frac{N}{d}\right] = \begin{cases} 1 & N < d \leq 2N \\ 0 & \frac{2N}{3} < d \leq N \\ 1 & \frac{2N}{4} < d \leq \frac{2N}{3} \\ \cdots, \end{cases}$$

and by analyzing this more carefully, we can actually prove Bertrand's postulate.

We state some nice theorems:

Theorem 2.4.

$$\sum_{n\le x}\frac{\Lambda(n)}{n} = \log x + O(1)$$

$$\sum_{p\le x}\frac{\log p}{p} = \log x + O(1).$$

Proof. Clearly, the first statement implies the second. We prove the first formula.

$$\begin{aligned} N\log N + O(N) = \log N! &= \sum_{d\le N}\Lambda(d)\left[\frac{N}{d}\right] = \sum_{d\le N}\Lambda(d)\left(\frac{N}{d}+O(1)\right) \\ &= N\sum_{d\le N}\frac{\Lambda(d)}{d} + O\left(\sum_{d\le N}\Lambda(d)\right) = N\sum_{d\le N}\frac{\Lambda(d)}{d} + O(N), \end{aligned}$$

so therefore $\sum_{d\le N}\frac{\Lambda(d)}{d} = \log N + O(1)$. ☐

Why can't we use partial summation to prove the prime number theorem from this? We have

$$\sum_{n\le x} n\left(\frac{\Lambda(n)}{n}\right) = \int_{2^-}^{x} t\, d\left(\sum_{n\le t}\frac{\Lambda(n)}{n}\right) = t\sum_{n\le t}\frac{\Lambda(n)}{n}\Bigg|_{2^-}^{x} - \int_{2^-}^{x}\left(\sum_{n\le t}\frac{\Lambda(n)}{n}\right)dt.$$

Plugging in our estimates just yields that this is $O(x)$, which isn't good enough. Instead of $\log x + O(1)$, we need a more precise version, such as $\log x + a + O(1/\log n)$. This would imply the prime number theorem. The point is that partial summation will not manufacture information; we need to feed in something precise to get precise information out.

Now, let's give a proof of Chebyshev's theorem 2.2.

Proof of theorem 2.2. Suppose that $\lim_{x\to\infty}\frac{\psi(x)}{x} = c$ exists. Then

$$\sum_{n\le x}\frac{\Lambda(n)}{n} = \int_{2^-}^{x}\frac{1}{t}\,d\psi(t) = \frac{\psi(x)}{x} + \int_{2^-}^{x}\frac{\psi(t)}{t^2}\,dt = (c+o(1))\log x + O(1),$$

so therefore $c = 1$. ☐

So in the prime number theorem, the possibility that is hard to eliminate is the case where this limit does not exist and instead oscillates between some values.

Selberg gave a nice result where he considered products of two primes, and he showed that $a + A = 2$. Around the same time, Erdos and Selberg gave an elementary proof (without complex analysis) that $a = A = 1$. This proof is actually much harder than any proof using complex analysis, however.

Proposition 2.5.

$$\sum_{p\le x}\frac{1}{p} = \log\log x + C + O\left(\frac{1}{\log x}\right).$$

Proof. This follows from what we have done and partial summation. We have

$$A(x) = \sum_{p\le x}\frac{\log p}{p} = \log x + E(x)$$

where $E(x) = O(1)$. Then

$$\sum_{p\le x} \frac{1}{p} = \int_{2^-}^{x} \frac{1}{\log t}\, d(A(t)) = \frac{A(x)}{\log x} - \int_{2^-}^{x} \frac{1}{t(\log t)^2} A(t)\, dt.$$

Plugging in $A(x) = \log x + E(x)$ gives us what we want. The error term is

$$\int_2^x \frac{E(t)}{t(\log t)^2}\, dt = \int_2^\infty \frac{E(t)}{t(\log t)^2}\, dt + O\left(\int_x^\infty \frac{dt}{t(\log t)^2}\right),$$

which gives us some estimates for the constant c. $\square$

3. 4/10

3.1. **Properties of the Mobius function.** There is one more statement that is equivalent to the prime number theorem. This is a bit harder than the statements we had before, but it involves an important function.

Recall that

$$\mu(n) = \begin{cases} 0 & p^2 \mid n \text{ for some } p \\ (-1)^k & n = p_1 \dots p_k \text{ for distinct primes } p_k. \end{cases}$$

This is a mysterious function, and we don't understand a lot about it. A guiding principle is that this is ± 1 with around equal probability; it is very much related to counting coin tosses. We can think of $\mu(n)$ as being random: not correlated with anything else. This is about as vague as a statement can be. Can we find a polynomial time algorithm to compute $\mu(n)$?

We expect a sum of $\mu(n)$ to more-or-less cancel out. We expect to have that

$$\sum_{n\le x} \mu(n) = o(x),$$

and in fact, we'll see later in this lecture that this is equivalent to the prime number theorem. In addition, $\sum_{n\le x} \mu(n) = O(x^{1/2+\varepsilon})$ is equivalent to the Riemann Hypothesis.

Sometimes in number theory we are interested in some arithmetic functions $f : \mathbb{N} \to \mathbb{C}$. There is the idea of the Dirichlet convolution, which is $(f \star g)(n) = \sum_{d|n} f(d)g(\frac{n}{d})$. We may find it convenient to study $F(s) = \sum_{n=1}^\infty \frac{f(n)}{n^s}$ and $G(s) = \sum_{n=1}^\infty \frac{g(n)}{n^s}$. Then Dirichlet convolution means that

$$F(s)G(s) = \sum_{n=1}^\infty \frac{(f \star g)(n)}{n^s}.$$

The Mobius function allows us to invert various convolutions. Let 1 be the function that is always one. Then

$$\sum_{d|n} \mu(d) = \sum_{d|n} \mu\left(\frac{n}{d}\right) = (1 \star \mu)(n) = \delta(n) = \begin{cases} 1 & n = 1 \\ 0 & \text{otherwise.} \end{cases}$$

There is also the Mobius inversion formula. Say we have a function f and $F = 1\star f = f\star 1$. Then

$$F(n) = \sum_{d|n} f\left(\frac{n}{d}\right) = \sum_{d|n} f(d).$$

We have $f(n) = (f \star \delta)(n) = (f \star 1 \star \mu)(n) = (F \star \mu)(n)$.

We've only dealt with this in one context so far, where we showed that $\log = 1 \star \Lambda$. So we can write that $\Lambda = (\mu \star \log)$, which means that $\Lambda(n) = \sum_{d|n} \mu(d) \log(n/d)$. This is sort of morally why the Mobius function is connected with primes: We understand logs, and if we understand something about the Mobius function, then we can get information about average values of $\Lambda(n)$, which is the prime number theorem.

One thing about the Mobius function is that $\mu = \mu \star \mu \star 1$.

3.2. **Divisor function.** Consider the divisor function $d(n) = \#\{d \mid n\} = \sum_{d|n} 1 = 1 \star 1$.

Eventually, we will consider $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$, which converges absolutely if $\operatorname{Re} s > 1$. Thinking about convolutions yields that

$$\sum_{n=1}^{\infty} \frac{d(n)}{n^s} = \zeta(s)^2.$$

Our goal is to consider how large the divisor function can get, and how big it is on average.

Note that $d(n) = \sum_{ab=n} 1 \leq 2 \sum_{a \leq \sqrt{n}} 1 + O(1)$, so therefore $d(n) \leq 2\sqrt{n}$.

Proposition 3.1. *$d(n) \ll n^{\varepsilon}$ for any $\varepsilon > 0$.*

Proof. For every ε and every n, we can look at $\frac{d(n)}{n^{\varepsilon}}$. Write $n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$; here, p_j are distinct primes and $\alpha_j > 0$. Note that d is a multiplicative function, i.e. $d(mn) = d(m)d(n)$ if $(m, n) = 1$. Then $d(n) = \prod_{j=1}^{k} (\alpha_j + 1)$. Therefore,

$$\frac{d(n)}{n^{\varepsilon}} = \prod_{j=1}^{k} \frac{\alpha_j + 1}{p_j^{\varepsilon a_j}}.$$

Given p, we have

$$\frac{\alpha + 1}{p^{\alpha \varepsilon}} \leq \frac{\alpha_p + 1}{p^{\alpha_p \varepsilon}}$$

for some $\alpha_p \geq 0$ and $\alpha_p \in \mathbb{N}$.

If p is large, then $\alpha_p = 0$ because $\frac{2}{p^{\varepsilon}}$ will be < 1. Then

$$\frac{d(n)}{n^{\varepsilon}} \leq \prod_{p} \frac{\alpha_p + 1}{p^{\alpha_p \varepsilon}} = \text{constant}(\varepsilon)$$

because the terms in the product will eventually all be 1. $\square$

We can make this more precisely. A result of Ramanujan is $d(n) \leq 2^{(1+o(1)) \frac{\log n}{\log \log n}}$.

Now, we will prove something about the average size of $d(n)$. In fact, on average, $d(n) \approx \log n$. First we consider a crude bound, and then we will refine it.

$$\begin{aligned} \sum_{n \leq x} d(n) &= \sum_{n \leq x} \sum_{d|n} 1 = \sum_{d \leq x} \sum_{\substack{n \leq x \\ d|n}} 1 = \sum_{d \leq x} \left(\frac{x}{d} + O(1) \right) \\ &= x \sum_{d \leq x} \frac{1}{d} + O(x) = x \left(\log x + \gamma + O(1/x) \right) + O(x) = x \log x + O(x). \end{aligned}$$

This error term is very bad because we are always accumulating an error of 1 for each term. There is a classical problem to find a better error term.

Problem 3.2 (Dirichlet's Divisor Problem)**.** Find a formula with a better (best?) error term.

Theorem 3.3.

$$\sum_{n\le x} d(n) = x\log x + (2\gamma-1)x + O(\sqrt{x}).$$

A conjecture of Dirichlet is that this $O(\sqrt{x})$ term is actually $O(x^{1/4+\varepsilon})$.

Remark. This is similar to a problem that Gauss had: Count the number of lattice points in a circle. This is like the area, with an error term of the order of the circumference.

Proof. Here, we are trying to count points (a,b) with $ab = n \le x$. So we are counting lattice points under a hyperbola $ab = x$.

Pick a point (A,B) on the hyperbola. Group the lattice points under the hyperbola into three categories:

I. $a \le A$
II. $b \le B$
III. $a \le A, b \le B$.

We want to count I + II − III. (Draw a picture.)

In region I, we have

$$\sum_{a\le A}\sum_{b\le \frac{x}{a}} 1 = \sum_{a\le A}\left(\frac{x}{a}+O(1)\right) = x\left(\log A+\gamma+O\left(\frac{1}{A}\right)\right)+O(A) = x\log A + x\gamma + O(A+B).$$

In region II, by symmetry, we have $x\log B + x\gamma + O(A+B)$. Finally, in region III, we have $(A+O(1))(B+O(1)) = x+O(A+B)$.

Then

$$\mathrm{I}+\mathrm{II}-\mathrm{III} = x\log AB + 2\gamma x - x + O(A+B) = x\log x + (2\gamma-1)x + O(\sqrt{x})$$

by choosing $A = B = \sqrt{x}$. $\square$

Note that $\log n$ is the average size of $d(n)$, but it is not the usual size. There are a few values of $d(n)$ which are very large, and they increase the average.

As another example, consider $\omega(n)$ is the number of prime divisors of n. This has a usual size equal to its average size, which is approximately $\log\log n$.

Another result of Ramanujan is that

$$\sum_{n\le x} d(n)^2 \sim Cx(\log x)^3,$$

which is not what we would expect if the usual value of $d(n)$ is $\log x$; the few large values get amplified even more.

4. 4/12

Today, we will prove that the prime number theorem is equivalent to the cancellation in the Mobius function: $\sum_{n\le x}\mu(n) = o(x)$.

4.1. **Mobius cancellation implies PNT.** Let $M(x) = \sum_{n \le x} \mu(n)$. First, we show that $M(x) = o(x)$ implies that $\psi(x) \sim x$. Note that $\Lambda = \mu \star \log$. We would like to replace $\log n$ by $d(n) + C$ for some constant C; the two sides have approximately equal averages. Then instead of considering $\mu \star \log$, we need to look at $(\mu \star d) + C(\mu \star 1)$. In fact, $\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}$ because $\zeta(s) \sum \frac{\mu(n)}{n^s} = \sum \frac{(1 \star \mu)(n)}{n^s} = 1$. So this means that we hope to replace $\Lambda = \mu \star \log$ with $1(n) + C\delta(n)$, which has average value 1 and hence gives us what we want.

We have

$$\sum_{n \le N} \log n = N \log N - N + O(\log N)$$
$$\sum_{n \le N} d(n) = N \log N + (2\gamma - 1)N + O(\sqrt{N}),$$

so therefore

$$\sum_{n \le N} (\log n - d(n) + 2\gamma) = O(\sqrt{N}).$$

So this is our candidate for approximating $\log n$: by $d(n) - 2\gamma$. Let $b(n) = \log n - d(n) + 2\gamma$, so we know that $\sum_{n \le x} b(n) = O(\sqrt{x})$. Also, $\sum_{n \le x} |b(n)| = O(x \log x)$.

So write $\log n = d(n) - 2\gamma + b(n)$. Then $\Lambda(n) = (\mu \star \log)(n) = 1(n) - 2\gamma\delta(n) + (\mu \star b)(n)$, so therefore

$$\psi(x) = \sum_{n \le x} \Lambda(n) = [x] - 2\gamma + \sum_{n \le x} (\mu \star b)(n).$$

This gives us the main term in the prime number theorem, and next we need to show that $\sum_{n \le x} (\mu \star b)(n) = o(x)$. For this, we use the hyperbola method. Then

$$(\mu \star b)(n) = \sum_{rs = n} \mu(r) b(s).$$

As before, we have three cases: (1) $r \le R$; (2) $s \le S$; (3) $r \le R, s \le S$.

For (1), we have

$$\sum_{r \le R} \mu(r) \sum_{s \le x/r} b(s) = O\left(\sum_{r \le R} \frac{\sqrt{x}}{\sqrt{r}} |\mu(r)| \right) = O(\sqrt{xR}) = O\left(\frac{x}{\sqrt{S}} \right).$$

For (2), we use our hypothesis that for all $\varepsilon > 0$, if y is sufficiently large then $|\sum_{n \le y} \mu(n)| \le \varepsilon y$. R is taken large enough so that if $y > R$ then this estimate holds. Then

$$\left| \sum_{s \le S} b(s) \sum_{r \le x/s} \mu(r) \right| \le \varepsilon \sum_{s \le S} \frac{x}{s} |b(s)| \le \varepsilon\sigma \sum_{s \le S} |b(s)| = O(\varepsilon x S(\log S)).$$

For (3), we could in fact have done the same thing:

$$\left(\sum_{s \le S} b(s) \right) \left(\sum_{r \le R} \mu(r) \right) = O(\sqrt{S} \varepsilon R) = O(\varepsilon x),$$

so this term doesn't matter very much.

Now we pick values for R and S. (1) can be made small by making S large. Choose $S = \frac{1}{\sqrt{\varepsilon}}$. Then (1) is bounded by $x\varepsilon^{1/4}$, and (2) is also sufficiently small; this completes the proof.

Note that we have been wasteful in this procedure by throwing away the $\frac{1}{s}$ in (2). We could use this information through partial summation. More precisely, we can prove that $\sum_{s\le S} \frac{|b(s)|}{s} \ll (\log S)^2$, which is better than what we got.

4.2. **PNT implies Mobius cancellation.** We will now show that $\psi(x) \sim x$ implies that $M(x) = o(x)$.

This starts out with a trick. Consider

$$M(x)\log x = \sum_{n\le x} \mu(n)\log x = \sum_{n\le x} \mu(n)\log n + \sum_{n\le x} \mu(n)\log(x/n).$$

It turns out that $\log x$ is around $\log n$ because most numbers are within a constant factor of x. So we expect the second sum to be small. That is,

$$\sum_{n\le x} \mu(n)\log(x/n) = O\left(\sum_{n\le x}\log(x/n)\right) = O\left([x]\log x - \sum_{n\le x}\log n\right) = O(x).$$

Then

$$M(x) = \frac{1}{\log x}\sum_{n\le x}\mu(n)\log n + O\left(\frac{x}{\log x}\right).$$

We claim that $-\mu(n)\log n = (\mu \star \Lambda)$. This is something that we can check easily. Here is another way of thinking about this: If $F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$ then $F'(x) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}(-\log n)$. Therefore, $-\mu(n)\log n$ are the coefficients of

$$\left(\frac{1}{\zeta(s)}\right)' = -\frac{\zeta'(s)}{\zeta(s)^2} = \left(-\frac{\zeta'}{\zeta}(s)\right)\left(\frac{1}{\zeta(s)}\right).$$

Note that $-\zeta'(s)$ has coefficients of log, so then $-\frac{\zeta'}{\zeta}(s)$ has coefficients of $\mu \star \log = \Lambda$, and hence this shows that $-\mu(n)\log n = (\mu \star \Lambda)$. Therefore, we have

$$M(x)\log x + O(x) = \sum_{n\le x}\mu(n)\log n = -\sum_{rs\le x}\mu(r)\Lambda(s) = -\sum_{r\le x}\mu(r)\left(\psi\left(\frac{x}{r}\right) - \frac{x}{r}\right) - \sum_{r\le x}\mu(r)\frac{x}{r}.$$

Observe that the final term is

$$\sum_{r\le x}\mu(r)\left(\sum_{s\le x/r} 1 + O(1)\right) = O(x) + \sum_{n\le x}\left(\sum_{n=rs}\mu(r)\cdot 1\right) = O(x).$$

Remark. So we've shown that

$$\sum_{n\le x}\frac{\mu(n)}{n} = O(1).$$

In fact, the prime number theorem is equivalent to $\sum_{n\le x}\frac{\mu(n)}{n} \to 0$ as $x\to\infty$.

Now, we have

$$|M(x)\log x| + O(x) = \left|\sum_{r\le x}\mu(r)\left(\psi\left(\frac{x}{r}\right) - \frac{x}{r}\right)\right|.$$

We use the hypothesis that if $y \ge z = z(\varepsilon)$ then $|\psi(y) - y| \le \varepsilon y$. That means that we can write this sum as (using the Chebyshev bounds for the second sum)

$$\begin{aligned}
&= \left|\sum_{r\le x/z}\mu(r)\left(\psi\left(\frac{x}{r}\right) - \frac{x}{r}\right) + \sum_{x/r<r\le x}\mu(r)\left(\psi\left(\frac{x}{r}\right) - \frac{x}{r}\right)\right| \\
&\le \varepsilon x\sum_{r\le x/z}\frac{|\mu(r)|}{r} + C\sum_{x/z<r\le x}|\mu(r)|\frac{x}{r} \le \varepsilon x\log x + Cx\log z.
\end{aligned}$$

Therefore, we have

$$M(x) \le O\left(\frac{x}{\log x}\right) + \varepsilon x + xC\frac{\log z}{\log x},$$

and when x is large enough, we have that $|M(x)| \le 2\varepsilon x$.

5. 4/17

Today we will discuss the mean and variance of arithmetic functions.

5.1. **Mean and variance of arithmetic functions.** Recall that $d(n)$ is the number of divisors of n and $\omega(n)$ is the number of prime divisors of n. As we'll see, these two functions behave very differently. We will also consider $\Omega(n)$ is the number of prime divisors counted with multiplicity.

