Analytic Number Theory and Combinatorics

Neil J. Calkin
(Notes by Shuai Wei)

September 18, 2023

Contents

Chapter 1

Introduction

Number Theory: algebraic properties of integers.
Analytic: tools from complex analysis, inequalities, issues of convergence and rearrangements of series.

Example 1.1. Let $S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots$. $1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \cdots$. $-\frac{1}{2} - \frac{1}{4} - \frac{1}{6} - \cdots$. $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{6} + \frac{1}{5} - \frac{1}{10} + \frac{1}{7} - \frac{1}{14} + \cdots$. $-\frac{1}{4} - \frac{1}{6} - \frac{1}{12} - \cdots$. $\frac{1}{2} + \frac{1}{6} + \frac{1}{10} + \frac{1}{14} + \cdots$. $-\frac{1}{4} - \frac{1}{8} - \frac{1}{12} - \cdots$. $\frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \frac{1}{14} + \cdots = \frac{1}{2}(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \cdots) = \frac{1}{2}S$.

Theorem 1.2. *Any conditionally but not absolutely convergent series of real numbers can be rearranged to converge to any number we like.*

Remark. Since $\sum_{p \text{ prime}} \frac{1}{p} = \infty$ and $\sum_{n=1}^{\infty} \frac{1}{n(\log n)^2} < \infty$, # primes $> \#\{n(\log n)^2 \mid n \in \mathbb{N}\}$. So $\#\{p < x\} > x \cdot \frac{x}{x(\log x)^2} = \frac{x}{(\log x)^2}$ since there are at least x primes from 1 to $x(\log x)^2$.

Definition 1.3. A system $a_1 \pmod{m_1}$, $a_2 \pmod{m_2}$, $\cdots$, $a_k \pmod{m_k}$ with $m_1 \leqslant m_2 \leqslant \cdots m_k$ and $0 \leqslant a_j < m_j$ is a *covering system of convergence* if for every $n \in \mathbb{Z}_{\geqslant 0}$, there is a j so that $n \equiv a_j \pmod{m_j}$. It is a *disjoint* covering system if every $n \in \mathbb{Z}_{\geqslant 0}$ is covered by exactly one congruence.

Remark. For any disjoint covering system of convergence, the two largest moduli must be equal.

Theorem 1.4. *Let* $a_1 \pmod{a_1}, a_2 \pmod{m_2}, \cdots, a_k \pmod{m_k}$ *with* $m_1 \leqslant m_2 \leqslant \cdots \leqslant m_k$ *be a disjoint covering system, then* $m_{k-1} = m_k$.

Proof. Consider

$$f_j(x) = x^{a_j} + x^{a_j+m_j} + x^{a_j+2m_j} + \cdots = \sum_{n \equiv a_j \pmod{m_j}} x^n, \forall j = 1, \ldots, k.$$

Since it is a disjoint covering system, $\frac{x^{a_1}}{1-x^{m_1}} + \frac{x^{a_2}}{1-x^{m_2}} + \cdots + \frac{x^{a_k}}{1-x^{m_k}} = f_1(x) + f_2(x) + \cdots + f_k(x) = \frac{1}{1-x}(*)$. Now let $w = e^{2\pi i/m_k}$, then $w^{m_k} = 1$. RHS is well-behaved at $x = w$, but LHS is not. If $m_{k-1} < m_k$, we have exactly one term $\frac{w^{a_k}}{0}$ and all other terms would be finite. Hence for (*) holds, we must have $m_{k-1} = m_k$. If $m_{k-2} < m_{k-1}$, then we must have $w^{a_{k-1}} = -w^{a_k}$ to cancel last two infinite terms. So $e^{2\pi i a_{k-1}/m_k} = -w^{2\pi i a_k/m_k}$. Hence m_k is even. In fact, if $m_1 < m_2 < \cdots < m_{k-1} = m_k$, then we must have $m_1 = 2$, $m_2 = 4$, $m_3 = 8$, $\cdots$, $m_{k-1} = m_k = 2^{k-2}$. ☐

Fact 1.5. Useful facts about $\mathbb{C}$. Let $0 \leqslant j \leqslant n-1$ with $n \in \mathbb{Z}_{\geqslant 2}$.

(a) Let $w = e^{2\pi i j/n}$, then $\sum_{k=0}^{n-1} w^k = \begin{cases} 0 & \text{if } j \neq 0 \\ n & \text{if } j = 0 \end{cases}$.

(b) If $k \in \mathbb{Z}$, $\int_0^1 e^{2\pi i k x} dx = \frac{1}{2\pi}\int_0^{2\pi} e^{ikx} dx = \begin{cases} 0 & \text{if } k \neq 0 \\ 1 & \text{if } k = 0 \end{cases}$.

Question: what happens if $k \notin \mathbb{Z}$. Nice fact: $a_n = \sum_{k=1}^n \sin(k)$. Exercise: a_k is bounded.

Recall 1.6. Let f and g be non-zero functions of x. $f(x) \sim g(x)$ as $x \to \infty$ means precisely that $\frac{f(x)-g(x)}{g(x)} \to 0$ as $x \to \infty$, it is equivalent to say $\frac{f(x)}{g(x)} \to 1$ as $x \to \infty$, i.e., $\frac{f(x)-g(x)}{f(x)} \to 0$ as $x \to \infty$.

$f(x) = o(g(x))$ means that $\frac{f(x)}{g(x)} \to 0$ as $x \to \infty$.

$f(x) = O(g(x))$ means there exists $c > 0$ and x_0 so that $|f(x)| < cg(x)$ for $x > x_0$.

Definition 1.7. Let $\pi(x)$ be the number of primes $\leqslant x$ for $x \in \mathbb{N}$.

Definition 1.8.

$$\mathrm{Li}(x) = \int_2^x \frac{1}{\log t} dt,$$

where log is $\log_e$.

Theorem 1.9 (Prime number theorem)**.**

$$\pi(x) \sim \mathrm{Li}(x).$$

Remark. Sometimes we see $\mathrm{Li}(x) = \int_0^x \frac{1}{\log t} dt = \lim_{\delta,\epsilon \to 0^+} \int_0^{1-\delta} \frac{1}{\log t} dt + \int_{1+\epsilon}^x \frac{1}{\log t} dt$.

Theorem 1.10.

$$\pi(x) \sim \frac{x}{\log x}.$$

But $\mathrm{Li}(x)$ *is a better approximation to* $\pi(x)$ *than* $\frac{x}{\log x}$ *is.*

Proof. Let $\epsilon > 0$ be small. Then if $x > t > x^{1-\epsilon}$, then $\log x > \log t > (1-\epsilon)\log x$, i.e., $\frac{1}{\log x} < \frac{1}{\log t} < \frac{1}{(1-\epsilon)\log x}$. So $\int_{x^{1-\epsilon}}^x \frac{1}{\log t} dt < \frac{1}{1-\epsilon}\frac{x - x^{1-\epsilon}}{\log x} < \frac{1}{1-\epsilon}\frac{x}{\log x}$ and $\mathrm{Li}(x) = \int_0^x \frac{1}{\log t} dt > \frac{x}{\log x}$. This means that $\mathrm{Li}(x) \sim \frac{x}{\log x}$. So prime number theorem implies $\pi(x) \sim \frac{x}{\log x}$. □

Lemma 1.11 (RH = Riemann Hypothesis (approximately))**.** For any $\epsilon > 0$,

$$|\pi(x) - \mathrm{Li}(x)| = O\big(x^{\frac{1}{2}+\epsilon}\big).$$

Fact 1.12. More careful analysis shows that $\mathrm{Li}(x) = \frac{x}{\log x} + \frac{x}{(\log x)^2} + O\big(\frac{x}{(\log x)^3}\big)$. So if RH is true,

$$\left|\pi(x) - \frac{x}{\log x}\right| \sim \frac{x}{(\log x)^2}.$$

Question 1.13. If we average a nice number theoretic function, how does it behave?

Definition 1.14.

$$\mathbb{1}_p(n) = \begin{cases} 1 & \text{if } n \text{ is prime} \\ 0 & \text{if } n \text{ is not prime} \end{cases}.$$

Fact 1.15.
$$\pi(x) = \sum_{n \leqslant x} \mathbb{1}_p(n).$$

Definition 1.16. Let
$$\tau(n) = \# \text{ of divisor of } n = \sum_{d|n} 1.$$

Question 1.17. Pick x uniformly in $\{1, \ldots, N\}$, what are $E[\tau(X)]$ and $\mathrm{Var}(\tau(X))$. Fix $x > 0$, how many integer lattice points in a circle of radius x?