5.1.1. *Mean of* $\omega(n)$. First, we want to find their mean. Recall that

$$\sum_{n\le x} d(n) = x\log x + O(x).$$

For $\omega(n)$, we have

$$\begin{aligned}
\sum_{n\le x}\omega(n) &= \sum_{n\le x}\sum_{p|n} 1 = \sum_{p\le x}\sum_{\substack{p|n\\ n\le x}} 1 = \sum_{p\le x}\left\lfloor\frac{x}{p}\right\rfloor \\
&= \sum_{p\le x}\left(\frac{x}{p} + O(1)\right) = x\sum_{p\le x}\frac{1}{p} + O\left(\frac{x}{\log x}\right) = x\log\log x + O(x).
\end{aligned}$$

This means that if $n \approx x$ then the average value of $\omega(n)$ is $\log\log x$.

Exercise 5.1. The same estimate is true for $\Omega(n)$: $\sum_{n\le x}\Omega(n) = x\log\log x + O(x)$.

5.1.2. *Variance of $\omega(n)$.* Next, we will compute the variance. That is,

$$\begin{aligned}\sum_{n\le x}(\omega(n)-\log\log x)^2 &= \sum_{n\le x}\omega(n)^2 - 2\log\log x\sum_{n\le x}\omega(n) + x(\log\log x)^2 + O((\log\log x)^2)\\ &= \sum_{n\le x}\omega(n)^2 - x(\log\log x)^2 + O(x\log\log x).\end{aligned}$$

Also, we have

$$\begin{aligned}\sum_{n\le x}\omega(n)^2 &= \sum_{n\le x}\left(\sum_{p\mid n}1\right)^2 = \sum_{n\le x}\sum_{p\mid n}\sum_{q\mid n}1 = \sum_{p,q\le x}\sum_{\substack{p\mid n, q\mid n\\ n\le x}}1\\ &= \sum_{\substack{p,q\le x\\ p\ne q, pq\le x}}\left(\frac{x}{pq}+O(1)\right) + \sum_{p\le x}\left(\frac{x}{p}+O(1)\right)\end{aligned}$$

noting that the last sum is $O(x\log\log x)$,

$$= \sum_{\substack{p,q\le x\\ pq\le x}}\left(\frac{x}{pq}+O(1)\right) - \sum_{\substack{p=q\\ p^2\le x}}\left(\frac{x}{p^2}+O(1)\right) + O(x\log\log x),$$

and observing that the middle sum is $O(x)$, we've shown that

$$\sum_{n\le x}\omega(n)^2 = \sum_{\substack{p,q\le x\\ pq\le x}}\left(\frac{x}{pq}+O(1)\right) + O(x\log\log x).$$

Now, we have

$$x\left(\sum_{p\le\sqrt{x}}\frac{1}{p}\right)^2 \le \sum_{\substack{p,q\le x\\ pq\le x}}\frac{x}{pq} \le x\left(\sum_{p\le x}\frac{1}{p}\right)^2$$

where on the left we undercount by cases were $p\approx x/2$ and on the right we overcount by cases where $p,q\approx x$. Approximating each side shows that they are both $x(\log\log x)^2 + O(x\log\log x)$, so therefore we see that $\sum_{n\le x}\omega(n)^2 = x(\log\log x)^2 + O(x\log\log x)$. Then

$$\sum_{n\le x}(\omega(n)-\log\log x)^2 = O(x\log\log x).$$

Remark. To contrast, this is very different from what the divisor function looks like. In this case, we have

$$\sum_{n\le x}d(n)^2 = Cx(\log x)^3 + O(x(\log x)^2),$$

and

$$\sum(d(n)-\log x)^2 \sim x(\log x)^3.$$

This means that the variance of $d(n)$ is much higher, and there are a few large values of $d(n)$ that are dominating the mean.

In fact, we claim that the variance of $\omega(n)$ is especially small, so there are very few large values of $\omega(n)$. Let $\mathcal{E}$ be the exceptionally large values of $\omega(n)$, i.e. $\mathcal{E} = \{n \leq x \mid |\omega(n) - \log\log x| \geq A\sqrt{\log\log x}\}$. Note that

$$x \log\log x \gg \sum_{n \leq x} (\omega(n) - \log\log x)^2 \geq \sum_{n \in \mathcal{E}} \left(A\sqrt{\log\log x}\right)^2 = |\mathcal{E}|^2 A^2 \log\log x,$$

so therefore $|\mathcal{E}| \ll \frac{x}{A^2}$. In fact, let $A = A(x)$ such that $A(x) \to \infty$ as $x \to \infty$, e.g. $A(x) = (\log\log x)^{1/4}$. Then

$$\#\{n \leq x \mid |\omega(n) - \log\log x| \geq (\log\log x)^{3/4}\} \ll \frac{x}{(\log\log x)^{3/4}}.$$

This means that asymptotically, almost no n have $|\omega(n) - \log\log x| \geq (\log\log x)^{3/4}$.

Exercise 5.2. We also have

$$\sum (\omega(n) - \log\log n)^2 \ll x \log\log x,$$

which can be done using the triangle inequality in the same way. Also,

$$\sum (\Omega(n) - \log\log x)^2 \ll x \log\log x.$$

Corollary 5.3. *For almost all n, we have $\omega(n) \sim \Omega(n) \sim \log\log n$.*

On the other hand, $d(n)$ has many large values. Recall that $\frac{1}{x}\sum_{n \leq x} d(n) = \log x + O(1)$, but most values of the divisor function are not that big.

Observe that $2^{\omega(n)} \leq d(n) \leq 2^{\Omega(n)}$ for all n. Then for almost all n, we have that $2^{\log\log n + O(1)} \leq d(n) \leq 2^{\log\log n + O(1)}$, so therefore, for almost all n, $d(n) = (\log n)^{\log 2 + o(1)}$, which is significantly less than its mean value. This is because we are throwing away a density zero set with large values of $d(n)$.

5.2. **Applications of Dirichlet convolution and Mobius inversion.** Recall that $(a \star b)(n) = \sum_{d|n} a(d) b(\frac{n}{d})$ and $\sum_{d|n} \mu(d) = \delta(n)$. Alternatively, $f(n) = \sum_{d|n} g(d)$ if and only if $g(n) = \sum_{d|n} \mu(d) f(\frac{n}{d})$. Morally, $1^{-1} = \mu$ in the Dirichlet convolution algebra, i.e. $g \star 1 = f$ if and only if $f \star \mu = g$.

Example 5.4. One application is square-free numbers. Let

$$1_{\text{sqfree}} = \begin{cases} 1 & n \text{ is square-free} \\ 0 & \text{otherwise.} \end{cases}$$

The number of square-free numbers $\leq x$ is $\sum_{n \leq x} 1_{\text{sqfree}}(n)$.

Every integer n is uniquely expressible as $n = a^2 b$ where b is squarefree. Then by Mobius inversion,

$$1_{\text{sqfree}}(n) = \delta(a) = \sum_{d|a} \mu(d) = \sum_{d^2|n} \mu(d),$$

so therefore

$$\# \text{ square-free} \le x = \sum_{n\le x} 1_{\text{sqfree}}(n) = \sum_{n\le x}\sum_{d^2|n}\mu(d) = \sum_{d\le\sqrt{x}}\mu(d)\sum_{\substack{d^2|n\\ n\le x}} 1$$
$$= \sum_{d\le\sqrt{x}}\mu(d)\left(\frac{x}{d^2}+O(1)\right) = x\sum_{d\le\sqrt{x}}\frac{\mu(d)}{d^2}+O(\sqrt{x}).$$

We know that $\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty}\frac{\mu(n)}{n^s}$, so therefore $\frac{1}{\zeta(2)} = \sum_{n=1}^{\infty}\frac{\mu(n)}{n^2}$. Then we have that $\sum_{d\le\sqrt{x}}\frac{\mu(d)}{d^2} = \frac{1}{\zeta(2)}+O(\frac{1}{\sqrt{x}})$. Finally, this shows that

$$= \frac{1}{\zeta(2)}x+O(\sqrt{x}).$$

This means that approximately $\frac{6}{\pi^2}$ of integers are square-free.

Example 5.5. Recall that the Euler φ-function is $\varphi(n) = \#\{d\le n \mid \gcd(d,n)=1\}$. For example, $\varphi(p)=p-1$. What is the average value of $\varphi(n)$? We again do this by Mobius inversion.

Proposition 5.6. *$\sum_{d|n}\varphi(d) = n$.*

Proof. φ is multiplicative, so it suffices to check this for prime powers. Then we just want that $p^k = \sum_{j=0}^{k}\varphi(p^j) = \sum_{j=0}^{k}(p^j-p^{j-1})$, which is true because it telescopes. $\square$

By Mobius inversion, we then see that

$$\varphi(n) = \sum_{d|n}\mu(d)\frac{n}{d}.$$

It is slightly more natural to look at

$$\sum_{n\le x}\frac{\varphi(n)}{n} = \sum_{n\le x}\sum_{d|n}\frac{\mu(d)}{d} = \sum_{d\le x}\frac{\mu(d)}{d}\sum_{\substack{d|n\\ n\le x}}1$$
$$= \sum_{d\le x}\frac{\mu(d)}{d}\left(\frac{x}{d}+O(1)\right) = x\sum_{d\le x}\frac{\mu(d)}{d^2}+O\left(\sum_{d\le x}\frac{1}{d}\right) = \frac{x}{\zeta(2)}+O(\log x).$$

To get information about $\sum\varphi(n)$, we apply partial summation:

$$\sum_{n\le x}\varphi(n) = \int_{1^-}^{x} t\,d\left(\sum_{n\le t}\frac{\varphi(n)}{n}\right) = x\sum_{n\le x}\frac{\varphi(n)}{n} - \int_{1}^{x}\sum_{n\le t}\frac{\varphi(n)}{n}\,dt$$
$$= \frac{x^2}{\zeta(2)}+O(x\log x) - \int_{1^-}^{x}\left(\frac{t}{\zeta(2)}+O(\log t)\right)dt$$
$$= \frac{1}{2\zeta(2)}x^2+O(x\log x).$$

Today we will discuss some more applications of Dirichlet convolution and the hyperbola method.

Consider the arithmetic function

$$d_k(n) = \sum_{n=a_1\dots a_k} 1.$$

This is the number of ways of writing n as a product of k factors. Note that $d(n) = d_2(n)$. Recall that $\sum_{n=1}^{\infty} \frac{d(n)}{n^s} = \zeta(s)^2$; it turns out that $\sum_{n=1}^{\infty} \frac{d_k(n)}{n^s} = \zeta^k(s)$.

Furthermore, let $1(n)$ denote the constant sequence of ones; then $d(n) = 1 \star 1$. Also, $d_k(n) = (1 \star d_{k-1})(n) = (1 \star \cdots \star 1)(n)$ (k times).

Recall the hyperbola method: Say $a(n)$ and $b(n)$ are arithmetic sequences. Let $c(n) = (a \star b)(n)$. Let $\sum_{n\le x} a(n) = A(x)$; $\sum_{n\le x} b(n) = B(n)$; $\sum_{n\le x} c(n) = C(n)$. Then

$$C(x) = \sum_{\ell \le y} a(\ell) B\left(\frac{x}{\ell}\right) + \sum_{m\le x/y} b(m) A\left(\frac{x}{m}\right) - A(y)B\left(\frac{x}{y}\right).$$

We want to find the mean value of $d_k(n)$.

Lemma 6.1.

$$\sum_{n\le x} d_k(n) = xP_{k-1}(\log x) + O\left(x^{1-\frac{1}{k}+\varepsilon}\right)$$

where P_{k-1} is a polynomial of degree $k-1$.

Proof. We proceed by induction. The base case is $k=2$, and we already showed that

$$\sum_{n\le x} d(n) = x\log x + (2\gamma - 1)(x) + O(x^{1/2}).$$

For the inductive step, we use the hyperbola method. Let $a(n) = 1(n)$ so $A(x) = [x]$; let $b(n) = d_{k-1}(n)$, so that $B(x) = xP_{k-2}(\log x) + O(x^{1-\frac{1}{k-1}+\varepsilon})$. This gives

$$\begin{aligned}\sum_{n\le x} d_k(n) &= \sum_{n\le x} (1\star d_{k-1})(n)\\ &= \sum_{\ell\le y}\left(\frac{x}{\ell}P_{k-2}\left(\log\frac{x}{\ell}\right) + O\left(\left(\frac{x}{\ell}\right)^{1-\frac{1}{k-1}+\varepsilon}\right)\right)\\ &+ \sum_{m\le x/y} d_{k-1}(m)\left[\frac{x}{m}\right] - [y]\left(\frac{x}{y}P_{k-2}\left(\log\frac{x}{y}\right) + O\left(\left(\frac{x}{y}\right)^{1-\frac{1}{k-1}+\varepsilon}\right)\right).\end{aligned}$$

The first term is

$$x\sum_{\ell\le y}\frac{1}{\ell}P_{k-2}(\log x - \log \ell) = xP_{k-1}(\log x, \log y).$$

Note that P_n is some polynomial that might change from line to line. The next term is

$$\sum_{\ell\le y} O\left(\left(\frac{x}{\ell}\right)^{1-\frac{1}{k-1}+\varepsilon}\right) = O\left(x\left(\frac{x}{y}\right)^{-\frac{1}{k-1}+\varepsilon}\right).$$

Then

$$\sum_{m\le x/y} d_{k-1}(m)\left[\frac{x}{m}\right] = x\sum_{m\le x/y}\frac{d_{k-1}(m)}{m} + \sum_{m\le x/y} d_{k-1}(m)O(1).$$

The second sum is $O((\frac{x}{y})^{1+\varepsilon})$. The first sum can be evaluated via the induction hypothesis and partial summation. That is,

$$\begin{aligned}\sum_{m\le x/y}\frac{d_{k-1}(m)}{m} &= \int_1^{x/y}\frac{1}{u}\,d\left(\sum_{m\le n} d_{k-1}(m)\right) = \frac{1}{x/y}\sum_{m\le x/y} d_{k-1}(m) + \int_1^{x/y}\sum_{m\le n} d_{k-1}(m)\frac{1}{u^2}\,du\\ &= xP_{k-1}\left(\log\frac{x}{y}\right) + O\left(\left(\frac{x}{y}\right)^{1+\varepsilon}\right)\end{aligned}$$

after a lot of bookkeeping. Finally, the last term is

$$[y]\frac{x}{y}P_{k-2}\left(\log\frac{x}{y}\right) = \left(x + O\left(\frac{x}{y}\right)\right)P_{k-2}\left(\log\frac{x}{y}\right) = xP_{k-2}\left(\log\frac{x}{y}\right) + O\left(\left(\frac{x}{y}\right)^{1+\varepsilon}\right).$$

Putting all of this together, we have

$$\sum_{n\le x} d_k(n) = xP_{k-1}(\log x, \log y) + xP_{k-1}\left(\log\frac{x}{y}\right) + O\left(x\left(\frac{x}{y}\right)^{-\frac{1}{k-1}+\varepsilon} + \left(\frac{x}{y}\right)^{1+\varepsilon}\right).$$

Now, pick y in terms of x to minimize the error term. Let $y = x^\alpha$ for some $\alpha > 0$; we will solve for α. So we want the two error terms to be the same size, i.e.

$$x^{1-\frac{1}{k-1}+\frac{\alpha}{k-1}} = x^{1-\alpha},$$

so solving for α yields $\alpha = \frac{1}{k}$. Using this, we have

$$\sum_{n\le x} d_k(n) = xP_{k-1}(\log x) + O\left(x^{1-\frac{1}{k}+\varepsilon}\right). \qquad \square$$

An application of this is the "Selberg formula":

Proposition 6.2.

$$\sum_{p\le x}(\log p)^2 + \sum_{\substack{p\le x, q\le x\\ pq\le x}}(\log p)(\log q) = 2x\log x + O(x).$$

The idea is that counting primes is hard, but counting products of two primes is easier. The Selberg formula is the key step in elementary proofs of the prime number theorem.

Proof. Define

$$\Lambda_2(n) = \mu\star(\log n)^2 = \begin{cases} 0 & n \text{ has more than 2 distinct prime factors}\\ (\log p)(\log q) & n = p^k q^\ell \text{ for } k,\ell > 0\\ (\log p)^2 & n = p^k.\end{cases}$$

To show this equality, observe that

$$(1 \star \Lambda_2)(n) = \sum_{d|n} \Lambda_2(d) = \sum_{\substack{p^k || n \\ q^\ell || n}} \left(\sum_{j=1}^{k} \log p \right) \left(\sum_{j=1}^{\ell} \log q \right) = \left(\sum_{p^k || n} \log p^k \right)^2 = (\log n)^2.$$

Now, we see that $\sum_{n \le x} \Lambda_2(n) = 2x \log x + O(x)$ would imply the Selberg formula; higher prime powers contribute almost nothing. So it suffices to prove this.

Here's the clever part of the argument: Let $b(n) = (\log n)^2 - 2d_3(n) - c_1 d(n) - c_2 1(n)$ where $c_1, c_2 \in \mathbb{R}$ are constants. We choose c_1, c_2 so that $\sum_{n \le x} b(n) \ll x^{2/3+\varepsilon}$. Why can we hope to do this? Recall that $\sum_{n \le x} d_3(n) = xQ(\log x) + O(x^{2/3+\varepsilon})$ where Q is some quadratic polynomial, and $\sum_{n \le x} (\log n)^2 = x(\log x)^2 - 2x \log x + 2x + O((\log x)^2)$ by Euler-Maclaurin summation, and $\sum_{n \le x} d(x) = x \log x + (2\gamma - 1)x + O(x^{1/2})$. So we can pick c_1 and c_2 so that all of the large terms cancel.

Now

$$\Lambda_2(n) = \mu \star (\log)^2 = \mu \star (2d_3 + c_1 d + c_2 1 + b) = 2d + c_1 1 + c_2 \delta + (\mu \star b),$$

so

$$\begin{aligned} \sum_{n \le x} \Lambda_2(n) &= 2x \log x + O(x) + \sum_{n \le x} \sum_{d|n} \mu(d) b\left(\frac{n}{d}\right) = 2x \log x + O(x) + \sum_{d \le x} \mu(d) \sum_{\substack{d|n \\ n \le x}} b\left(\frac{n}{d}\right) \\ &= 2x \log x + O(x) + \sum_{d \le x} \mu(d) O\left(\left(\frac{x}{d}\right)^{0.7}\right) = 2x \log x + O(x) + O\left(x^{0.7} \sum_{d \le x} d^{-0.7}\right) \\ &= 2x \log x + O(x). \end{aligned}$$

□

7. 4/24

7.1. **Applications.**

Recall from last week that we showed that $\omega(n)$ is usually of size $\log \log n$. Here is a nice application of this.

7.1.1. *Multiplication table.* Take an $N \times N$ multiplication table. We get numbers $\le N^2$. How many distinct numbers are in the table?

Theorem 7.1 (Erdos). *The number of distinct elements in a multiplication table is $o(N^2)$.*

Proof. The typical number up to N^2 has $\sim \log \log(N^2) = \log(2 \log N) = \log \log N + \log 2$ prime factors. The typical entry in the multiplication table has $\sim 2 \log \log N$ prime factors ($\log \log N$ from each factor). So therefore a typical entry in the multiplication table is not a typical number up to N^2. □

Remark. In 2004, it was proved that the number of entries in a multiplication table is actually of size

$$\asymp \frac{N^2}{(\log N)^\alpha (\log \log N)^{3/2}}$$

where $\alpha = 0.086 \cdots = 1 - \frac{1}{\log 2}(1 - \log \frac{1}{\log 2})$.

7.1.2. *Consequence of Selberg.* Last time, we also proved Selberg's Identity 6.2. We have that $\vartheta(x) = \sum_{p\le x} \log p$ and hence $\sum_{p\le x}(\log p)^2 \sum \vartheta(x)\log x$. Then Selberg's Identity becomes

$$\vartheta x \log x + \sum_{p\le x}(\log p)\vartheta\left(\frac{x}{p}\right) \sim 2x\log x.$$

This is consistent with the prime number theorem, and it is some sort of average version of the prime number theorem. In fact, with sufficient cleverness, this can be used to prove the prime number theorem.

We can actually say a little more than that. Recall the Chebyshev bounds 2.1, where we defined

$$a = \liminf \frac{\psi(x)}{x}, \quad A = \limsup \frac{\psi(x)}{x}.$$

A nice consequence of Selberg's identity is that $a + A = 2$. This is because we can rewrite Selberg's identity as

$$\begin{aligned} 2x\log x &\sim \vartheta(x)\log x + \sum_{p\le x}(\log p)\vartheta\left(\frac{x}{p}\right) \\ &\le \vartheta(x)\log x + \sum_{p\le x}(\log p)(A+\varepsilon)\frac{x}{p} \le \vartheta(x)\log x + x(A+\varepsilon)\log x, \end{aligned}$$

so therefore $\vartheta(x) \ge (2 - A - \varepsilon)x$ and hence $a \ge 2 - A$. The other argument is similar and yields that $A \le 2 - a$.

7.1.3. *Smooth numbers.* Factoring algorithms often use smooth numbers.

We want to count $\Psi(x,y)$ is the number of $n \le x$ such that all primes $p \mid n$ satisfy $p \le y$. Write

$$\Psi(x,y) = [x] - \sum_{y<p\le x}\sum_{\substack{n\le x \\ p\mid n}} 1 = x + O(1) - \sum_{y\le p\le x}\left(\frac{x}{p} + O(1)\right) = x + O\left(\frac{x}{\log x}\right) - x\sum_{y<p\le x}\frac{1}{p}.$$

Here $\sum_{y\le p\le x}\frac{1}{p} = \log\frac{\log x}{\log y} + O(\frac{1}{\log x})$, so this is

$$= x\left(1 - \log\frac{\log x}{\log y}\right) + O\left(\frac{x}{\log x}\right)$$

In the case where $y = x^{1/u}$ for $1 \le u \le 2$, we've shown that $\Psi(x,y) = x(1 - \log u) + O(x/\log x)$. So if we pick a random number x, it has a prime factor of size $> \sqrt{x}$ with probability $\log 2$.

What happens if $2 \le u \le 3$? We apply the principle of inclusion-exclusion:

$$x - \sum_{y\le p\le x}\left(\frac{x}{p} + O(1)\right) + \sum_{\substack{y<p,q\le x \\ pq\le x}}\left(\frac{x}{pq} + O(1)\right).$$

This will become a mess that we can tackle using partial summation.

Theorem 7.2. *$\Psi(x,y) = (\rho(u) + o(1))x$ where $y = x^{1/u}$. Here $\rho(u)$ is the Dickman-de Bruijn function given by $u\rho'(u) = -\rho(u-1)$ with the initial condition $\rho(u) = 1$ if $0 \le u \le 1$.*

Note that solving this for $1 \leq u \leq 2$ yields $\rho(u) = 1 - \log u$. The function $\rho(u)$ is always positive but is small for large u.

7.1.4. *Connection to permutations.* These kinds of elementary things that we've discussed are actually very robust phenomena that appear in other parts of mathematics.

Consider the symmetric group S_n of permutations on n elements. There are $n!$ such permutations, and we can decompose each into cycles. The number of n-cycles is $(n-1)! = \frac{1}{n}n!$.

How many cycles are there in a typical permutation? This is approximately $\log \log n! \approx \log n$. Also, the number of permutations with a cycle of length $\geq n/2$ is $\sim (\log 2)n!$. In analogy to above, we actually get the Dickman function again.

Problem 7.3. Here's a real world application. There are a hundred prisoners numbered from 1 to 100. There is a room with 100 boxes numbered 1 to 100, with slips of paper numbered 1 to 100 in some permutation. They can each open 50 boxes, hoping to see their own number inside. If they all see their number, they are all set free. They want to pick a strategy to maximize the probability that they can be set free.

If they did this at random, they have a $\frac{1}{2^{100}}$ chance of success. But if they pick the right strategy, they can win with probability 30%. This is kind of mind-boggling, and it is a cool problem to think about.

7.2. **Analytic methods.** We will discuss more general analytic methods.

Recall that when discussing Dirichlet series, we wrote down formal series $F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$. Now, we want to consider convergence.

One of the main things that we will study is the zeta function $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$. We can write $s = \sigma + it$ and ask: When does this converge?

When $\sigma > 1$, this series converges absolutely, and it defines a holomorphic function in this region. We can differentiate term-by-term to get

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

which will be absolutely convergent in the same region. This isn't a proof, but it should give us some confidence that we can justify this differentiation.

We can also express

$$\zeta(s) = \prod_p \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots \right) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}$$

This is true at least formally, because each number is expressed as a product of prime powers in precisely one way, and we can sum the geometric series. This also converges absolutely when $\operatorname{Re} s = \sigma > 1$. What does this mean?

In general, we relate the convergence $\prod_{n=1}^{\infty}(1+a_n)$ to the convergence of $\sum_{n=1}^{\infty} a_n$, because taking the log of the product will approximately yield the sum.

Definition 7.4. Absolute convergence of $\prod_{n=1}^{\infty}(1+a_n)$ means that $\sum_{n=1}^{\infty} |a_n| < \infty$.

This means that an absolutely convergent product equals zero if and only if one of the terms is zero.

Corollary 7.5. *$\zeta(\sigma + it) \neq 0$ if $\sigma > 1$.*

We want to get a definition of $\zeta(s)$ that is valid for all $s \in \mathbb{C}$ (except for $s \neq 1$). The key words here are analytic continuation.

We start with $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ for $\sigma > 1$, and we think of doing partial summation:

$$\zeta(s) = \int_{1^-}^{\infty} \frac{1}{y^s}\, d([y]) = \frac{[y]}{y^s}\bigg|_{1^-}^{\infty} - \int_{1^-}^{\infty} \left(\frac{-s}{y^{s+1}}\right)[y]\, dy = \int_{1^-}^{\infty} \frac{s}{y^{s+1}}(y - \{y\})\, dy$$
$$= s\int_1^{\infty} \frac{dy}{y^s} - s\int_1^{\infty} \frac{\{y\}}{y^{s+1}}\, dy = \frac{s}{s-1} - s\int_1^{\infty} \frac{\{y\}}{y^{s+1}}\, dy.$$

The second integral converges absolutely when $\operatorname{Re} s = \sigma > 0$, and it defines a holomorphic function in that region.

Proposition 7.6. *$\zeta(s)$ has a meromorphic continuation to the region $\sigma > 0$. It has a simple pole at $s = 1$ with residue 1. The Laurent series around 1 is*

$$\frac{1}{s-1} + \left(1 - \int_1^{\infty} \frac{\{y\}}{y^2}\, dy\right) + c_1(s-1) + \cdots = \frac{1}{s-1} + \gamma + c_1(s-1) + \cdots .$$

8. 4/26

8.1. **Analytic continuation of the zeta function.** We want to continue the zeta function into an even larger region. To do this, we use Euler-Maclaurin summation on the integral that remained from the preceding calculation:

$$\int_1^{\infty} \frac{\{y\}}{y^{s+1}}\, dy = \int_1^{\infty} \frac{\{y\} - \frac{1}{2}}{y^{s+1}}\, dy + \frac{1}{2}\int_1^{\infty} \frac{dy}{y^{s+1}}$$
$$= \int_1^{\infty} \frac{1}{y^{s+1}}\, d\left(\int_1^y (\{t\} - 1/2)\, dt\right) + \frac{1}{2s} = (s+1)\int_1^{\infty} \frac{\int_1^y (\{t\} - 1/2)\, dt}{y^{s+2}}\, dy + \frac{1}{2s}.$$

Now, this is absolutely convergent if $\operatorname{Re} s > -1$, so we have a further analytic continuation of $\zeta(s)$. Note that we do not have a pole at 0; we pick up a $\frac{1}{s}$, but it is multiplied by s. Note that $\zeta(0) = -\frac{1}{2}$.

Now, we can just repeat this calculation; $\int_1^y (\{t\} - 1/2)\, dt = \frac{1}{2}(\{y\}^2 - \{y\})$, and we can subtract out the average value and do the same procedure again. Note that we do not pick up any more poles.

In this way, we get an analytic continuation of $\zeta(s)$ to all of $\mathbb{C}$ except for $s = 1$, where there is a simple pole.

It is interesting to compute various values of $\zeta(s)$. For example, $\zeta(-1) = \frac{-1}{12}$, which was written by Ramanujan as $1 + 2 + 3 + 4 + \cdots = \frac{-1}{12}$. An interesting feature of the computation that we have done is that $\zeta(s)$ is rational at each of the negative integers.

We can now make sense of the Riemann Hypothesis. Note that $\zeta(-2) = \zeta(-4) = \cdots = 0$, but we don't know much about the complex zeros.

Here's another way to write down this analytic continuation. We can write that

$$\begin{aligned}\zeta(s) &= \sum_{n\le N}\frac{1}{n^s}+\int_{N^+}^{\infty}\frac{d[y]}{y^s}=\sum_{n\le N}\frac{1}{n^s}+\left[\frac{[y]}{y^s}\right]_{N^+}^{\infty}+s\int_{N^+}^{\infty}\frac{[y]}{y^{s+1}}\,dy\\ &=\sum_{n\le N}\frac{1}{n^s}-\frac{N}{N^s}+s\int_N^{\infty}\frac{dy}{y^s}-s\int_{N^+}^{\infty}\frac{\{y\}}{y^{s+1}}\,dy\\ &=\sum_{n\le N}\frac{1}{n^s}-N^{1-s}+\frac{s}{s-1}N^{1-s}-s\int_{N^+}^{\infty}\frac{\{y\}}{y^{s+1}}\,dy=\sum_{n\le N}\frac{1}{n^s}-\frac{N^{1-s}}{1-s}-s\int_{N^+}^{\infty}\frac{\{y\}}{y^{s+1}}\,dy.\end{aligned}$$

This is a small generalization of what we did before, where we did the same thing, but with $N=1$. This is an expression for $\zeta(s)$ whenever $\operatorname{Re}s>0$. This also gives us a way to compute $\zeta(\frac{1}{2}+25i)$; we can compute the sum up to a million terms, and bound the tail of the integral.

8.2. **Bounds for $\zeta(s)$.** Now, we want to get bounds on $\zeta(s)$ when $s=\sigma+it$ for $\sigma>0$. Consider the simplest case: $\sigma>1$. Then

$$|\zeta(\sigma+it)|=\left|\sum_{n=1}^{\infty}\frac{1}{n^{\sigma+it}}\right|\le\sum_{n=1}^{\infty}\frac{1}{n^{\sigma}}=\zeta(\sigma).$$

This bound goes to infinity as $\sigma\to 1$.

Another question is to get bounds for $|\zeta(1+it)|$. When $|t|$ is very small, we can use the Laurent expansion:

$$\zeta(1+it)=\frac{1}{it}+\gamma+c_1(it)+\cdots.$$

What about bounds for $|\zeta(1+it)|$ when $|t|\ge 1$? Or more generally, we might want a bound for $\zeta(s)$ in the region left after removing a neighborhood of radius $1/2$ around 1.

We will now bound $|\zeta(1+it)|$. We have

$$|\zeta(1+it)|\le\left|\sum_{n\le N}\frac{1}{n^{1+it}}\right|+\frac{1}{|it|}+|1+it|\int_N^{\infty}\frac{1}{y^2}\,dy\le\log N+O(1)+\frac{1+|t|}{N}\le\log(1+|t|)+O(1)$$

by choosing $N=[1+|t|]$.

We can use this to bound $\zeta(\sigma+it)$ for $\sigma>0$ and $|\sigma+it-1|\ge 1/2$ (i.e. away from the pole). That is,

$$\begin{aligned}|\zeta(\sigma+it)| &\le\sum_{n\le N}\frac{1}{n^{\sigma}}+\frac{N^{1-\sigma}}{|1-\sigma+it|}+(\sigma+|t|)\int_N^{\infty}\frac{dy}{y^{1+\sigma}}\\ &\le\frac{N^{1-\sigma}-1}{1-\sigma}+\frac{N^{1-\sigma}}{(1-\sigma)+|t|}+\frac{\sigma+|t|}{\sigma}N^{-\sigma}.\end{aligned}$$

As before, choose $N=[1+|t|]$ to get a reasonable bound. Restricting ourselves to the case $1-\varepsilon>\sigma>\varepsilon$, we get

$$|\zeta(\sigma+it)|\ll(1+|t|)^{1-\sigma}$$

To get a uniform bound for $\sigma>\varepsilon$, we can write

$$|\zeta(\sigma+it)|\ll(1+(1+|t|)^{1-\sigma})\log(1+|t|).$$

We will find something of this sort very useful.

8.2.1. *Perron's formula.* We want to do something with the zeta function, such as to prove the prime number theorem or to understand problems on arithmetic functions.

For example, we can use the zeta function to consider the k-divisor function $d_k(n)$ and estimate $\sum_{n \leq x} d_k(n)$.

The first step to doing this is called *Perron's formula*:

Proposition 8.1.

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{y^s}{s}\, ds = \begin{cases} 1 & y > 1 \\ \frac{1}{2} & y = 1 \\ 0 & y < 1 \end{cases}$$

for $y > 0$ and $c > 0$.

Remark. What do the bounds mean? This is a contour integral, and we evaluate on the line $c + it$. Does this integral make sense? The integral does not converge absolutely, and we have to be careful. We think of this integral as $\int_{c-i\infty}^{c+i\infty} = \lim_{T\to\infty} \int_{c-iT}^{c+iT}$.

Let's first do a simple case, when $y = 1$. In this case, we can compute the integral:

$$\frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{1}{s}\, ds = \frac{1}{2\pi}\int_{-T}^{T} \frac{dt}{c+it} = \frac{1}{2\pi}\int_0^T \left(\frac{1}{c+it} + \frac{1}{c-it}\right) dt = \frac{1}{2\pi}\int_0^T \frac{2c}{c^2+t^2}\, dt,$$

which is some arctangent that we can easily compute.

Note that this doesn't depend on c. Why should we expect this to happen? If we integrated along some other line, we can shift from one contour to the other. When we do this, we don't cross any singularities because we assume that $c > 0$, so therefore the integral should be the same along any contour.

We'll make this precise next time, but loosely, here's the plan: For $y < 1$, we have that $y^c \to 0$ as $c \to +\infty$. Then we move the line of integration to the right, so the integrand becomes small, so the integral is 0. For $y > 1$, we have $y^c \to 0$ as $c \to -\infty$, so we move the line of integration to the left. But the difference here is that we cross a singularity at 0, and the residue of the singularity is $y^0 = 1$.

If we do this, then we can try to understand $\sum_{n\leq x} a(n)$ by analyzing $A(s) = \sum_{n=1}^{\infty} \frac{a(n)}{n^s}$. To do this, we look at

$$\frac{1}{2\pi i}\int_{(c)} A(s)\frac{x^s}{s}\, ds = \frac{1}{2\pi i}\sum_{n=1}^{\infty} a(n)\int_{(c)} \left(\frac{x}{n}\right)^s \frac{ds}{s} = \sum_{n\leq x} a(n).$$

This is what we are aiming for, and we'll make this rigorous soon.

9. 5/1

Last time, we wanted to have a way of analytically identifying whether $n \leq n$ or $n > x$. To do this, we had the Perron formula.

9.1. **Results from complex analysis.**

Example 9.1. We compute

$$\frac{1}{2\pi i}\int_{(c)} y^s \frac{ds}{s(s+1)} = \begin{cases} 1 - \frac{1}{y} & y > 1 \\ 0 & y \leq 1. \end{cases}$$

for $c > 0$, $y > 0$.

When $y > 1$, we move the line of integration to the left, and pick up residues for the poles at $s = 0$ and $s = -1$; this yields a residue of $1 - \frac{1}{y}$.

When $y < 1$, we move the line of integration to the right and pass through no poles, so we just get zero.

What happens when $y = 1$? We still get 0, because we can just move to the right. Then the integral is bounded in size by $\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1}{c^2+t^2}\,dt \to 0$ as $c \to \infty$.

Here, our answer is nicer (continuous) because we integrated something that was nicer than in the Perron formula. If we integrate something even nicer, we can get out something better.

Example 9.2. We can define the *Gamma function* as

$$\Gamma(s) = \int_0^\infty e^{-t} t^s \frac{dt}{t}.$$

This converges for $\operatorname{Re} s > 0$. We also have the functional equation

$$\Gamma(s+1) = \int_0^\infty t^s\, d(-e^{-t}) = s\int_0^\infty t^{s-1} e^{-t}\,dt = s\Gamma(s).$$

From this, we can get a meromorphic continuation of $\Gamma(s)$ to $\mathbb{C}$.

This gives us a function with simple poles at $0, -1, -2, -3, \dots$. Then we have $\operatorname{Res}_{s=0}\Gamma(s) = 1$, $\operatorname{Res}_{s=1}\Gamma(s) = -1$, and in general, $\operatorname{Res}_{s=-n}\Gamma(s) = \frac{(-1)^n}{n!}$.

Note that directly from the definition, we have $|\Gamma(1+it)| \le \Gamma(1)$. This means that

$$|\Gamma(1+it)| = \frac{|\Gamma(2+it)|}{|1+it|} = \frac{\Gamma(2)}{|1+it|},$$

and we can repeat this. In fact, this yields that $|\Gamma(1+it)| \le \frac{C(k)}{(1+|t|)^k}$.

Now, we can compute

$$\frac{1}{2\pi i}\int_{(c)} y^s\Gamma(s)\,ds.$$

Given that the integrand is very rapidly decreasing, we expect that the result will be C^∞. In fact, we move the line of integration to the left. Here, we run into a bunch of singularities, but $\Gamma(s)$ is small on average. Moving to $-\infty$ yields

$$\frac{1}{2\pi i}\int_{(c)} y^s\Gamma(s)\,ds = \sum_{n=0}^\infty y^{-n}\frac{(-1)^n}{n!} = e^{-1/y}.$$

This looks like some smoothed version of the functions that we were considering before.

Here's another result that we'll do just for fun.

Example 9.3. We won't rigorously justify this, but it can be done and the answer is still correct. Consider $c > 1$, and compute

$$\frac{1}{2\pi i}\int_{(c)} \zeta(s)\Gamma(s)x^{-s}\,ds = \sum_{n=1}^\infty \frac{1}{2\pi i}\int_{(c)}\Gamma(s)(nx)^{-s}\,ds = \sum_{n=1}^\infty e^{-nx} = \frac{e^{-x}}{1-e^{-x}} = \frac{1}{e^x - 1}.$$

Next we evaluate this integral by moving the contour to the left. This is the step that is not obvious and needs proof, but we can do the calculation anyway. As we move the contour, we encounter poles at $1, 0, -1, -2, \ldots$, and pick up the corresponding residues. This gives

$$\frac{1}{2\pi i}\int_{(c)} \zeta(s)\Gamma(s)x^{-s}\,ds = \frac{1}{e^x - 1} = \frac{1}{x} + \sum_{n=0}^{\infty} \zeta(-n)\frac{(-1)^n}{n!}x^n.$$

So multiplying through by x yields

$$\frac{x}{e^x - 1} = 1 + \sum_{n=0}^{\infty} \zeta(-n)\frac{(-1)^n}{n!}x^{n+1}.$$

Forgetting the complex analysis now, it seems plausible that the Taylor series expansions on each side agree. In fact, Bernoulli numbers show up in the Taylor series expansion on the left hand side, and these are related to $\zeta(-n)$ somehow.