Theorem 1.18. *There are infinitely many primes.*

Proof. Method 1. We construct an infinite list $p_1, p_2, \ldots, p_i, \cdots$ of primes. First $p_1 = 2$ is prime. Now given $p_1, p_2, \ldots, p_k$, let p_{k+1} be the least prime factor of $p_1 p_2 \cdots p_k + 1$, so $p_2 = 3$, $p_3 = 7$, $p_4 = 43$. Then $p_1, p_2, \ldots, p_k, \cdots$ is an infinite list of primes. Question: Which primes are in/not in this list? Is 5 in the list? Heuristically speaking, how dense do we expect this list to be relative to the list of all primes?

Method 2. Claim. There are infinitely many primes congruent to 3 (mod 4). If $n = 4k+3$, then at least one of its prime factors is congruent to 3 (mod 4). Hence if $p_1 = 3, p_2 = 7, \ldots, p_k$ is a list of primes congruent to 3 (mod 4), then $4p_1p_2 \cdots p_k - 1$ is congruent to 3 (mod 4) and is not divisible by $p_1, p_2, \ldots, p_k$. Hence it has a prime factor $p_{k+1} \equiv 3 \pmod 4$. So we construct $p_1, \ldots, p_k, \cdots$ an infinite list.

Method 3(Euler). **Lemma**. Let $H_n = \sum_{k=1}^{n} \frac{1}{k}$. Then $H_n \sim \log n$ as $n \to \infty$. Let $s > 1$, and consider $\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}$. Then for all $\mathrm{Re}(s) > 1$, $\zeta(s)$ converges and since $H_n \to \infty$, $\zeta(s)$ diverges at $s = 1$, hence $\lim_{s \to 1^+} \zeta(s) = \infty$. But if $s > 1$, $\zeta(s) = \prod_p (1 - \frac{1}{p^s})^{-1}$. So if there are only finitely many primes, then the RHS would be bounded at 1, but $\lim_{s\to 1^+} \zeta(s) \to \infty$, a contradiction. Question: Can we use the rate at which H_n approaches ∞ to estimate $\pi(x)$?

Method 4. Define $F_n = 2^{2^n} + 1$, the n^{th} Fermat number, for $n \geqslant 0$. Claim. (Fermat) F_n is prime for all n. FALSE. ABOUT AS FALSE AS IS POSSIBLE. Probable fact: F_n is compositve if $n \geqslant 5$?6? **Lemma.** If $m \neq n$, then $\gcd(F_m, F_n) = 1$, and

$$\begin{aligned} F_{n+1} - 2 &= 2^{2^{n+1}} - 1 = (2^{2^n})^2 - 1 = (2^{2^n} + 1)(2^{2^n} - 1) = F_n(F_n - 2) = F_n F_{n-1}(F_{n-1} - 2) \\ &= F_n F_{n-1} \cdots F_3 F_2 F_1 F_0 (F_0 - 2) = F_n F_{n-1} \cdots F_1 F_0 \end{aligned}$$

i.e., $F_{n+1} = F_0 F_1 \cdots F_n + 2$. Since F_n is odd for all $n \geqslant 0$, $\gcd(F_n, F_{n+1}) = 1$. By induction, $\gcd(F_m, F_n) = 1$ for $m \neq n$. Let p_n = least prime factor of F_n. Then $p_0, p_1, p_2, \ldots, p_n, \ldots,$ is an infinite list of primes.

What more can we say about prime factors of F_n? Let $p \mid F_n$, $2^{2^n} = F_n - 1 \equiv -1 \pmod p$ and $2^{2^{n+1}} = (2^{2^n})^2 \equiv 1 \pmod p$. Since $2^n \mid 2^{n+1}$, the order of the residue class of 2 is 2^{n+1}. Also, since $|(\mathbb{Z}/p\mathbb{Z})^\times| = p - 1$, $2^{n+1} \mid p - 1$. So $p = 2^{n+1}k + 1$ for some $k \geqslant 1$. □

Corollary 1.19. If $q_1, q_2, \cdots$ is a list of all primes with $q_n \leqslant p_n$ for $n \in \mathbb{N}$, then $q_n \leqslant F_n$ for $n \in \mathbb{N}$. So $\pi(x) > \log_2 \log_2 x$.

Exercise 1.20. Proof Method 1-2 give a lower bound too. So does defining p_n = least prime factor of $(n! + 1)$.

Many of the things we discuss will be "average theorem". e.g., if we pick a random integer $n \leqslant x$,

(a) What is the probability it is prime?