Exercise 9.4. We can use this to prove that $\zeta(-2) = \zeta(-4) = \cdots = 0$.

9.2. **Rigorous proof of Perron formula.** Now, we will prove the Perron formula rigorously.

Proof of Perron formula 8.1. We will work with the integral

$$\frac{1}{2\pi i}\int_{c-iT}^{c+iT} y^s \frac{ds}{s}.$$

First, consider the case $y < 1$; we want to move the contour to the right. We can write

$$\int_{c-iT}^{c+iT} = \int_{c-iT}^{d-iT} + \int_{d-iT}^{d+iT} - \int_{c+iT}^{d+iT}.$$

We just have to estimate the integrals on these three other sides. First, look at the horizontal integral (and write $s = \sigma - iT$)

$$\left|\int_{c-iT}^{d-iT} \frac{y^s}{s}\,ds\right| \le \int_c^d \frac{y^\sigma}{T}\,d\sigma \le \frac{1}{T}\int_c^\infty y^\sigma\,d\sigma = \frac{y^c}{|\log y|T}.$$

Exactly the same bound holds for $\int_{c+iT}^{d+iT}$. The last thing to think about is

$$\left|\frac{1}{2\pi i}\int_{d-iT}^{d+iT} \frac{y^s}{s}\,ds\right| \le Cy^d \int_{-T}^{T} \frac{dt}{1+|t|} \le Cy^d \log T.$$

Now, let $d \to \infty$. This term goes to zero, and we already had good estimates for the horizontal terms. So we have that if $0 < y < 1$ and $c > 0$ then

$$\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{y^s}{s}\,ds\right| \le \frac{y^c}{\pi T|\log y|}.$$

Taking the limit as $T \to \infty$ yields that

$$\lim_{T\to\infty} \frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{y^s}{s}\,ds = 0.$$

So not only have we proved this, we've even proved this in a more quantitative way.

Let's think a bit more about this $\log y$ term. It is natural to expect to get a term like this, because we know that there is a discontinuity at $y = 1$, and we'll run into problems there.

Now, we evaluate the integral for $y > 1$; we want to move the contour to the left. So we can write

$$\frac{1}{2\pi i}\int_{c-iT}^{c+iT} = \frac{1}{2\pi i}\left(\int_{c-iT}^{-d-iT} + \int_{-d-iT}^{-d+iT} + \int_{-d+iT}^{c+iT}\right) + 1,$$

where we have picked up a pole at $s = 0$. Now, we have the same type of argument as before, estimating each integral separately. We have

$$\left|\int_{-d-iT}^{-d+iT} \frac{y^s}{s}\, ds\right| \le Cy^{-d}\int_{-T}^{T} \frac{dt}{1+|t|} \le Cy^{-d}\log T \to 0$$

as $d \to \infty$. For the horizontal integrals, we have precisely the same bounds as before: they are bounded by

$$\le \int_{-d}^{c} \frac{y^\sigma}{T}\, d\sigma \le \frac{y^c}{T|\log y|}.$$

So therefore when $y > 1$ and $c > 0$, we have

$$\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{y^s}{s}\, ds - 1\right| \le \frac{y^c}{\pi T|\log y|},$$

so therefore

$$\lim_{T\to\infty} \frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{y^s}{s}\, ds = 1.$$

□

As $y \to 1$, our error terms currently blow up. We can show that our error terms are also bounded by a constant, which makes sense since when $y = 1$ we should get $\frac{1}{2}$. We can also deform contours in other ways, e.g. by taking a large semicircle.

$$\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{y^s}{s}\, ds\right| \le \frac{y^c}{|c+iT|}|c+iT| \le y^c.$$

So we've obtained the quantitative version of Perron's formula:

Proposition 9.5.

$$\frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{y^s}{s}\, ds = \begin{cases} 1 & y > 1 \\ 0 & y < 1 \end{cases} + O\left(y^c \min\left(1, \frac{1}{T|\log y|}\right)\right).$$

We can now use this to count something.

Example 9.6. We can write $\zeta(s) = \sum \frac{1}{n^s}$, so we have

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \zeta(s)\frac{x^s}{s}\, ds = \sum_{n=1}^{\infty} \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \left(\frac{x}{n}\right)^s \frac{ds}{s} = \sum_{n\le x} 1.$$

(note that here if $x \in \mathbb{Z}$ then the last value should be evaluated with weight 1/2). How do we compute this integral? We move the line of integration to the left and run into a pole at $s = 1$; here the residue is x. We'll get some bounds out, and eventually get $x + O(x^\theta)$ for some $\theta < 1$.

This example is sort of a joke, but we can apply this method to more interesting examples.

We want to apply Perron's formula to prove some asymptotics.

10.1. A toy example.

Example 10.1. Here's a toy example. We want to count $\sum_{n\le x} 1$ by considering

$$\frac{1}{2\pi i}\int_{c-iT}^{c+iT}\zeta(s)\frac{x^s}{s}\,ds=\sum_{n=1}^{\infty}\frac{1}{2\pi i}\int_{c-iT}^{c+iT}\left(\frac{x}{n}\right)^s\frac{ds}{s}$$

We need to choose c and T. We want T to be large, and we need $c>1$ in order for this to make sense. Applying the quantitative version of Perron's formula 9.5, we can write (where $\delta(y)=1$ if $y>1$ and $\delta(y)=0$ if $y<1$)

$$=\sum_{n=1}^{\infty}\delta\left(\frac{x}{n}\right)+O\left(\sum_{n=1}^{\infty}\left(\frac{x}{n}\right)^c\min\left(1,\frac{1}{T|\log\frac{x}{n}|}\right)\right).$$

On the other hand, we want to move the line of integration to the left, to $\int_{d-iT}^{d+iT}$ for some $0<d<1$. Here, there is a pole of $\zeta(s)$ at $s=1$, with residue 1. The remaining integrals will contribute an error term, so we need to bound the two horizontal integrals and the vertical integral. The choice of T will balance these error terms.

First, we assume that $x=\text{integer}+\frac{1}{2}$. This allows us to drop the minimum in our error term, leaving us

$$\sum_{n=1}^{\infty}\left(\frac{x}{n}\right)^c\frac{1}{T|\log\frac{x}{n}|}.$$

So we split into the following cases: (1) $0.9x\le n\le 1.1x$, and (2) $n<0.9x$ or $n>1.1x$. Let's deal with Case (2) first. Here, the error term is

$$\sum_{\substack{n<0.9x\\ n>1.1x}}\left(\frac{x}{n}\right)^c\frac{1}{T}\ll\frac{x^c}{T}\sum_{n=1}^{\infty}\frac{1}{n^c}=\frac{x^c\zeta(c)}{T}\ll\left(\frac{x^c}{c-1}\right)\frac{1}{T}\ll\frac{x\log x}{T}$$

where we have chosen $c=1+\frac{1}{\log x}$. Now, we do case (1). Write $n=[x]+k$, where $|k|\le 0.1x$. Then we are interested in

$$\left|\log\frac{x}{n}\right|=\left|\log\left(\frac{[x]+\frac{1}{2}}{[x]+k}\right)\right|=\left|\log\left(1+\frac{k-\frac{1}{2}}{[x]+k}\right)\right|\asymp\frac{|k-\frac{1}{2}|}{x}$$

For these terms, $(\frac{x}{n})^c$ is just a constant, and we have that the terms are

$$\ll\sum_{|k|\le 0.1x}\frac{x}{T|k-\frac{1}{2}|}\ll\frac{x\log x}{T}.$$

Now we bound the vertical and horizontal integrals. For both of these integrals, we need some estimate for the zeta function to the left of 1. Recall that we have $|\zeta(\sigma+it)|\ll(1+|t|)^{1-\sigma}\log(1+|t|)+1$ for $\sigma>0$.

For the vertical integral, we have

$$\int_{-T}^{T}|\zeta(d+it)|\frac{x^d}{|d+it|}\,dt\ll\int_{-T}^{T}\frac{(1+|t|)^{1-d}x^d}{1+|t|}(\log T)\,dt\ll x^dT^{1-d}\log T.$$

Now, consider the horizontal integrals:

$$\left|\int_d^c \zeta(\sigma+iT)\frac{x^{\sigma+iT}}{\sigma+iT}\,d\sigma\right| \ll \frac{1}{T}\int_d^c x^\sigma|\zeta(\sigma+iT)|\,d\sigma \ll \frac{1}{T}\int_d^c x^\sigma(1+T)^{1-\sigma}(\log T)\,d\sigma$$
$$\ll \frac{\log T}{T}(x^cT^{1-c}+x^dT^{1-d}) \ll \frac{x\log x}{T}+\frac{x^dT^{1-d}\log T}{T}.$$

Putting everything together, we have that

$$\sum_n \delta\left(\frac{x}{n}\right) = \sum_{n\le x} 1 + O\left(\frac{x\log x}{T}\right) = x + O\left(\frac{x\log x}{T} + x^dT^{1-d}\log T\right)$$

where the first equality comes from the quantitative form of Perron's formula and the second comes from shifting contours. Choose $d=\varepsilon$ to be as small as possible and pick $T=\sqrt{x}$ to get that

$$= x + O\left(\frac{x\log x}{T} + x^\varepsilon T^{1-\varepsilon}\log T\right) = x + O(x^{1/2+\varepsilon}).$$

This is a pretty bad estimate, but it demonstrates the general method to balance various error terms.

10.2. **k-divisor function.** Now let's use this method to prove a new result.

Example 10.2. Let $d_k(n)$ be the kth divisor function (where $k\in\mathbb{N}$); these satisfy

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

This is absolutely convergent for $\sigma>1$ because $d_k(n)\ll n^\varepsilon$ for any $\varepsilon>0$.

We can try to understand $\sum_{n\le x} d_k(n)$ by estimating

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} x^s\zeta(s)^k\frac{ds}{s}.$$

More precisely, we have

$$\frac{1}{2\pi i}\int_{c-iT}^{c+iT} x^s\zeta(s)^k\frac{ds}{s} = \sum_{n\le x} d_k(n) + O\left(\sum_n \left(\frac{x}{n}\right)^c \frac{d_k(n)}{T|\log\frac{x}{n}|}\right).$$

We want to move the line of integration to the left. The main term will be

$$\operatorname*{Res}_{s=1}\left(\zeta(s)^k\frac{x^s}{s}\right).$$

How should we compute this?

Example 10.3. First, think about the Laurent expansion in the case of $k=2$. This gives

$$\left(\frac{1}{s-1}+\gamma+\cdots\right)^2 \frac{x(1+(s-1)\log x+\frac{(s-1)^2(\log x)^2}{2}}{1+(s-1)}$$
$$= x\left(\frac{1}{(s-1)^2}+\frac{2\gamma}{s-1}+\cdots\right)(1+(s-1)\log x+\cdots)(1-(s-1)+\cdots).$$

Then the residue at $s = 1$ is $x \log x + 2\gamma x - x = x \log x + (2\gamma - 1)x$, which is the asymptotic that we proved for the sum of the divisor function.

Now, the Laurent expansion of $\zeta(s)^k \frac{x^s}{s}$ looks like

$$\left(\frac{1}{(s-1)^k} + \cdots\right) x \left(1 + (s-1)\log x + \cdots + \frac{((s-1)\log x)^{k-1}}{(k-1)!} + \cdots\right)(1 - (s-1) + \cdots),$$

so therefore

$$\operatorname*{Res}_{s=1}\left(\zeta(s)^k \frac{x^s}{s}\right) = xP_k(\log x)$$

for some polynomial P_k of degree k.

What is the error involved in doing this? We have

$$\sum_n \left(\frac{x}{n}\right)^c \frac{d_k(n)}{T|\log \frac{x}{n}|},$$

and as before, we break into two pieces. When $n < 0.9x$ or $n > 1.1x$, the error is

$$\ll \frac{x^c}{T}\sum \frac{d_k(n)}{n^c} \ll \frac{x^c \zeta(c)^k}{T} \ll \frac{x(\log x)^k}{T}$$

where we have chosen $c = 1 + \frac{1}{\log n}$. Now, in the other case, for $0.9x < n < 1.1x$, we have

$$\ll \frac{x^c}{T}\sum_{0.9x<n<1.1x} \frac{1}{|\log \frac{x}{n}|} \ll \frac{x^{1+\varepsilon}}{T}.$$

We still have to bound the error from the contour shifts. The horizontal integrals won't be too bad, and it should be almost the same as what we did before. Let's focus on the vertical integral:

$$|\text{vertical integral}| \ll x^d \int_{-T}^{T} \frac{|\zeta(d+iT)|^k}{1+|T|}\,dt \ll x^d \int_{-T}^{T} \frac{(1+|t|)^{k(1-d)+\varepsilon}}{1+|t|}\,dt \ll x^d T^{k(1-d)+\varepsilon}.$$

So the total error will be something of the form (choosing $d = \varepsilon$ to be small and $T = x^{1/(k+1)}$)

$$\frac{x^{1+\varepsilon}}{T} + x^d T^{k(1-d)+\varepsilon} = \frac{x^{1+\varepsilon}}{T} + x^\varepsilon T^k = x^{1-\frac{1}{k+1}+\varepsilon}.$$

11. 5/10

11.1. **Examples of asymptotics.** Last time, we found asymptotics for $\sum_{n\le x} d_k(n)$.

Example 11.1. Consider $\sum_{n\le x} d(n)^2$, which we know is of size $x(\log x)^3$. We have

$$F(s) = \sum_{n=1}^{\infty} \frac{d(n)^2}{x^s} = \prod_p \left(1 + \frac{d(p)^2}{p^s} + \frac{d(p^2)^2}{p^{2s}} + \cdots\right) = \prod_p \left(1 + \frac{4}{p^s} + \frac{9}{p^{2s}} + \cdots\right).$$

Recall that $\zeta(s) = \prod_p (1 + \frac{1}{p^s} + \cdots)$. Then $\zeta(s)^4 = \prod_p (1 + \frac{4}{p^s} + \cdots)$. So to kill off the lower order terms, we can write

$$F(s) = \zeta(s)^4 G(s) = \zeta(s)^4 \prod_p \left(1 + \frac{4}{p^s} + \frac{9}{p^{2s}} + \cdots\right)\left(1 - \frac{1}{p^s}\right)^4.$$

Where is $G(s)$ absolutely convergent? This product has no $\frac{1}{p^s}$ term, so we only have higher order terms. We have to be slightly careful; note that the coefficients of these terms only grow polynomially, so we don't have to worry about them. So we have

$$G(s) = \prod_p \left(1 + \sum_{k \geq 2} \frac{a(k)}{p^{ks}}\right)$$

where $a(k)$ grows at most polynomially in k. This converges absolutely if $\operatorname{Re} s > 1/2$.

Then carrying out the same argument that we did before, we will get that

$$\frac{1}{2\pi i} \int_{c-iT}^{c+iT} F(s) \frac{x^s}{s} \, ds = \sum_{n \leq x} d(n)^2 + O\left(\frac{x^{1+\varepsilon}}{T}\right)$$

where the error term comes from $\sum \frac{d(n)^2 (\frac{x}{n})^c}{T |\log \frac{x}{n}|}$ and we use $c = 1 + 1/\log x$. Now, we shift contours to some $d > 1/2$, and we pick up a residue:

$$\operatorname*{Res}_{s=1} \left(x^s \frac{\zeta(s)^4 G(s)}{s}\right).$$

We expect this to be x times a polynomial of degree 3 in $\log x$. We can expand in Taylor series to get $(s-1)^3 \frac{(\log x)^3}{3!} \frac{1}{(s-1)^4} \frac{G(1)}{1}$and this gives leading term $\frac{1}{3!} x (\log x)^3 G(1)$.

To bound the vertical line, we have (for $d \geq 0.51$, note that $|G(d+it)|$ is bounded)

$$x^d \int_{-T}^{T} \frac{|\zeta(d+it)|^4 |G(d+it)|}{1+|t} \, dt \ll x^d \int_{-T}^{T} \frac{(1+|t|)^{4(1-d)+\varepsilon}}{1+|t|} \, dt \ll x^d T^{4(1-d)+\varepsilon}.$$

We should also check the horizontal integrals, which turn out to be negligible.

What should we choose d and T to be? We want d as close to $1/2$ as possible, so choose $d = 1/2 + \varepsilon$. Then we have a bound of $\ll x^{1/2+\varepsilon} T^{2+\varepsilon}$. Choose $T = x^{1/6}$ to see that

$$\sum_{n \leq x} d(n)^2 = \operatorname*{Res}_{s=1} \left(x^s \frac{\zeta(s)^4 G(s)}{s}\right) + O\left(x^{5/6+\varepsilon}\right).$$

Example 11.2. What if we wanted to compute $\sum_{n \leq x} d_k(n)^2$? What power of log is this? We write down an Euler product

$$\prod_p \left(1 + \frac{k^2}{p^s} + \cdots\right).$$

This means that we compare to $\zeta(s)^{k^2}$, which means that we expect to see a $(\log x)^{k^2-1}$ term.

Example 11.3. Life is not always so simple. Suppose we wanted to consider $\sum_{n \leq x} d_\pi(x)$. This is some multiplicative function, with $d_\pi(p) = \pi$, for example. In this case, we look at

$$\sum_{n=1}^{\infty} \frac{d_\pi(n)}{n^s} = \zeta(s)^\pi = \exp(\pi \log \zeta(s)).$$

We no longer have a polar singularity at $s = 1$, and this becomes a complicated object. At the moment, we only know that this is defined to the right of 1. If we try to shift contours, we pass through a singularity at $s = 1$, which is a logarithmic singularity which has to be treated carefully. This can be done, and in fact,

$$\sum_{n\le x} d_\pi(n) \sim \frac{x(\log x)^{\pi-1}}{\Gamma(\pi)}.$$

Example 11.4. Let $a(n)$ be the number of abelian groups of order n. This is a multiplicative function, so we only have to observe that $a(p^k) = p(k)$ is the number of partitions of k. Then we can write

$$\sum_{n=1}^\infty \frac{a(n)}{n^s} = \prod_p \left(\sum_{k=0}^\infty \frac{p(k)}{p^{ks}}\right).$$

To do this, set $z = p^{-s}$ and look at

$$\sum_{n=0}^\infty p(n)z^n = \prod_{j=1}^\infty (1 + z^j + z^{2j} + \cdots) = \prod_{j=1}^\infty \frac{1}{1-z^j}.$$

This product in fact makes sense for $|z| < 1$. Using this, the Euler product becomes (assuming $\operatorname{Re} s > 1$)

$$\prod_p \left(\sum_{k=0}^\infty \frac{p(k)}{p^{ks}}\right) = \prod_p \prod_{j=1}^\infty \frac{1}{1-\frac{1}{p^{js}}} = \prod_{j=1}^\infty \zeta(js).$$

With a bit of thought, this converges for $\operatorname{Re} s > 1$. We want to extend this to a larger domain. In $\operatorname{Re} s > 1/2$, this is analytic except for a simple pole at $s = 1$. If $\operatorname{Re} s > 1/3$, we have poles at $s = 1/2$ and $s = 1$. So we actually have a meromorphic continuation to $\operatorname{Re} s > 0$, and in this region we have poles at $1, \frac{1}{2}, \frac{1}{3}, \ldots$.

We can evaluate

$$\sum_{n\le x} a(n) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \prod_{j=1}^\infty \zeta(js) x^s \frac{ds}{s} \sim x\zeta(2)\zeta(3)\zeta(4)\cdots.$$

11.2. **Return to prime numbers.** Recall that we want to study $\psi(x) = \sum_{n\le x}\Lambda(n)$.

In $\operatorname{Re} s > 1$, we have

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

$$\log\zeta(s) = \sum_p \log\left(1-\frac{1}{p^s}\right)^{-1} = \sum_p \sum_{k=1}^\infty \frac{1}{kp^{ks}}.$$

Then

$$\frac{\zeta'}{\zeta}(s) = \sum_p \sum_{k=1}^\infty \frac{-k\log p}{kp^{ks}} = -\sum_p\sum_k \frac{\Lambda(p^k)}{p^{ks}} = -\sum_n \frac{\Lambda(n)}{n^s}.$$

Now we have that (for $c > 1$)

$$\psi(x) = \sum_{n\le x}\Lambda(n) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\left(-\frac{\zeta'}{\zeta}(s)\right)\frac{x^s}{s}\,ds.$$

We proceed somewhat nonrigorously at the moment; we'll do a rigorous proof later. We now move the line of integration to the left and compute residues. The singularities of the integrand are at $s = 1$ (a pole of $\zeta(s)$), $s = \rho$ for zeros of $\zeta(s)$, and $s = 0$. We compute residues at these poles.

At $s = 1$, we have $\zeta(s) \sim \frac{1}{s-1}$, and hence $-\frac{\zeta'}{\zeta}(s) \sim \frac{1}{s-1}$. So then

$$\operatorname*{Res}_{s=1}\left(\frac{x^s}{s}\left(-\frac{\zeta'}{\zeta}(s)\right)\right) = x.$$

This is the content of the prime number theorem: the main term comes from the pole at $s = 1$.

Let's compute the other residues. At ρ, we have $\zeta(s) \sim c(s-\rho)^{m(\rho)}$, so then $-\frac{\zeta'}{\zeta}(s) \sim \frac{m(\rho)}{s-\rho}$.

$$\operatorname*{Res}_{s=\rho}\left(\frac{x^s}{s}\left(-\frac{\zeta'}{\zeta}(s)\right)\right) = -\frac{x^\rho}{\rho}.$$

Here, we count all of the zeros with multiplicity. Finally, the residue at 0 gives $-\frac{\zeta'}{\zeta}(0)$.

This gives us the *explicit formula*:

$$\psi(x) = x - \sum_\rho \frac{x^\rho}{\rho} - \frac{\zeta'}{\zeta}(0).$$

We have to be careful here: does the sum converge? In fact, it doesn't converge absolutely, but we can make sense of this. This sort of formula is morally true.

At the moment, we don't know much about the zeros of $\zeta(s)$. There are the trivial zeros of $\zeta(s)$ at $-2, -4, -6, \cdots$. We usually ignore these; they contribute $-\sum_{n=1}^\infty \frac{x^{-2n}}{-2n} = \frac{1}{2}\log\frac{1}{1-x^{-2}}$. We are left with the nontrivial zeros. We will later see that these zeros $\rho = \beta + i\gamma$ all satisfy $0 \le \beta \le 1$. We claim that $\zeta(\overline{s}) = \overline{\zeta(s)}$. This can be seen from the Schwarz reflection principle. So this means that if $\beta + i\gamma$ is a zero, then $\beta - i\gamma$ is too. In addition, there is a relation connecting $\zeta(s)$ to $\zeta(1-s)$, called the *functional equation*:

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \pi^{-(1-s)/2}\Gamma((1-s)/2)\zeta(1-s).$$

So the zeros of $\zeta(s)$ are symmetric about both the real axis and $\operatorname{Re} s = 1/2$. The Riemann hypothesis states that all nontrivial zeros satisfy $\beta = 1/2$.

The Riemann hypothesis is nice because looking at the explicit formula, the sum over zeros will be as small as it can be, and hence the error in the prime number theorem is small.

12. 5/15

Recall from last time that we are studying the asymptotics of $\psi(x)$ through Perron's formula and shifting contours. We hope that for large x, the contribution from the zeros is $o(x)$. The explicit formula can be made precise, i.e.

$$\psi(x) = x - \sum_{|\rho| \le T} \frac{x^\rho}{\rho} + O\left(\frac{x^{1+\varepsilon}}{T} + \log x\right).$$

Theorem 12.1. *If $\beta + i\gamma$ is a zero of $\zeta(s)$ then $\beta < 1$ (i.e. $\zeta(1+it) \neq 0$).*

Proof (Hadamard and de la Vallee-Poussin). Look at $\sigma > 1$; then we claim that

$$\zeta(\sigma)^3|\zeta(\sigma+it)|^4|\zeta(\sigma+2it)| \ge 1$$

for all $t \in \mathbb{R}$. To show this, take logs; we wish to show that

$$3 \log \zeta(\sigma) + 4 \operatorname{Re} \log \zeta(\sigma + it) + \operatorname{Re} \log \zeta(\sigma + 2it) \geq 0.$$

Now, consider $\log \zeta(s) = \sum_{p^k} \frac{1}{kp^{ks}}$. The quantity that we want is then

$$\sum_{p^k} \frac{1}{k} \left(\frac{3}{p^{k\sigma}} + 4 \operatorname{Re} \frac{1}{p^{k(\sigma+it)}} + \operatorname{Re} \frac{1}{p^{k(\sigma+2it}} \right) = \sum_{p,k} \frac{1}{kp^{k\sigma}} \left(3 + 4\cos(kt \log p) + \cos(2tk \log p)\right).$$

Now, the heart of the proof is a trigonometric identity:

$$3 + 4\cos\theta + \cos(2\theta) = 3 + 4\cos\theta + 2\cos^2\theta - 1 = 2(1 + \cos\theta)^2 \geq 0.$$

This implies the result that we needed.

Now we can complete the proof that $\zeta(1+it) \neq 0$. Firstly, observe that $\zeta(s) = \frac{1}{s-1} + \gamma + \cdots$, so in a neighborhood of 1, then ζ is not zero. So we can assume that $|t| \geq \delta$.

Suppose then that $\zeta(1+it) = 0$. Then for σ close to 1, we have

$$\zeta(\sigma + it) = \zeta(\sigma - 1 + 1 + it) = C(\sigma - 1) + \text{higher order terms},$$

so hence $|\zeta(\sigma + it)| \leq C(\sigma - 1)$ for some constant $C = C(t)$. Also, $|\zeta(\sigma + 2it)| \leq \log(2 + |t|)$. We also know that $\zeta(\sigma) = \frac{1}{\sigma - 1} + \gamma + \cdots$. So putting these estimates together, we see that

$$1 \leq \zeta(\sigma)^3 |\zeta(\sigma + it)|^4 |\zeta(\sigma + 2it)| \leq \left(\frac{1}{\sigma - 1} \right)^3 (\sigma - 1)^4 C(t),$$

so $C(t)(\sigma - 1) \geq 1$ for all $\sigma > 1$, which is false if σ is sufficiently close to 1. ☐

Remark. Why did we look at this inequality? We have three poles at $\sigma = 1$ and four zeros, and then some regular holomorphic thing. So there are more zeros than poles, so we cannot have $\zeta(\sigma)^3 |\zeta(\sigma + it)|^4 |\zeta(\sigma + 2it)| \geq 1$.

What's a good heuristic for this? We have

$$|\zeta(1 + it)| = \prod_p \left| 1 - \frac{1}{p^{1+it}} \right|^{-1} = \left| 1 - \frac{p^{-it}}{p} \right|^{-1}$$

$|\zeta(1+it)|$ is large when $p^{it} \approx 1$ for many p, and $|\zeta(1+it)|$ is small when $p^{it} \approx -1$ for many p. When $\zeta(1+it) = 0$, then $p^{it} \approx -1$ for many primes p. This means that $p^{2it} \approx 1$, so $\zeta(1+2it)$ will be large. But we know that $\zeta(s)$ cannot have a pole there. This is the heuristic that suggests the inequality that we considered above.

So now we know that there are no zeros of $\zeta(s)$ on $\operatorname{Re} s = 1$. Let's try to complete the proof of the prime number theorem.

Instead of looking at $\psi(x)$, we will look at a slightly smoother thing:

$$\psi_1(x) = \sum_{n \leq x} \Lambda(n)(x - n) = \int_0^x \psi(t)\, dt.$$

We hope to show that $\psi_1(x) \sim \frac{x^2}{2}$.

Proposition 12.2. $\psi_1(x) \sim \frac{x^2}{2}$ *implies* $\psi(x) \sim x$.

Proof. For $h \le x$, consider

$$h\psi(x+h) \ge \psi_1(x+h) - \psi_1(x) = \int_x^{x+h} \psi(t)\,dt \ge h\psi(x).$$

By hypothesis,

$$\psi_1(x+h) - \psi_1(x) = \frac{(x+h)^2}{2} + o((x+h)^2) - \frac{x^2}{2} + o(x^2) = hx + \frac{h^2}{2} + o(x^2).$$

From this, we see that

$$\psi(x) \le x + \frac{h}{2} + o\left(\frac{x^2}{h}\right) \le x + \frac{h}{2} + \frac{\varepsilon x^2}{h}$$

Choosing $h = \sqrt{\varepsilon}x$ gives then that this error term is $\frac{h}{2} + o(\frac{x^2}{h}) \le \sqrt{\varepsilon}x$. □

Theorem 12.3. $\psi_1(x) \sim \frac{x^2}{2}$.

Proof. For $c > 1$, write

$$\psi_1(x) = \left(\frac{1}{2\pi i}\int_{(c)} -\frac{\zeta'}{\zeta}(s)x^s\frac{ds}{s(s+1)}\right)x.$$

This is true because it is

$$\sum_n \Lambda(n)\left(\frac{1}{2\pi i}\int_{(c)}\left(\frac{x}{n}\right)^s\frac{ds}{s(s+1)}\right)x = \sum_{n\le x}\Lambda(n)\left(1-\frac{n}{x}\right).$$

Now, we want to show that

$$\frac{1}{2\pi i}\int_{(c)} -\frac{\zeta'}{\zeta}(s)x^s\frac{ds}{s(s+1)} \sim \frac{x}{2}.$$

To see that we're on the right track, we expect the main contribution to be from the residue at 1, which is precisely $\frac{x}{2}$.

We shift the contour to be along

$$1 - i\infty \to 1 - iT \to 1 - \delta(T) - iT \to 1 - \delta(T) + iT \to 1 + iT \to 1 + i\infty$$

where T is some large parameter. This allows us to shift contours while only passing through the pole at $s = 1$. To do this, find $\delta(T)$ such that $\zeta(s) \ne 0$ in the region $\operatorname{Re} s > 1 - \delta(T)$ and $|\operatorname{Im}(s)| \le T$. This can be done because there exists a neighborhood of $[1 - iT, 1 + iT]$ that is free of the zeros of $\zeta(s)$.

So now our integral becomes

$$\frac{1}{2\pi i}\int_{(c)} -\frac{\zeta'}{\zeta}(s)x^s\frac{ds}{s(s+1)} = \frac{1}{2\pi i}\int_{\text{contour}} -\frac{\zeta'}{\zeta}(s)x^s\frac{ds}{s(s+1)} + \frac{x}{2}$$

where we've picked up the residue at $s = 1$. So it remains to bound this contour integral, which we can do in five pieces.

First, consider the middle vertical piece:

$$\left|\int_{1-\delta(T)-iT}^{1-\delta(T)+iT} -\frac{\zeta'}{\zeta}(s)x^s\frac{ds}{s(s+1)}\right| \le x^{1-\delta(T)}\left|\int_{1-\delta(T)-iT}^{1-\delta(T)+iT}\frac{\left|-\frac{\zeta'}{\zeta}(s)\right|}{|s(s+1)|}|ds|\right| \le F(T)x^{1-\delta(T)}$$

because the remaining integral depends only on T and not on x. For example, we could take $F(T) = \max_{[1-\delta(T)-iT, 1-\delta(T)+iT]} |\frac{\zeta'}{\zeta}(s)|$.

Now, consider the horizontal pieces:

$$\left|\int_{1-\delta(T)+iT}^{1+iT} -\frac{\zeta'}{\zeta}(s)x^s \frac{ds}{s(s+1)}\right| \leq F(T)\int_{1-\delta(T)}^{1} x^\sigma \, d\sigma \leq F(T)\frac{x}{\log x}.$$

The last piece to consider are the remaining vertical integrals, where we have to be more careful. Consider

$$\left|\int_{1+iT}^{1+i\infty} -\frac{\zeta'}{\zeta}(s)\frac{x^s}{s(s+1)}\, ds\right| \ll x\int_T^\infty \frac{\left|\frac{\zeta'}{\zeta}(1+it)\right|}{1+t^2}\, dt$$

This is where we used the extra smoothness; otherwise working with Perron's formula would be hopeless. At this point, we want an upper bound for

$$\left|\frac{\zeta'}{\zeta}(1+it)\right| = \frac{|\zeta'(1+it)|}{|\zeta(1+it)|}.$$

We already know that $|\zeta'(1+it)| = O((\log(1+|t|))^2)$. We still need some kind of lower bound for $|\zeta(1+it)|$. We'll do this later, but for now, let's assume that $|\zeta(1+it)| \gg (\log(2+|t|))^{-10}$. Making this assumption, we have that

$$\left|\frac{\zeta'}{\zeta}(1+it)\right| = \frac{|\zeta'(1+it)|}{|\zeta(1+it)|} \ll (\log(2+|t|))^{12}.$$

This would mean that we have

$$\left|\int_{1+iT}^{1+i\infty} -\frac{\zeta'}{\zeta}(s)\frac{x^s}{s(s+1)}\, ds\right| \ll x\int_T^\infty \frac{\left|\frac{\zeta'}{\zeta}(1+it)\right|}{1+t^2}\, dt \ll x\int_T^\infty \frac{(\log t)^{12}}{t^2}\, dt \ll x\frac{(\log T)^{12}}{T}.$$

Now, we've proved that

$$\frac{1}{2\pi i}\int_{(c)} -\frac{\zeta'}{\zeta}(s)\frac{x^s}{s(s+1)}\, ds = \frac{x}{2} + O\left(x^{1-\delta(T)}F(T) + \frac{x}{\log x}F(T) + x\frac{(\log T)^{12}}{T}\right).$$

We want to make the error term $\leq \varepsilon x$ for large x. To do this, first choose T such that $\frac{(\log T)^{12}}{T} < \frac{\varepsilon}{100}$. For such a choice of T, choose x so large that $(\frac{x}{\log x} + x^{1-\delta(T)})F(T) \leq \frac{\varepsilon x}{100}$. That completes our proof of the Prime Number Theorem. To estimate the error term more carefully, we need to analyze $F(T)$, which is something that can be done. □

13. 5/17

13.1. **Finishing the proof of the Prime Number Theorem: Lower bounds for $|\zeta(1+it)|$.**

Recall that we are proving the prime number theorem. Last time, we completed the proof, but we assumed that

$$\left|\frac{\zeta'}{\zeta}(1+it)\right| \ll (\log(2+|t|))^{12}$$

in the range $|t| \geq 1$. This is what we will now consider.

Recall from the homework that we know that $|\zeta'(1+it)| \ll (\log|t|+1)^2$. Finally, we need a lower bound for $|\zeta(1+it)| \gg (\log|t|)^{-10}$. Note that $|\zeta(1+it)|$ is not bounded below, and in fact, $\liminf |\zeta(1+it)| = 0$.

Last time, we showed that

$$\zeta(\sigma)^3|\zeta(\sigma+it)|^4|\zeta(\sigma+2it)| \geq 1,$$

which gives us a lower bound

$$\begin{aligned}|\zeta(\sigma+it)| \geq |\zeta(\sigma+2it)|^{-1/4}\zeta(\sigma)^{-3/4} &\geq (C\log(1+|t|))^{-1/4}\left(\frac{1}{\sigma-1}\right)^{-3/4}\\ &\gg (\sigma-1)^{3/4}(\log(1+|t|))^{-1/4}.\end{aligned}$$

This is already fairly good; for example, this is $\geq (\log(1+|t|)^{-9})$ so long as $\sigma \geq 1+\frac{1}{(\log(1+|t|))^{10}}$. We can also write

$$\zeta\left(1+\frac{1}{(\log|t|)^{10}}+it\right) \geq \frac{C}{(\log|t|)^{7.75}}.$$

Then

$$\zeta(1+it) = \zeta\left(1+\frac{1}{(\log t)^{10}}+it\right) - \int_0^{1/(\log t)^{10}} \zeta'(1+\lambda+it)\,d\lambda,$$

and hence

$$\begin{aligned}|\zeta(1+it)| &\geq \left|\zeta\left(1+\frac{1}{(\log t)^{10}}+it\right)\right| - \int_0^{1/(\log t)^{10}} |\zeta'(1+\lambda+it)|\,d\lambda\\ &\geq \frac{C}{(\log|t|)^{7.75}} - \frac{C}{(\log|t|)^{10}}(\log|t|)^2 \gg \frac{1}{(\log|t|)^{7.75}}.\end{aligned}$$

The point of the proof is that we are still making use of the lemma that we proved last time, where we said that four is more than three. At this stage, we have a complete proof of the prime number theorem.

13.2. **Making the error term quantitative.** It is true that $\zeta(\beta+it) \neq 0$ if $\beta \geq 1-\frac{c}{\log|t|+2}$. Moreover in this region we have

$$\left|\frac{\zeta'}{\zeta}(\beta+it)\right| = O((\log|t|+2)^{12}).$$

Knowing these facts, we can choose $\delta(T) = \frac{C}{\log T}$ in the proof of the prime number theorem. We can then estimate our integrals more carefully. Then

$$\begin{aligned}\left|\int_{1-\delta(T)-iT}^{1-\delta(T)+iT} \frac{x^s}{s(s+1)}\left(-\frac{\zeta'}{\zeta}(s)\right)ds\right| &\ll x^{1-\delta(T)}\\ \left|\int_{1-\delta(T)+iT}^{1+iT} \frac{x^s}{s(s+1)}\left(-\frac{\zeta'}{\zeta}(s)\right)ds\right| &\ll \frac{x}{T^2}(\log T)^{12}\\ \left|\int_{1+iT}^{1+i\infty} \frac{x^s}{s(s+1)}\left(-\frac{\zeta'}{\zeta}(s)\right)ds\right| &\ll x\frac{(\log T)^{12}}{T}.\end{aligned}$$

So the error is of size

$$x^{1-\frac{C}{\log T}} + \frac{x}{T}(\log T)^{12}.$$

Choose $T = x^{\frac{C}{\log T}}$ and $\log T = \frac{C\log x}{\log T}$, so that $T = \exp(\sqrt{\log x})$. Then the error is $O(xe^{-c\sqrt{\log x}})$. Hence we have the following result:

Theorem 13.1.

$$\begin{aligned}\psi_1(x) &= \frac{x^2}{2} + O\left(x^2 \exp(-c\sqrt{\log x})\right)\\ \implies \psi(x) &= x + O\left(x \exp(-c\sqrt{\log x})\right)\\ \implies \pi(x) &= \mathrm{li}(x) + O\left(x \exp(-c\sqrt{\log x})\right).\end{aligned}$$

The best known result has an error term something like $O(x\exp(-c(\log x)^{3/5}))$, and it is an outstanding problem to show an error term of $O(x^{1-\delta})$, for any $\delta > 0$. The Riemann Hypothesis says that the error should be $O(x^{1/2})$. Proving better error terms is therefore related to progress on the Riemann Hypothesis by expanding the zero-free region for $\zeta(s)$.

Question. Does there exist a prime p in $[n^2, (n+1)^2]$?