(b) What is the distribution of the value of a nice function of n?

Recall $\pi(x) = \sum_{n\leqslant x} \mathbb{1}_p(n)$. $\mathbb{1}_p(n)$ is a complicated function but it is computable in time polynomial in $\log(n)$. (Agrawal et al, AKS, 2002) But thanks to the Prime Number Theory, we know quite a lot about the "smoothed" function $\pi(x)$. So long as we don't ask for too preicise an estimate for $\pi(x)$, we can approximate it.

Since we know $\pi(x) \sim \mathrm{Li}(x)$ or $x/\log(x)$, can we use this fact to estimate other "smoothed functions"? For example,

$$\sum_{n\leqslant x} \mathbb{1}_p(n)\log n = \sum_{p\leqslant x} \log p,\ \sum_{n\leqslant x} \mathbb{1}_p(n)\frac{1}{n} = \sum_{p\leqslant x}\frac{1}{p},\ \sum_{n\leqslant x}\mathbb{1}_p(n)n = \sum_{p\leqslant x} p.$$

Roughly speaking (Heuristic), PNT says is that for a large number N,

$$P(x \text{ is prime}) \approx \left(\text{the average of } \{\mathbb{1}_p(n) \mid n \leqslant x\} = \frac{1}{x}\sum_{n\leqslant x}\mathbb{1}_p(n) = \right)\frac{\pi(x)}{x} \approx \frac{1}{\log x}.$$

Note

$$E\left[\sum_{p\leqslant x} f(p)\right] = E\left[\sum_{n\leqslant x}\mathbb{1}_p(n)f(n)\right] = \sum_{n\leqslant x} E[\mathbb{1}_p(n)]f(n) = \sum_{n\leqslant x} f(n)P(n \text{ prime}) \approx \sum_{2\leqslant n\leqslant x} f(n)\frac{1}{\log n}.$$

So $\sum_{p\leqslant x} f(p)$ should behave like $\sum_{2\leqslant n\leqslant x} f(n)\frac{1}{\log n}$. Similarly, $\sum_{p\leqslant x}\log p$ should behave like $\sum_{2\leqslant n\leqslant x}\frac{\log n}{\log n} = x - 1 \sim x$. So we expect $\prod_{p\leqslant x} p \approx e^x$. More precisely, we expect

(a) for every $c > 1$, $\prod_{p\leqslant x} p = O(e^{cx})$,

(b) for every $c < 1$, $e^{cx} = O(\prod_{p\leqslant x} p)$.

Theorem 1.21 (Abel summation)**.** *Suppose $\{a_k\}_{k\geqslant 1}$ and $\{b_k\}_{k\geqslant 1}$ are sequences (here a_k will be weired but have a reasonable "smooth" estimate, and b_k will be a nice smooth function of k). Suppose we know $A(x) = \sum_{k\leqslant x} a_k$ (to some precision): $\sum_{k\leqslant x} a_k b_k = A(x)b_x + \sum_{k\leqslant x-1} A(k)(b_k - b_{k+1})$.*

Here if we are lucky, $b_k - b_{k+1}$ behaves nicely, and is small enough that we can bound the sum and control errors in our approximation of $A(k)$.

Proof.

$$\begin{aligned}\sum_{k\leqslant x} a_k b_k &= \sum_{k\leqslant x}(A(k) - A(k-1))b_k = \sum_{k\leqslant x} A(k)b_k - \sum_{k\leqslant x} A(k-1)b_k \\ &= \sum_{k\leqslant x} A(k)b_k - \sum_{k\leqslant x-1} A(k)b_{k+1} = A(x)b_x + \sum_{k\leqslant x-1} A(k)(b_k - b_{k+1}).\end{aligned}$$ □

Example 1.22. Let $a_k = \mathbb{1}_p(k)$ for $k \in \mathbb{N}$. Then $A(x) = \sum_{k=1}^{x} a_k = \pi(x) \sim \frac{x}{\log x}$. Let $b_k = \log k$ for $k \in \mathbb{N}$. Then $b_k - b_{k+1} = \log k - \log(k+1) = \log(\frac{k}{k+1}) = -\log(\frac{k+1}{k}) = -\log(1 + \frac{1}{k})$. If $\epsilon < 1$,