We don't know how to do this, even assuming the Riemann Hypothesis. Here, we have

$$\psi(x+h) - \psi(x) = h + O(E(x))$$

for some error term. If the Riemann Hypothesis is true, then this is $h + O(x^{1/2+\varepsilon})$, so hence $[x, x + x^{1/2+\varepsilon}]$ contains a prime. But this isn't good enough; $[n^2, (n+1)^2]$ looks like $[x, x + 2\sqrt{x}]$. But we can get a related result.

Theorem 13.2. *There exists a prime in* $[n^3, (n+1)^3]$.

This uses the explicit formula and more information on the zeros of $\zeta(s)$.

13.3. **Functional equation for $\zeta(s)$.** Earlier, we mentioned the symmetry of the zeta function about the line $\operatorname{Re} s = \frac{1}{2}$:

Theorem 13.3 (Functional Equation for $\zeta(s)$)**.**

$$s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s) = s(s-1)\pi^{-(1-s)/2}\Gamma((1-s)/2)\zeta(1-s).$$

Let $\xi(s) = s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s)$*; then* $\xi(s) = \xi(1-s)$ *and* $\xi(s)$ *is analytic in* $\mathbb{C}$.

The fact that $\xi(s)$ is analytic in $\mathbb{C}$ should be clear from the functional equation. When $\operatorname{Re} s > 0$, $\zeta(s)$ has the only pole at $s = 1$, and this is canceled by $(s-1)$. Also, Γ has no poles here, so $\xi(s)$ is analytic. But $\xi(s) = \xi(1-s)$, so $\xi(s)$ is also analytic when $\operatorname{Re} s < 1$. These two regions overlap, so therefore ξ is analytic in $\mathbb{C}$.

We also know that $\Gamma(s/2)$ has poles at $0, -2, -4, -6, \dots$. The pole at 0 is canceled by the term s; the other poles must all be canceled by the zeros of $\zeta(s)$. That's the reason why $\zeta(-2) = \zeta(-4) = \cdots = 0$. These are the trivial zeros of the zeta function.

Recall that we also know that $\zeta(-n) \in \mathbb{Q}$ for $n = 1, 2, 3, \dots$. The functional equation gives $\zeta(1+n) = \zeta(-n)$, which produces an evaluation of $\zeta(2), \zeta(4), \dots$ as some rational multiples of powers of π. So the functional equation also encodes Euler's evaluation of these values.

The proof of the functional equation for $\zeta(s)$ requires the Poisson Summation Formula, which we will recall here.

Suppose that $f : \mathbb{R} \to \mathbb{R}$ is smooth and rapidly decreasing at $|x| \to \infty$, e.g. $f \in \mathcal{S}(\mathbb{R})$ is a Schwartz class function. Then we can define the *Fourier transform*:

Definition 13.4.

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \xi}\, dx.$$

Here, if $f \in \mathcal{S}(\mathbb{R})$ then $\hat{f} \in \mathcal{S}(\mathbb{R})$ as well.

Theorem 13.5 (Poisson Summation Formula)**.**

$$\sum_{n \in \mathbb{Z}} f(n) = \sum_{k \in \mathbb{Z}} \hat{f}(k).$$

Proof. Think of $F(x) = \sum_{n \in \mathbb{Z}} f(x+n)$. This is a nice and convergent sum, so the sum is absolutely convergent. Also, note that

$$F(x+1) = \sum_{n \in \mathbb{Z}} f(x+1+n) = \sum_{n \in \mathbb{Z}} f(x+n) = F(x).$$

So $F(x)$ is a nice smooth function with period 1; $F : \mathbb{R}/\mathbb{Z} \to \mathbb{R}$. Now we can write down the Fourier series for F. The Fourier coefficients are

$$\hat{F}(n) = \int_0^1 F(x) e^{-2\pi i n x}\, dx$$

for $n \in \mathbb{Z}$, so therefore the Fourier series is

$$\sum_{k \in \mathbb{Z}} \hat{F}(k) e^{2\pi i k x} = F(x).$$

Now

$$\begin{aligned}\hat{F}(k) &= \int_0^1 \left(\sum_{n \in \mathbb{Z}} f(x+n) \right) e^{-2\pi i k (x+n)}\, dx = \sum_{n \in \mathbb{Z}} \int_0^1 f(x+n) e^{-2\pi i k (x+n)}\, dx \\ &= \sum_{n \in \mathbb{Z}} \int_n^{n+1} f(y) e^{-2\pi i k y}\, dy = \int_{-\infty}^{\infty} f(y) e^{-2\pi i k y}\, dy = \hat{f}(k).\end{aligned}$$

So now

$$F(0) = \sum_{n \in \mathbb{Z}} f(n) = \sum_{k \in \mathbb{Z}} \hat{F}(k) e^{2\pi i k 0} = \sum_{k \in \mathbb{Z}} \hat{f}(k).$$

□

A function and its Fourier transform has some sort of inverse relationship; a function with a peak will have a Fourier transform that is spread out, and conversely. In some sense, we can think of $\sum f(n)$ as some sort of Riemann sum; $\hat{f}(0) = \int f(x)\, dx$, so $\sum \hat{f}(k)$ is like saying that the integral is equal to the Riemann sum, with some error term.

Example 13.6. Consider the Gaussian $f(x) = e^{-\pi x^2}$. Then the Fourier transform is (via completing the square)

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} e^{-\pi x^2 - 2\pi i x \xi}\, dx = \int_{-\infty}^{\infty} e^{-\pi (x + i\xi)^2 - \pi \xi^2}\, dx = e^{-\pi \xi^2} \int_{-\infty + i\xi}^{\infty + i\xi} e^{-\pi z^2}\, dz.$$

Since the integrand is analytic, we can move this contour and see that

$$\hat{f}(\xi) = e^{-\pi \xi^2} \left(\int_{-\infty}^{\infty} e^{-\pi x^2}\, dx \right) = e^{-\pi \xi^2}.$$

Here, we either remember that this final integral is 1, or apply the Poisson Summation Formula:

$$\sum_{n\in\mathbb{Z}} e^{-\pi n^2} = C\sum_{k\in\mathbb{Z}} e^{-\pi k^2}$$

implies that $C = 1$.

14. 5/22

We finish the proof of the functional equation for $\zeta(s)$:

14.1. **Proof of functional equation 13.3.** Last time, we computed the Fourier transform of a Gaussian. Through substituting $y = \sqrt{t}x$, we have that $f_t(x) = e^{-\pi x^2 t}$ has Fourier transform $\frac{1}{\sqrt{t}}e^{-\pi x^2/t}$.

Proposition 14.1.

$$\theta(t) = \sum_{n\in\mathbb{Z}} e^{-\pi n^2 t} = \frac{1}{\sqrt{t}}\theta\left(\frac{1}{t}\right).$$

Proof.

$$\sum_{n\in\mathbb{Z}} f_t(n) = \sum_{k\in\mathbb{Z}} \frac{1}{\sqrt{t}}e^{-\pi k^2/t} = \frac{1}{\sqrt{t}}\theta\left(\frac{1}{t}\right).$$

□

This should make sense. For large t, we have $\theta(t) \approx 1$. For small t, we can think of replacing the sum by an integral, and we get $\int e^{-\pi x^2 t}\,dx = \frac{1}{\sqrt{t}}$.

This implies the functional equation for $\zeta(s)$. Recall that we are interested in the quantity $\pi^{-s/2}\Gamma(s/2)\zeta(s)$. Start with the situation $\operatorname{Re} s > 1$. Then we have

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \pi^{-s/2}\Gamma(s/2)\sum_{n=1}^{\infty}\frac{1}{n^s} = \sum_{n=1}^{\infty}\int_0^{\infty} e^{-y}\left(\frac{y}{\pi n^2}\right)^{s/2}\frac{dy}{y}$$

and substituting $y = \pi n^2 z$ yields

$$= \sum_{n=1}^{\infty}\int_0^{\infty} e^{-\pi n^2 z} z^{s/2}\frac{dz}{z} = \int_0^{\infty}\left(\sum_{n=1}^{\infty} e^{-\pi n^2 z}\right) z^{s/2}\frac{dz}{z}.$$

Note that $\omega(z) = \sum_{n=1}^{\infty} e^{-\pi n^2 z} = \frac{\theta(z)-1}{2}$, so therefore

$$= \int_0^{\infty}\frac{\theta(z)-1}{2}z^{s/2}\frac{dz}{z}.$$

We want an expression for this that makes sense for all $s \neq 1$. But for now, the integral makes sense for z large and for any s. We have to worry about the integral near 0. So we have to look at

$$= \int_0^{1}\frac{\theta(z)-1}{2}z^{s/2}\frac{dz}{z} + \int_1^{\infty}\frac{\theta(z)-1}{2}z^{s/2}\frac{dz}{z}.$$

For the first integral, we use the relation $\theta(z) = \frac{1}{\sqrt{z}}\theta(\frac{1}{z})$. So then

$$\begin{aligned}
\int_0^1 \frac{\theta(z)-1}{2} z^{s/2}\frac{dz}{z} &= \int_0^1 \frac{\frac{1}{\sqrt{z}}\theta(\frac{1}{z})-1}{2} z^{s/2}\frac{dz}{z} \\
&= \int_0^1 \frac{\frac{1}{\sqrt{z}}(\theta(\frac{1}{z})-1)}{2} z^{s/2}\frac{dz}{s} + \int_0^1 \frac{1}{2\sqrt{z}} z^{s/2}\frac{dz}{z} - \int_0^1 \frac{z^{s/2}\,dz}{2z} \\
&= \frac{1}{2}\int_0^1 \left(\theta\left(\frac{1}{z}\right)-1\right) z^{\frac{s-1}{2}}\frac{dz}{z} + \frac{1}{s(s-1)}
\end{aligned}$$

and applying the substitution $y = \frac{1}{z}$ gives

$$= \frac{1}{s(s-1)} + \int_1^\infty \frac{\theta(y)-1}{2} y^{\frac{1-s}{2}}\frac{dy}{y}.$$

Adding our two integrals together, we have

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \frac{1}{s(s-1)} + \int_1^\infty \left(\frac{\theta(y)-1}{2}\right)\left(y^{\frac{s}{2}} + y^{\frac{1-s}{2}}\right)\frac{dy}{y}.$$

This is holomorphic for all s. Replacing s by $1-s$, we notice that the right hand side does not change, so therefore this is symmetric under $s \leftrightarrow 1-s$, which proves our functional equation.

14.2. **Further ideas about $\zeta(s)$.** This was one of Riemann's original proofs, and it also gives another proof for the analytic continuation of $\zeta(s)$. This also shows that if ρ is a zero, then $1-\rho$ is also a zero.

Now, from this, we can also see that

$$\xi(s) = s(s-1)\Gamma(s/2)\pi^{-s/2}\zeta(s)$$

has moderate growth properties: $|\xi(s)| \le \exp(C|s|\log|s|)$ if $|s|$ is large.

Functions of the type $F(s) = O(\exp(|s|^A))$ are called functions of finite order. So here, $\xi(s)$ is an entire function of order 1.

From this, we can also say one more thing about the zeta function. We use a fact from complex analysis, called Hadamard's factorization formula. Then

$$\xi(s) = e^{Bs}\prod_\rho \left(1-\frac{s}{\rho}\right)e^{s/\rho}$$

where the product is taken over all nontrivial zeros of $\zeta(s)$. Using this type of fact, we see that

$$\#\{\rho = \beta + i\gamma : 0 < \beta < 1, 0 < \gamma \le T\} = \frac{T}{2\pi}\log\frac{T}{2\pi e} + O(\log T).$$

The Riemann Hypothesis has been checked for trillions of zeros already. How can we tell if a zero has $\operatorname{Re} s = \frac{1}{2}$ instead of just being very close? We can check that

$$\xi\left(\frac{1}{2}+it\right) = \xi\left(\frac{1}{2}-it\right) = \overline{\xi\left(\frac{1}{2}+it\right)},$$

so therefore, if $t \in \mathbb{R}$ then $\xi(\frac{1}{2}+it) \in \mathbb{R}$ as well. Then we can count sign changes of $\xi(s)$ while varying s from $\frac{1}{2}$ to $\frac{1}{2}+it$. By counting sign changes, we get a guaranteed number of zeros on the $\operatorname{Re} s = \frac{1}{2}$ line. We can also estimate precisely how many zeros we should have, by further estimating the $O(\log T)$ error above; then, we can just see if we get the right number of sign changes.

Riemann computed 3 or 4 zeros in this way, and Turing computed around 500 zeros using one of the first computer calculations.

If we get a double zero, then this technique for verifying the Riemann Hypothesis would fail; however, all zeros found so far have been simple zeros, and it is conjectured that all zeros are simple zeros.

14.3. **Dirichlet's theorem on primes in arithmetic progressions.** This is all a special case of a more general class of functions called L-functions. These come up in the proof of Dirichlet's theorem on primes in arithmetic progressions. There, we have $(a, q) = 1$, and the theorem says that

$$\psi(x; q, a) = \sum_{\substack{n \le x \\ n \equiv a \bmod q}} \Lambda(n) \sim \frac{x}{\psi(q)}$$

as $x \to \infty$. To do this, Dirichlet introduced Dirichlet characters, which are some functions $\chi : (\mathbb{Z}/q\mathbb{Z})^* \to \mathbb{C}^*$ satisfying $\chi(n+q) = \chi(n)$, $\chi(n) = 0$ if $(n, q) > 1$, and $\chi(mn) = \chi(m)\chi(n)$.

In the case $q = 4$, we have characters

$$\chi_0(n) = \begin{cases} 1 & n \text{ is odd} \\ 0 & n \text{ is even,} \end{cases}$$

or more interestingly,

$$\chi_{-4}(n) = \begin{cases} 1 & n \equiv 1 \bmod 4 \\ -1 & n \equiv 3 \bmod 4 \\ 0 & n \text{ is even.} \end{cases}$$

In this second case, we can define the L-function

$$L(s, \chi_{-4}) = \sum_{n=1}^{\infty} \frac{\chi_{-4}(n)}{n^s} = \prod_{p \equiv 1 \bmod 4} \left(1 - \frac{1}{p^s}\right)^{-1} \prod_{p \equiv 3 \bmod 4} \left(1 + \frac{1}{p^s}\right)^{-1} = \prod_p \left(1 - \frac{\chi_{-4}(p)}{p^s}\right)^{-1}.$$

This also satisfies the functional equation

$$\left(\frac{4}{\pi}\right)^{\frac{s+1}{2}} \Gamma\left(\frac{s+1}{2}\right) L(s, \chi_{-4}) = \left(\frac{4}{\pi}\right)^{s/2} \Gamma(-s/2) L(1-s, \chi_{-4}).$$

Now, we see that some combinations of $\chi_{-4}(n)$ picks out arithmetic progressions:

$$\frac{\chi_{-4}(n) + \chi_{-4}(n)}{2} = \begin{cases} 1 & n \equiv 1 \bmod 4 \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\chi_{-4}(n) - \chi_{-4}(n)}{2} = \begin{cases} 1 & n \equiv 3 \bmod 4 \\ 0 & \text{otherwise.} \end{cases}.$$

We can consider the logarithmic derivative

$$-\frac{L'}{L}(s, \chi_{-4}) = \sum \frac{\Lambda(n)\chi_{-4}(n)}{n^s}.$$

Then $\sum_{n\le x}\Lambda(n)\chi_{-4}(n)$ corresponds to singularities of $\frac{L'}{L}(s,\chi_{-4})$. It turns out that $L(s,\chi_{-4})$ has no pole, and the crucial fact is that $L(1+it,\chi_{-4})\neq 0$. How would we show this? We can write that

$$L(1+it,\chi_{-4})=\prod\left(1-\frac{\chi_{-4}(p)}{p^{1+it}}\right)^{-1}.$$

For this to be nonzero, we expect that

$$\frac{\chi_{-4}(p)}{p^{it}}\approx -1$$

or equivalently, that $\frac{\chi_{-4}(p)^2}{p^{2it}}\approx 1$ a lot of the time. Now, we consider

$$\zeta(\sigma)^3|L(\sigma+it,\chi_{-4})|^4|\zeta(\sigma+2it)|\ge 1$$

We need to check that $L(1,\chi_{-4})\neq 0$, which is true because it's an alternating series: $L(1,\chi_{-4})=1-\frac{1}{3}+\frac{1}{5}-\cdots=\frac{\pi}{4}$. This is a crucial point, and this gives that there are no zeros on $\operatorname{Re} s=1$.

This argument generalizes for all q. There are $\varphi(q)$ characters $\chi \bmod q$. They form an abelian group isomorphic to $(\mathbb{Z}/q\mathbb{Z})^*$, and every arithmetic progression mod q can be written in terms of these characters:

$$\frac{1}{\varphi(q)}\sum_{x \bmod q}\chi(n)\overline{\chi(a)}=\begin{cases}1 & n\equiv a \bmod q\\ 0 & \text{otherwise.}\end{cases}$$

Then we have

$$\frac{1}{\varphi(q)}\sum_{x \bmod q}\overline{\chi(a)}\left(\sum_{n\le x}\Lambda(n)\chi(n)\right)-\sum_{\substack{n\le x\\ n\equiv a \bmod q}}\Lambda(n).$$

We consider

$$L(s,\chi)=\sum\frac{\chi(n)}{n^s}.$$

There is always a principal character

$$\chi_0=\begin{cases}1 & (n,q)=1\\ 0 & \text{otherwise,}\end{cases}$$

which looks almost like $\zeta(s)$. It happens that $L(s,\chi_0)$ has a pole at $s=1$, and all other L-functions are analytic. This pole gives the main term, and everything else gives an error term. The main crucial point is that $L(1+it,\chi)\neq 0$, which we do by a generalization of the argument above; we consider

$$\zeta(\sigma)^3|L(\sigma+it,\chi)|^4|L(\sigma+2it,\chi^2)|.$$

This requires knowing that $L(1,\chi)\neq 0$, and pushing through this calculation yields Dirichlet's theorem for primes in arithmetic progressions.

These L-functions all have functional equations like that for $\zeta(s)$, and they are also expected to satisfy the Riemann Hypothesis; their zeros should also all lie on $\operatorname{Re} s=\frac{1}{2}$.

15. 5/24

For the next few lectures, we will discuss the number of partitions of n, which we will denote $p(n)$.

Recall that this had a generating function:

$$\sum_{n=0}^{\infty} p(n)z^n = \prod_{j=1}^{\infty}(1-z^j)^{-1}.$$

Hardy and Ramanujan proved an asymptotic for the partition function.

Theorem 15.1 (Hardy and Ramanujan (1918))**.**

$$p(n) \sim \frac{e^{\pi\sqrt{\frac{2}{3}}\sqrt{n}}}{4n\sqrt{3}}$$

15.1. **Rough estimates for an upper bound.** First, where do the product and series converge? The product converges absolutely in $|z|<1$, and the series also converges absolutely in $|z|<1$.

Let

$$F(z) = \prod_{j=1}^{\infty}(1-z^j)^{-1},$$

and we will think of this as a function of a complex variable z. If z is a real number $r<1$, an immediate upper bound is $p(n) \le C(\varepsilon)e^{\varepsilon n}$. This is because if $z = e^{-\varepsilon}$, then we are studying $\sum p(n)e^{\varepsilon n} < \infty$, which converges.

We can even formalize this like so: For any $0 < r < 1$, we can say that $p(n)r^n \le \sum p(j)r^j = F(r)$, which means that $p(n) \le F(r)r^{-n}$. In particular, the best that we can do is $p(n) \le \min_{0<r<1}(F(r)r^{-n})$. So we want to find the value of r where this minimum is attained; this will allow us to obtain some bounds for $p(n)$. The point here is that from bounds for $F(r)$ as $r \to 1^-$, we can obtain bounds for $p(n)$.