$\log(1+\epsilon) = \epsilon - \frac{\epsilon^2}{2} + \frac{\epsilon^3}{3} - \frac{\epsilon^4}{4} + \cdots$. This is an alternating series, so $\epsilon - \frac{\epsilon^2}{2} < \log(1+\epsilon) < \epsilon$. So $b_k - b_{k+1} = -\frac{1}{k} + \frac{1}{2k^2} + O(\frac{1}{k^3}) = -\frac{1}{k} + O(\frac{1}{k^2})$ for $k \geqslant 2$. Hence

$$\sum_{p\leqslant x} \log p = \sum_{k\leqslant x} \mathbb{1}_p(k)\log k = \sum_{k\leqslant x} a_k b_k = A(x)b_x + \sum_{k\leqslant x-1} A(k)(b_k - b_{k+1})$$

$$\sim \pi(x)\log(x) + \sum_{k\leqslant x-1} \pi(k)\Big(-\frac{1}{k} + O\Big(\frac{1}{k^2}\Big)\Big) \approx x - \sum_{k\leqslant x-1}\frac{1}{\log k} + O\left(\sum_{k\leqslant x-1}\frac{1}{k\log k}\right).$$

Example 1.23. Estimate $\sum_{k=2}^n \frac{1}{\log k}$. How to estimate this? Various techniques: We will see Euler-Maclourin, a very general method for estimating $\sum_{k\leqslant n} f(k)$ when f is a nice smooth function in terms of $\int_a^n f(x)dx$.

What can we do with Abel summation? How big is $n!$? (Stirling approximation: $n! \sim (\frac{n}{e})^n\sqrt{2\pi n}$) We'll estimate

$$\log(n!) = n\log n - n + \frac{1}{2}\log n + \log\sqrt{2\pi} + o(1) = n\log n - n + \frac{1}{2}\log n + O(1).$$

Let $a_k = 1$, $b_k = \log k$ for $k \geqslant 1$, $A_n = \sum_{k=1}^n a_k = n$ and $b_k - b_{k-1} = -\log(1+\frac{1}{k})$ for $k = 2, \dots, n$. Then

$$\log(n!) = \sum_{k=1}^n \log k = \sum_{k=1}^n a_k b_k = A_n b_n + \sum_{k=1}^{n-1} A_k(b_k - b_{k-1}) = n\log n - \sum_{k=1}^{n-1} k\log(1+\frac{1}{k})$$

$$= n\log n - \log 2 - \sum_{k=2}^{n-1} k\log(1+\frac{1}{k}) = n\log n - \log 2 - \sum_{k=2}^{n-1}\sum_{j=1}^{\infty} k\frac{(-1)^{j-1}}{jk^j}$$

$$= n\log n - \log 2 - \sum_{k=2}^{n-1}\frac{k}{k} + \sum_{k=2}^{n-1}\frac{k}{2k^2} - \sum_{k=2}^{n-1}\sum_{j=3}^{\infty} k\frac{(-1)^{j-1}}{jk^j}$$

$$= n\log n - \log 2 - (n-2) + \frac{1}{2}\left(\sum_{k=1}^n \frac{1}{k} - 1 - \frac{1}{n}\right) - \sum_{k=2}^{n-1}\sum_{j=3}^{\infty} k\frac{(-1)^{j-1}}{jk^j}$$

$$= n\log n - n + \frac{1}{2}\log n - \log 2 + \frac{3}{2} + \frac{1}{2}\gamma - \sum_{k=2}^{n-1}\sum_{j=3}^{\infty} k\frac{(-1)^{j-1}}{jk^j} + o(1)$$

$$= n\log n - n + \frac{1}{2}\log n + C + o(1),$$

where $\gamma = \lim_{n\to\infty}(-\log n + \sum_{k=1}^n \frac{1}{k}) = 0.577$ is the Euler-Mascheroni costant and

$$C = -\log 2 + \frac{3}{2} + \frac{1}{2}\gamma - \sum_{k=2}^{n-1}\sum_{j=3}^{\infty} k\frac{(-1)^{j-1}}{jk^j} = \log\sqrt{2\pi}.$$

Hence $n! \sim (\frac{n}{e})^n\sqrt{n}C$ (*).

Aside: to show $C = \sqrt{2\pi}$: use $2^{2n} = \sum_{k=-n}^n \binom{2n}{n+k}$. Use $*$ to estimate $\binom{2n}{n}$. Now approximate $\frac{\binom{2n}{n+k}}{\binom{2n}{n}}$ for $k = o(n)$. After easy but somewhat tedious calculation, we see $C = \sqrt{2\pi}$.