Our first guess is that the minimum is attained at $r = r(n) \to 1$ as $n \to \infty$. We expect that $F(r)$ will be increasing as r increases, and it should become very big. But r^{-n} decreases, and by calculus, we should be able to find a minimum.

For the moment, let's adopt the notation that $r = e^{-1/N}$; then letting $r \to 1^-$ is equivalent to letting $N \to \infty$. Then

$$F(r) = \prod_{j=1}^{\infty}(1-e^{-j/N})^{-1}.$$

There are two things here, when j is large or j is small. When j is large, we don't have to care too much, and when j is small, this looks like some sort of expansion, and we can write

$$F(r) = \prod_{j=1}^{\infty}(1-e^{-j/N})^{-1} \approx \prod_{l\le N}\left(\frac{N}{j}\right) \approx e^N.$$

So then $F(r)r^{-n} \approx e^{N+n/N}$. Setting $N \approx \sqrt{n}$ yields that $F(r)r^{-n} \approx e^{2\sqrt{n}}$, which explains the $e^{\sqrt{n}}$ term in the asymptotic. Here, this was a very rough estimate, and we didn't keep track of our constants.

Now we do a more precise version of this. Here, we have

$$\log F(z) = \sum_{j=1}^{\infty} -\log(1-z^j) = \sum_{j=1}^{\infty}\sum_{\ell=1}^{\infty} \frac{z^{j\ell}}{\ell} = \sum_{n=1}^{\infty} z^n \left(\sum_{j\ell=n} \frac{1}{\ell}\right).$$

Here,

$$\sum_{j\ell=n} \frac{1}{\ell} = \frac{1}{n}\sum_{j\ell=n} j = \frac{\sigma(n)}{n}$$

where $\sigma(n) = \sum_{d|n} d$. So therefore

$$\log F(z) = \sum_{n=1}^{\infty} \frac{\sigma(n)}{n} z^n.$$

We want to minimize $\log(F(r)r^{-n}) = \log F(r) - n\log r$, and by calculus, this means that we want to solve for $\frac{F'}{F}(r) = \frac{n}{r}$. Here, we have

$$z\frac{F'}{F}(z) = \sum_{n=1}^{\infty} \sigma(n) z^n,$$

so we are trying to solve

$$\sum_{n=1}^{\infty} \sigma(n) r^n = n.$$

15.2. **Obtaining an asymptotic formula.** So far, we were just trying to get upper bounds, but we claim that this can also give us an asymptotic formula. We have $F(z) = \sum p(n)z^n$. Then we have

$$\frac{1}{2\pi i}\int_{|z|=r} F(z)\frac{dz}{z^{n+1}} = p(n).$$

We write $z = re^{i\theta}$ and $-\pi < \theta \le \pi$. Then this integral that we want to evaluate will be

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} F(re^{i\theta}) r^{-n} e^{-in\theta}\, d\theta.$$

Observe that we have a bound $|F(re^{i\theta})| \le \sum p(n)r^n = F(r)$, but we might have some cancellation; for example, if $\theta = \pi$, this might be some alternating series. Using this bound, we have that our integral is bounded by $F(r)r^{-n}$. Our goal is to evaluate this integral asymptotically for the value of r which minimizes $F(r)r^{-n}$.

The method by which we will do this is called the *saddle point method* or the *method of stationary phase*. The integrand is $F(re^{i\theta})r^{-n}e^{-in\theta}$. If we go along the real axis, we know that this function attains a minimum at some $r = r_0(n)$ that is a solution to $\sum_{n=1}^{\infty} \sigma(n)r^n = n$. What happens if we fix r and vary θ? Here, $r^{-n}e^{-in\theta}$ doesn't change size; $F(re^{i\theta})$ has a local maximum at $\theta = 0$. This explains the name "saddle point"; there is a local minimum on the real axis, and a local maximum upon varying θ. Note also that

$$\log(F(re^{i\theta})) = \log(F(r)) + \theta\,\frac{d}{d\theta}(\log(F(re^{i\theta})))\bigg|_{\theta=0} + \frac{\theta^2}{2}\,\frac{d^2}{d\theta^2}(\log(F(re^{i\theta})))\bigg|_{\theta=0} + \cdots.$$

The first derivative comes out to be purely imaginary:

$$\frac{F'}{F}(r) \cdot r \cdot i = in$$

at $r = r_0$. This justifies the name "stationary phase". The second derivative looks like $-(\cdots)\frac{\theta^2}{2}$, which means that this drops off quickly as θ moves away from 0. For small values of θ, we will see that the higher derivatives are negligible, and we will end up with some Gaussian $\int e^{-(\cdots)\theta^2/2}\, d\theta$.

15.3. **Determining the saddle point.** Now, the question is to determine the saddle point.

We have

$$\log F(r) = \sum_{n=1}^{\infty} \frac{\sigma(n)}{n} r^n,$$

and we want to solve $\sum_{n=1}^{\infty} \sigma(n) r^n = N$. We write $r = e^{-\frac{2\pi}{x}}$, so that $r \to 1^-$ corresponds to $x \to \infty$. Now, we have

$$\log F(e^{-\frac{2\pi}{x}}) = \sum_{n=1}^{\infty} \frac{\sigma(n)}{n} e^{-2\pi \frac{n}{x}}$$

As a bit of a review, by Perron's formula, we can write

$$\sum_{n \le x} \frac{\sigma(n)}{n} = \frac{1}{2\pi i} \int_{(c)} \left(\sum \frac{\sigma(n)}{n^{1+s}} \right) x^s \frac{ds}{s}.$$

Then our sum converges when $\operatorname{Re} s > 1$, and we can write $\sigma(n)$ as a Dirichlet convolution $\sigma = 1 \star n$, and therefore,

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

So hence

$$\sum_{n \le x} \frac{\sigma(n)}{n} = \frac{1}{2\pi i} \int_{(c)} \zeta(s)\zeta(s+1) x^s \frac{ds}{s} \sim \zeta(2) x.$$

since that is approximately the residue for the first pole that we encounter.

Then we can do partial summation to evaluate

$$\begin{aligned}\log F(e^{-\frac{2\pi}{x}}) = \sum_{n=1}^{\infty} \frac{\sigma(n)}{n} e^{-2\pi \frac{n}{x}} &= \int_0^{\infty} e^{-2\pi \frac{t}{x}}\, d\left(\sum_{n \le t} \frac{\sigma(n)}{n} \right) \\ &\approx \zeta(2) \int_0^{\infty} e^{-2\pi \frac{t}{x}}\, dt = \frac{\zeta(2) x}{2\pi} = \frac{\pi x}{12}.\end{aligned}$$

At the moment, we haven't kept track of our error terms, but let's assume that something like this is correct. So differentiating, we see that

$$\frac{F'}{F}\left(e^{-2\pi/x}\right) \cdot \left(\frac{2\pi}{x^2}\right) e^{-2\pi/x} = \frac{\pi}{12},$$

so therefore

$$e^{-2\pi/x} \cdot \frac{F'}{F}\left(e^{-2\pi/x}\right) = N = \frac{x^2}{24},$$

which means that we should pick $x \approx \sqrt{24N}$ and hence $r \approx e^{-2\pi/\sqrt{24N}}$. This means that

$$p(N) \le F(r)r^{-N} = \exp\left(\frac{\pi\sqrt{24N}}{12} + \frac{2\pi}{\sqrt{24}}\sqrt{N}\right) = \exp\left(\frac{\pi\sqrt{2}\sqrt{N}}{\sqrt{3}}\right),$$

which is already very close to the right asymptotic. One thing that we've neglected here is to keep track of the error terms.

Now, we go back to the real problem, which is

$$\log F\left(e^{-2\pi/x}\right) = \sum_{n=1}^{\infty} \frac{\sigma(n)}{n} e^{-2\pi n/x}.$$

The answer is much nicer than we might expect. Recall that for $c > 0$

$$\frac{1}{2\pi i}\int_{(c)} \Gamma(s)x^s\,ds = e^{-1/x}.$$

The point is that the Gamma function has rapid decay, and we pick up the residues of the Gamma function at its poles, picking up the Taylor expansion for $e^{-1/x}$. That means that we can write

$$\begin{aligned}\log F\left(e^{-2\pi/x}\right) &= \sum_{n=1}^{\infty} \frac{\sigma(n)}{n} e^{-2\pi n/x} = \frac{1}{2\pi i}\int_{(c)} \sum_{n=1}^{\infty} \frac{\sigma(n)}{n^{1+s}}(2\pi)^{-s}\Gamma(s)x^s\,ds \\ &= \frac{1}{2\pi i}\int_{(c)} (2\pi)^{-s}\Gamma(s)\zeta(s)\zeta(s+1)x^s\,ds.\end{aligned}$$

We do this carefully next time, but crudely, the main term should be the residue from the pole at $s = 1$, which gives us $\frac{\zeta(2)x}{2\pi}$. There is also a double pole at $s = 0$, which we have to evaluate. A miracle that happens is that we will be able to exactly evaluate this. We will find that there is a relation between $\log F(e^{-2\pi/x})$ and $\log F(e^{-2\pi x})$.

16. 5/29

We continue the discussion from last time. We wanted to understand

$$\log F\left(e^{-2\pi/x}\right) = \sum_{n=1}^{\infty} \frac{\sigma(n)}{n} e^{-2\pi n/x} = \frac{1}{2\pi i}\int_{(c)} (2\pi)^{-s}\Gamma(s)\zeta(s)\zeta(s+1)x^s\,ds.$$

We do this by shifting contours. There is a pole at $s = 1$ from $\zeta(s)$, there is a double pole at $s = 0$ from $\zeta(s+1)$ and $\Gamma(s)$, and there is a pole at $s = -1$ from $\Gamma(s)$. These are the only poles.

Let $X(s) = (2\pi)^{-s}\Gamma(s)\zeta(s)\zeta(s+1)$.

Proposition 16.1. *$X(s) = X(-s)$.*

Proof. We claim that we can write

$$X(s) = C\left(\pi^{-s/2}\Gamma(s/2)\zeta(s)\right)\left(\pi^{-(1+s)/2}\Gamma((1+s)/2)\zeta(1+s)\right).$$

This uses the *duplication formula* for $\Gamma(s)$:

$$\frac{2^{2z-1}}{\sqrt{\pi}}\Gamma(z)\Gamma(z+1/2) = \Gamma(2z),$$

and we can rewrite this to become

$$\Gamma(s/2)\Gamma((s+1)/2) = \Gamma(s)\frac{\sqrt{\pi}}{2^{s-1}},$$

which gives us the product that we want.

Using the functional equation for the zeta function yields that

$$X(s) = C\left(\pi^{-(1-s)/2}\Gamma((1-s)/2)\zeta(1-s)\right)\left(\pi^{-(-s)/2}\Gamma(-s/2)\zeta(-s)\right) = X(-s),$$

which is what we wanted to show. $\square$

Now, for $c > 1$, we have to evaluate

$$\log F(-e^{2\pi/x}) = \frac{1}{2\pi i}\int_{(c)} X(s)x^s\, ds.$$

Moving contours to $(-c)$, we pick up residues at $s = 1, 0, -1$ of $X(s)x^s$. Let's evaluate these residues.

The residue at 1 is

$$\operatorname*{Res}_{s=1} X(s)x^s = \frac{1}{2\pi}\zeta(2)x = x\frac{\pi}{12}.$$

The residue at -1 is

$$\operatorname*{Res}_{s=-1} X(s)x^s = -\frac{\pi}{12}\frac{1}{x}$$

because of our expression $X(s) = X(-s)$. Finally, we consider the residue at $s = 0$. Since $X(s)$ has a double pole, we should have $X(s) = \frac{c}{s^2} + c' +$ higher order terms (since $X(s)$ is also even). Then we have that

$$\operatorname*{Res}_{s=0} X(s)x^s = \zeta(0)\log x = -\frac{1}{2}\log x.$$

So the residues give us that the residues are $\frac{\pi}{12}(x - \frac{1}{x}) - \frac{1}{2}\log x$. We do a change of variable $w = -s$ on the remaining integral:

$$\frac{1}{2\pi i}\int_{(-c)} X(s)x^s\, ds = \frac{1}{2\pi i}\int_{c+i\infty}^{c-i\infty} X(w)x^{-w}\,(-dw) = \frac{1}{2\pi i}\int_{(c)} X(w)x^{-w}\, dw. = \log F(e^{-2\pi x}).$$

Theorem 16.2. *For all $x > 0$,*

$$\log F(e^{-2\pi/x}) - \log F(e^{-2\pi x}) = \frac{\pi}{12}\left(x - \frac{1}{x}\right) - \frac{1}{2}\log x.$$

Why is this useful? We want to understand $\log F(e^{-2\pi/x})$ for large x, but when x is large, we observe that

$$\log F(e^{-2\pi x}) = -\sum \frac{\sigma(n)}{n}e^{-2\pi i n x},$$

which is incredibly small. So as a corollary, we have:

Corollary 16.3.

$$\log F(e^{-2\pi/x}) = \frac{\pi}{12}\left(x - \frac{1}{x}\right) - \frac{1}{2}\log x + O\left(e^{-2\pi x}\right).$$

Now we can understand how to choose the radius r. Differentiating gives

$$\left(\frac{F'}{F}\left(e^{-2\pi/x}\right)e^{-2\pi/x}\right)\left(\frac{2\pi}{x^2}\right)+\frac{F'}{F}\left(e^{-2\pi x}\right)e^{-2\pi x}(2\pi)=\frac{\pi}{12}\left(1+\frac{1}{x^2}\right)-\frac{1}{2x}.$$

We ignore the exponentially small term, and we want to solve

$$N\frac{2\pi}{x^2}=\frac{\pi}{12}\left(1+\frac{1}{x^2}\right)-\frac{1}{2x}.$$

Then canceling things properly, we have

$$N=\frac{x^2}{24}+\frac{1}{24}-\frac{x}{4\pi},$$

so therefore

$$x=\frac{3}{\pi}+\frac{1}{2}\sqrt{4(24N-1)+\left(\frac{6}{\pi}\right)^2}.$$

So now we know precisely how to find the saddle point, and evaluate our function at the saddle point. We make this choice of x, and we have an asymptotic formula for $\log F(e^{-2\pi/x})$. Then

$$F\left(e^{-2\pi/x}\right)\left(e^{2\pi/x}\right)^N=\exp\left(\frac{\pi x}{6}-\frac{1}{2}-\frac{1}{2}\log x\right)=\frac{\exp\left(\frac{\pi}{12}\sqrt{4(24N-1)+\left(\frac{6}{\pi}\right)^2}\right)}{\sqrt{x}}.$$

This could be a good upper bound for the number of partitions. At the moment, we have that

$$p(n)\ll\frac{e^{\pi\sqrt{\frac{2}{3}N}}}{N^{1/4}},$$

but the true asymptotic has $4N\sqrt{3}$ in the denominator. We need to figure out where this comes from.

The last step is to understand (for this particular choice of r)

$$\frac{1}{2\pi i}\int_{|z|=r}F(z)z^{-N}\,\frac{dz}{z}=\frac{1}{2\pi}\int_{-\pi}^{\pi}F(re^{i\theta})e^{-iN\theta}r^{-N}\,d\theta.$$

Here, if we look at

$$\log F(re^{i\theta})=\sum_{n=1}^{\infty}\frac{\sigma(n)}{n}r^ne^{in\theta}.$$

This is maximized at $\theta=0$, and as we move away from $\theta=0$, this takes on substantially lower values.

To understand this better, we will consider the cases where $|\theta|$ is small and $|\theta|$ not small separately.

First, let's consider the case where $|\theta|$ is small. We write out the first few terms in the Taylor series of $e^{in\theta}$. So we have

$$\begin{aligned}\log F(re^{i\theta}) &= \sum_{n=1}^{\infty} \frac{\sigma(n)}{n} r^n \left(1 + in\theta - \frac{n^2\theta^2}{2} + O(n^3|\theta|^3)\right) \\ &= \log F(r) + iN\theta - \frac{\theta^2}{2}\left(\sum_{n=1}^{\infty} \sigma(n) n r^n\right) + O\left(|\theta|^3 \sum_{n=1}^{\infty} n^2 \sigma(n) r^n\right).\end{aligned}$$

Observe that we have $\sum \frac{\sigma(n)}{n} e^{-2\pi n/x} \sim cx$, and $\sum \sigma(n) e^{-2\pi n/x} = N \sim cx^2$, so the final two sums should be of size cx^3 and cx^4 respectively. Of course, we have to work out what this constant is.

So for small $|\theta|$, we have that the integrand

$$F(re^{i\theta})e^{-iN\theta} = F(r)e^{-\frac{\theta^2}{2}(cx^3) + O(|\theta|^3 x^4)} = F(r)e^{-\frac{\theta^2}{2}(cN^{3/2}) + O(|\theta|^3 N^2)}.$$

Then we have

$$\frac{F(r)}{2\pi} \int_{|\theta| < ?} e^{-\frac{\theta^2}{2}(\cdots) + O(|\theta|^3 N^2)}\, d\theta,$$

which suggests that we should consider $|\theta| \le N^{-\frac{2}{3}-\delta}$ so that the error is small. So long as $|\theta| \le N^{-\frac{3}{4}+\delta}$, this integral will contribute substantially. So we consider $|\theta| \le N^{-0.7}$. Then this integral is almost the full Gaussian integral, so we have

$$\frac{F(r)}{2\pi} \int_{-N^{-0.7}}^{N^{-0.7}} e^{-\frac{\theta^2}{2}(cN^{3/2}) + O(N^{-0.1})}\, d\theta \sim \frac{\sqrt{2\pi}}{\sqrt{cN^{3/2}}}$$

We had that $F(r)r^{-N} \sim \frac{e^{\pi\sqrt{\frac{2}{3}N}}}{cN^{1/4}}$, and we have that this is significant only on an interval $|\theta| \le \frac{1}{N^{3/4}}$. Multiplying these gives the asymptotic that we want.

We now have to consider the second case, where $\pi \ge |\theta| \ge N^{-0.7}$. Here, we can't just write down a Taylor expansion or the error terms will become big. Here, we want bounds on

$$\frac{1}{2\pi} \int_{\pi \ge |\theta| \ge N^{-0.7}} |F(re^{i\theta})| r^{-N}\, d\theta$$

We expect that $|F(re^{i\theta})|$ is much smaller than $|F(r)|$, so therefore

$$\log |F(re^{i\theta})| = \sum \frac{\sigma(n)}{n} r^n \cos(n\theta).$$

Now θ is not too close to zero, so we expect $\cos(n\theta)$ should often be away from 1. In particular,

$$\begin{aligned}&\log F(r) - \log|F(re^{i\theta})| = \sum_{n=1}^{\infty} \frac{\sigma(n)}{n} r^n (1 - \cos n\theta) \ge \sum_{n=1}^{\infty} r^n (1 - \cos n\theta) \\ &= \frac{r}{1-r} - \mathrm{Re}\left(\frac{re^{i\theta}}{1 - re^{i\theta}}\right) = \mathrm{Re}\, \frac{r - r^2 e^{i\theta} - re^{i\theta} + r^2 e^{i\theta}}{(1-r)(1-re^{i\theta})} = \frac{r}{1-r} \mathrm{Re}\left(\frac{1 - e^{i\theta}}{1 - re^{i\theta}}\right),\end{aligned}$$

and now it should be straightforward to give a good lower bound, which we expect to look like $\theta^2 N^{3/2}$. There are some details here to work out, but the idea of the proof should be intuitive. We want to understand the value of something by integrating around a circle. We have a saddle point which is a maximum in θ and a minimum in r. As we go around the circle, it falls off in θ at some rate that we can control.

Observe that

$$\frac{1}{2\pi i}\int_{|z|=r} e^z z^{-N-1}\,dz = \frac{1}{N!}.$$

So maybe we can use the saddle point method to find some asymptotics for $\frac{1}{N!}$. We want to find r so that $e^r r^{-N}$ is minimized, which occurs when $r = N$. Then we choose $r = N$, and consider the integral as a function of θ. That is, we hope to control

$$\frac{1}{2\pi}\int_{\theta=-\pi}^{\pi} e^{Ne^{i\theta}} N^{-N} e^{-iN\theta}\,d\theta.$$

We write $e^{i\theta} = 1 + i\theta - \frac{\theta^2}{2} - \cdots$ and note that the phase cancels out. So now we get the integral of a Gaussian over an interval of length $\sqrt{N}$, which is another way of finding Stirling's formula.

17. 5/31

These notes are typed from Ravi's notes.

17.1. **Recap of the saddle point method.** We start with the example that we started last time: Finding Stirling's formula via the stationary phase method.

Example 17.1. We have that

$$\frac{1}{N!} = \frac{1}{2\pi i}\int_{|z|=r}\frac{e^z}{z^n}\frac{dz}{z} = \frac{1}{2\pi i}\int_{-\pi}^{\pi} e^{re^{i\theta}}(r^{-N}e^{-iN\theta})\,d\theta.$$

We choose r to minimize $e^r r^{-N}$, so $r = N$. Then we want to control

$$\frac{1}{2\pi i}\int_{-\pi}^{\pi} e^{Ne^{i\theta}-iN\theta}N^{-N}\,d\theta.$$

When $|\theta|$ is small, we use $Ne^{i\theta} - iN\theta = N - \frac{1}{2}N\theta^2 + O(N|\theta|^3)$, while for $|\theta|$ not too small, we have $|e^{Ne^{i\theta}}| = e^{N\cos\theta} \le e^N e^{-cN\theta^2}$. On $|\theta| < N^{0.4}$, the error is small but the other terms are large. On this interval, we then get

$$\frac{1}{2\pi}\left(\frac{e}{N}\right)^N\int_{|\theta|\le N^{-0.4}} e^{-N\theta^2/2}(1+O(N^{-0.2}))\,d\theta = \frac{\sqrt{2\pi}}{2\pi}\frac{N^{-N}e^N}{\sqrt{N}}(1+O(N^{-0.2})).$$

Now, we return to the example of partitions. Recall that we defined

$$F(z) = \sum p(n)z^n = \prod_{j=1}^{\infty}(1-z^j)^{-1}.$$

We chose $r = e^{-2\pi/x}$, and we determined that

$$\log F(e^{-2\pi/x}) - \log F(e^{-2\pi x}) = \frac{\pi}{12}\left(x - \frac{1}{x}\right) - \frac{1}{2}\log x.$$

We then showed that

$$x = \frac{3}{\pi} + \sqrt{24N - 1 + \left(\frac{3}{\pi}\right)^2}$$

is optimal, and from there,

$$F(r)r^{-N} \sim \frac{\exp(\frac{\pi}{6}\sqrt{24N})}{(24N)^{1/4}}$$

From here, we did a more careful estimate. The ideas behind this are connected to modular forms.

17.2. **Modular forms.** Let $\mathbb{H}$ be the upper half plane. There is a group action of $SL_2(\mathbb{R})$ on $\mathbb{H}$. If $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ for $ad - bc = 1$, then g acts by

$$gz = \frac{az+b}{cz+d}.$$

These are the linear fractional transformations.

Exercise 17.2. Check that this acts on $\mathbb{H}$. To do this, observe that $\operatorname{Im}(gz) = \frac{\operatorname{Im} z}{|cz+d|^2}$.

$SL_2(\mathbb{R})$ has a nice subgroup $SL_2(\mathbb{Z})$.

Our goal is to understand functions on $\mathbb{H}$ which transform nicely under $SL_2(\mathbb{Z})$. For example, we might want functions such that $f(\gamma z) = f(z)$ for all $\gamma \in SL_2(\mathbb{Z})$. We might ask for meromorphic, holomorphic, or C^∞ functions satisfying this condition. It turns out that there are no holomorphic functions for which this is true, so we instead ask for $f(\gamma z) = j(\gamma, z)f(z)$ for some multiplier systems j.

Observe that we need to have

$$f(\gamma_1\gamma_2 z) = j(\gamma_1\gamma_2, z)f(z) = j(\gamma_1, \gamma_2 z)j(\gamma_2, z)f(z),$$

so we must have $j(\gamma_1\gamma_2, z) = j(\gamma_1, \gamma_2 z)j(\gamma_2, z)$, which looks like a chain rule. What kinds of j satisfy this? The main example that we will consider is $j(\gamma, z) = (cz+d)^k$ where $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$.

Exercise 17.3. Check that this satisfies the relation above.

For fixed $k > 0$, we will now study holomorphic functions $f : \mathbb{H} \to \mathbb{C}$ satisfying $f(\gamma z) = (cz+d)^k f(z)$ for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$.

Note that I and $-I$ have the same action on $\mathbb{H}$, so therefore $f(-Iz) = f(z)$, and hence k is even.

There are a nice class of elements of $SL_2(\mathbb{Z})$ given by $\left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right)$. These send $z \mapsto z+n$. In this case, we have $(cz+d)^k = 1$, so therefore $f(z+n) = f(z)$ is periodic with period 1, among many other relations.

Another nice matrix is $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, representing the map $z \mapsto -1/z$ and hence giving that $f(-1/z) = z^k f(z)$. This relation should look like the functional equation for $\theta(x, \chi)$.

Definition 17.4. If f satisfies the above conditions, then f is a *modular function.*

Definition 17.5. A *modular form* is a modular function that is also holomorphic at ∞.

Why is this definition necessary? Suppose $z \in \mathbb{H}$, and consider $q = e^{2\pi i z} \in D(0,1)^*$. Here, $i\infty$ corresponds to 0, so define "hole at ∞" to mean "$f(e^{2\pi i z})$ has a removable singularity at 0". Equivalently, this condition means that f is bounded as $\operatorname{Im} z \to \infty$. Here, k is called the *weight* of the modular form.

Example 17.6. Recall that

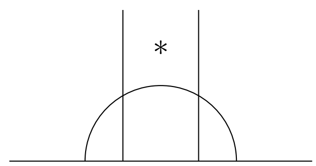
$$\sum p(n)q^n = \prod_{j=1}^{\infty}(1-q^j)^{-1}.$$

A related example is

$$q\prod_{n=1}^{\infty}(1-q^n)^{24} = \sum_{n=1}^{\infty}\tau(n)q^n = \sum_n \tau(n)e^{2\pi inz}$$

where $\tau(n)$ is Ramanujan's tau function. This is a modular form of weight 12.

Theorem 17.7. *$PSL_2(\mathbb{Z}) = SL_2(\mathbb{Z})/\{\pm I\}$ is generated by $\left(\begin{smallmatrix}1 & n\\ 0 & 1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$. Moreover, there exists a* fundamental domain*. Here, draw the lines $\operatorname{Re} z = \pm 1/2$, and take the region between them and above $|z| = 1$. Include the right half but not the left half of the boundary.*



Remark. This is related to the reduction of binary quadratic forms, which was studied by Gauss.

The idea of the proof is that we start with z, and we translate to $(\frac{1}{2}, \frac{1}{2}] + i\mathbb{R}$. If necessary, we invert. Repeat this if necessary. We claim that after finitely many operations, we'll get into the fundamental domain, and we claim that no two points in the domain are equivalent.

18. 6/5

Last time, we had the upper half plane $\mathbb{H} = \{z = x + iy : y > 0\}$. There is a group

$$SL_2(\mathbb{R}) = \left\{\begin{pmatrix}a & b\\ c & d\end{pmatrix} : ad = bc = 1, a, b, c, d \in \mathbb{R}\right\}$$

that acts on the upper half plane via

$$gz = \frac{az+b}{cz+d},$$

sending $\mathbb{H} \to \mathbb{H}$.

We in particular considered $SL_2(\mathbb{Z}) \subseteq SL_2(\mathbb{R})$, and we considered *modular forms* of weight k, which are given in the following way: Take $\gamma = \left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right) \in SL_2(\mathbb{Z})$, and let $f(\gamma z) = (cz+d)^k f(z)$. We require that for $\left(\begin{smallmatrix}1 & n\\ 0 & 1\end{smallmatrix}\right)$, $n \in \mathbb{Z}$, we have $f(z+n) = f(z)$, and that k is even, f is holomorphic in $\mathbb{H}$, and as $\operatorname{Im} z \to \infty$, $|f(z)|$ is bounded.

For $q = e^{2\pi iz}$, we have $f(z) = \sum_{k=0}^{\infty} a_k e^{2\pi ikz}$, and $f(q) = a_0 + a_1 q + a_2 q^2 + \cdots$.

Theorem 18.1. *$SL_2(\mathbb{Z})/\pm I$ is generated by $\left(\begin{smallmatrix}1 & n\\ 0 & 1\end{smallmatrix}\right) = \left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)^n$, and inversion $\left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$. There is a fundamental domain $\mathcal{F}$, which is the region between $\operatorname{Re} z = \pm 1/2$ and above $|z| = 1$.*

More precisely, if $\Gamma_1 = \langle \left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right), \left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)\rangle$ then for all $z \in \mathbb{H}$, there exists $\gamma \in \Gamma_1$ with $\gamma z \in \mathcal{F}$. Further, if $z_1, z_2 \in \mathcal{F}$ with $z_1 \neq z_2$ then $z_1 \neq \gamma z_2$ for all $\gamma \in SL_2(\mathbb{Z})$.

Proof. Note that for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, we have

$$\operatorname{Im}(\gamma z) = \frac{\operatorname{Im}(z)}{|cz+d|^2}.$$

Note that if either c or d gets large, then the imaginary part is small. Now, look at all matrices where c and d lie in some bounded set; then the imaginary parts will lie in some bounded set.

So there is some maximal value for $\operatorname{Im}(\gamma z)$. Pick the largest value of $\operatorname{Im}(\gamma z)$.

By translating, we can assume that the $-1/2 \le \operatorname{Re} z \le 1/2$. We claim that γz now lies in the fundamental domain. If not, replace $z \to -1/z$, and observe that $\operatorname{Im}(-1/z) = \operatorname{Im}(z)/|z|^2$, which increases the imaginary part above, which contradicts the assumption that we picked the largest value of $\operatorname{Im}(\gamma z)$. So we can translate to get inside the fundamental domain. $\square$

Now, we give examples of modular forms and see how they look. We only have to check that $f(z+1) = f(z)$ and $f(-1/z) = z^k f(z)$.

Example 18.2. A nice set of modular forms are called *Eisenstein series*. Here, we take $k \ge 4$ to be even, and let

$$G_k(z) = \sum_{(m,n)\neq(0,0)} \frac{1}{(mz+n)^k}.$$

The first thing to notice is that this series actually converges absolutely. To see this, observe that we are summing over a two-dimensional lattice, and we can sum over annuli. In fact, this converges for $k > 2$, and barely fails to converges for $k = 2$.

We claim that $G_k(z)$ is a modular form of weight k. First, observe that trivially, $G_k(z+1) = G_k(z)$. Also,

$$G_k(-1/z) = \sum_{(m,n)\neq(0,0)} \frac{1}{(-\frac{m}{z}+n)^k} = z^k \sum \frac{1}{(-m+nz)^k} = z^k G_k(z).$$

What does this have to do with what we have considered in this class? Let's compute the Fourier expansion for $G_k(z)$.

Note that

$$G_k(z) = 2\zeta(k) + 2\sum_{m=1}^{\infty}\sum_{n=-\infty}^{\infty} \frac{1}{(mz+n)^k}.$$

So let

$$F(w) = \sum_{n=-\infty}^{\infty} \frac{1}{(w+n)^k}$$

for $w = mz \in \mathbb{H}$; we hope to understand this formula. By the Poisson summation formula 13.5, we can write

$$F(w) = \sum_{\ell\in\mathbb{Z}} \left(\int_{-\infty}^{\infty} \frac{1}{(w+t)^k} e^{-2\pi i t\ell}\, dt \right)$$

The Fourier coefficients can be rewritten as

$$\int_{-\infty}^{\infty} \frac{1}{(w+t)^k} e^{-2\pi i t\ell}\, dt = \int_{i\operatorname{Im}(w)-\infty}^{i\operatorname{Im}(w)+\infty} \frac{1}{z^k} e^{-2\pi i (z-w)\ell}\, dz = e^{2\pi i w\ell} \int_{i\operatorname{Im}(w)-\infty}^{i\operatorname{Im}(w)+\infty} \frac{1}{z^k} e^{-2\pi i z\ell}\, dz.$$

This now has a singularity at $z = 0$, so we can try to evaluate this using contour shifts. We should either move up to go to zero and pass through no poles, or move down to pick up a pole. When should we do each?

In the integrand we have

$$\left|e^{2\pi i z\ell}\right| = e^{2\pi(\operatorname{Im} z)\ell}$$

So if $\ell \leq 0$, we should move the line of integration upwards, encounter no poles, and get a result of 0. But if $\ell > 0$, if we move upwards, we get something growing exponentially, which is no good. So we move downwards. Note that we flip orientation when we do this; then we get

$$-(2\pi i)e^{2\pi i w\ell} \operatorname*{Res}_{z=0} \frac{1}{z^k} e^{-2\pi i z\ell} = -(2\pi i)\frac{(-2\pi i\ell)^{k-1}}{(k-1)!} e^{2\pi i w\ell} = \frac{(2\pi i)^k \ell^{k-1}}{(k-1)!} e^{2\pi i \ell w}.$$

So now we have a very pretty formula. Poisson summation tells us that

$$F(w) = \sum_{n\in\mathbb{Z}} \frac{1}{(w+n)^k} = \sum_{\ell=1}^{\infty} \frac{(2\pi i)^k}{(k-1)!} \ell^{k-1} e^{2\pi i \ell w}.$$

Plugging this in one step further, the Eisenstein series becomes

$$\begin{aligned} G_k(z) &= 2\left(\zeta(k) + \frac{(2\pi i)^k}{(k-1)!} \sum_{m=1}^{\infty}\sum_{\ell=1}^{\infty} \ell^{k-1} e^{2\pi i \ell m z}\right) \\ &= 2\left(\zeta(k) + \frac{(2\pi i)^k}{(k-1)!} \sum_{n=1}^{\infty} e^{2\pi i n z}\left(\sum_{n=\ell m} \ell^{k-1}\right)\right). \end{aligned}$$

Letting $\sigma_{k-1}(n) = \sum_{d\mid n} d^{k-1}$, we see that

$$G_k(z) = 2\left(\zeta(k) + \frac{(2\pi i)^k}{(k-1)!} \sum_{n=1}^{\infty} \sigma_{k-1}(n) e^{2\pi i n z}\right).$$

We know that $\zeta(k)$ is a rational number times π^k. So for example, in the case that $k = 4$, we have that

$$G_4(z) = 2\left(\frac{\pi^4}{90} + \frac{16\pi^4}{4} \sum \sigma_3(n) e^{2\pi i n z}\right).$$

Sometimes it is convenient to define this series without the $2\zeta(k)$ term, and we might write

$$\begin{aligned} E_4(z) &= \frac{G_4(z)}{2\zeta(4)} = 1 + 240 \sum_{n=1}^{\infty} \sigma_3(n) e^{2\pi i n z} \\ E_6(z) &= 1 - 504 \sum_{n=1}^{\infty} \sigma_5(n) q^n \\ E_8(z) &= 1 + 480 \sum_{n=1}^{\infty} \sigma_7(n) q^n \\ E_{12}(z) &= 1 + \frac{65520}{691} \sum_{n=1}^{\infty} \sigma_{11}(n) q^n. \end{aligned}$$

Here, the constants preceding each summation have a connection with Bernoulli numbers; there are $\frac{-2k}{B_k}$ where $\frac{x}{e^x-1} = \sum B_k \frac{x^k}{k!}$.

If f has weight k and g has weight ℓ, then fg is a modular form of weight $k+\ell$.

For example,

$$\frac{E_4^3 - E_6^2}{1728} = q + (\cdots)q^2 + \cdots = \Delta(z) = q\prod_{n=1}^{\infty}(1-q^n)^{24} = \sum_{n=1}^{\infty} \tau(n) q^n$$

is a modular form of weight 12. Here $\tau(1) = 1$, $\tau(2) = 24$, $\cdots$ is known as *Ramanujan's τ function*, and it is the beginning of a rich theory. In fact, $\tau(m)\tau(n) = \tau(mn)$ if $(m,n) = 1$, and $|\tau(n)| \ll n^{11/2+\varepsilon}$, and for prime p, we have $|\tau(p)| \le 2p^{11/2}$.

First, we prove that this object is a modular form. Recall that we considered

$$F(z) = \prod_{j=1}^{\infty}(1-z^j)^{-1}.$$

Define

$$\tilde{F}(z) = F(e^{2\pi i z}) = \prod (1-q^n)^{-1}.$$

Then previously, we had previously obtained a connection between $\tilde{F}(iy)$ and $\tilde{F}(i/y)$. This is exactly what the modularity condition is about: if $z \to -1/z$ then $iy \to i/y$. So the relation that we proved

$$\log F(e^{-2\pi/x}) - \log F(e^{-2\pi x}) = \frac{\pi}{12}\left(x - \frac{1}{x}\right) - \frac{1}{2}\log x$$

can be converted with a little bit of algebra; defining the *Dedekind eta function* $\eta(z) = q^{1/24}\prod(1-q^n)$, we get that $|\eta(iy)|y^{1/2} = |\eta(i/y)|$, so η is like a modular form of weight $1/2$. So we've shown that $\Delta(z) = \eta(z)^{24}$ and $\Delta(i/y) = (iy)^{12}\Delta(iy)$. If Δ is holomorphic, we claim that it must satisfy this relation for all values of z. To see this, observe that $\Delta(-1/z) = z^{12}\Delta(z) = 0$ on $i\mathbb{R}$, which implies that this is true for all $z \in \mathbb{H}$. So this implies that Δ is a modular form of weight 12.

There is one more feature that makes Δ more interesting than Eisenstein series. There is a cusp form of Δ that makes the constant Fourier coefficient equal to 0. This is a special feature of Δ.

These modular forms are sources of functions that are very similar to zeta and L-functions; these have nice functional equations. We saw the theta function, which can also be thought of in this framework. We can define

$$L(s,\Delta) = \sum_{n=1}^{\infty}\frac{\tau(n)}{n^s}.$$

This has an Euler product because the $\tau(n)$ are multiplicative. We also claim that this has a nice functional equation. Consider $(2\pi)^{-s}\Gamma(s)L(s,\Delta)$. Then since $\Delta(iy) = \sum \tau(n)e^{-2\pi n y}$, we have

$$\begin{aligned}(2\pi)^{-s}\Gamma(s)L(s,\Delta) &= \int_0^{\infty}\Delta(iy)y^s\frac{dy}{y} = \sum_{n=1}^{\infty}\tau(n)\int_0^{\infty}e^{-2\pi n y}y^s\frac{dy}{y} \\ &= \int_1^{\infty}\Delta(iy)y^s\frac{dy}{y} + \int_1^{\infty}\Delta(i/y)y^{-s}\frac{dy}{y} = \int_1^{\infty}\Delta(iy)(y^s + y^{12-s})\frac{dy}{y},\end{aligned}$$

which is now invariant under $s \leftrightarrow 12 - s$. There is even a version of the Riemann hypothesis for this.

Observe that

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

Replacing $s \to k-s$ makes this look like the functional equation for zeta, which gives another nice connection.

We've only scratched the surface of the theory of modular forms.

E-mail address: `moorxu@stanford.edu`