Euler-Maclourin summation give not just the same result but a different formula for $\log C$. plus an explicit asymptotic formula for subsequenct terms in $\frac{1}{n}, \frac{1}{n^2}$.

Exercise 1.24. (a) If $|x| < \frac{1}{2}$, give explicit bounds in

(1) $|-\log(1-x) - x| < cx^2$

i. for $x \in (0, \frac{1}{2})$;

ii. for $x \in (-\frac{1}{2}, 0)$.

Plot to compare reality.

(2) $\left|-\log(1-x) - x - \frac{x^2}{2}\right| < cx^3$

i. for $x \in (0, \frac{1}{2})$;

ii. for $x \in (-\frac{1}{2}, 0)$.

(3) $\left|\cdots - x - \frac{x^2}{2} - \frac{x^3}{3}\right|$

i. for $x \in (0, \frac{1}{2})$;

ii. for $x \in (-\frac{1}{2}, 0)$.

(b) Consider $\sum_{k=2}^{n} \frac{1}{\log k}$. Take $a_k = 1$, $b_k = \frac{1}{\log k}$.

(1) Given careful estimates for $b_k - b_{k-1}$.

(2) Use abel summation to estimate $\sum_{k=2}^{n} \frac{1}{\log k}$.

(c) Can you improve this? Compare to doing integrating by parts twice. You may wish to consider something like $\sum (k-1)(\frac{1}{\log k} - \frac{1}{\log(k+1)})$.

Side note. $\int_2^x \frac{1}{\log t} dt \sim x \sum_{k=0}^{\infty} \frac{k!}{(\log x)^{k+1}}$ as an asymptotic series, meaning that for any m, as $x \to \infty$, $\left|\int_2^x \frac{1}{\log t} dt - x \sum_{k=0}^{m} \frac{k!}{(\log x)^{k+1}}\right| = O(\frac{x(m+1)!}{(\log x)^{m+2}})$. In particular, for any x, there is a "best" m to use to approximate $\text{Li}(x)$, i.e., to minimize $\left|\int_2^x \frac{1}{\log t} dt - x \sum_{k=0}^{m} \frac{k!}{(\log x)^{k+1}}\right|$.

(d) (1) For $x = 10^2, 10^3, 10^4, 10^5, 10^6, 10^7, (...?)$, find m to minimize $\left|\text{Li}(x) - \sum_{k=0}^{m} \frac{xk!}{(\log x)^{k+1}}\right|$. Once the sum exceeds $\text{Li}(x)$, if we take one more term, we've gone too far. So we know when to stop. Question: how often is the best approximation when $\sum < \text{Li}(x)$ and how often when $\text{Li}(x) < \sum$?

(2) How big is m as a function of x when $\left|\int - \sum\right|$ is minimized?

(3) How big is $\frac{xk!}{(\log x)^{k+1}}$ in this region. How good is the approximation? (If $\log x$ is big, say 10^6, then the first few terms of the sum start out very well, decrease very rapidly, then finally start increasing and diverge to ∞ very rapidly. How does this relate to the optimal choice for m?). The graph of $x \sum_{k=0}^{m} \frac{k!}{(\log x)^{k+1}}$ is the following:

Replace $k!$ by the Γ to obtain a continuous of y: $F(m) = \int_0^{m+1} \frac{\Gamma(y+1)}{(\log x)^{y+1}} dy$. Plot $F(m)$.

Chapter 2

Arithmetic, multiplicative and completely multiplicative functions

Let $\mathbb{N} = \{1, 2, 3, \cdots\}$.

Definition 2.1. Define the *Möbius function* $\mu : \mathbb{N} \to \{0, \pm 1\}$ by

$$\mu(n) = \begin{cases} 0 & \text{if } n \text{ is divisible by the square of a prime} \\ (-1)^k & k \text{ is the number of prime factors of } n \end{cases}.$$

Remark. The Möbius function is fundamental to the theory of multiplicative functions.

2.1 Möbius inversion

Theorem 2.2 (Möbius 1832)**.** *Assume $F : [1, t] \to \mathbb{C}$. For all $x \leqslant t$, define $G(x) = \sum_{n \leqslant x} F(\frac{x}{n})$, then $F(x) = \sum_{m \leqslant x} \mu(m) G(\frac{x}{m})$.*

Theorem 2.3 (Dirichlet,1857)**.** *If arithmetic functions g and f satisfying $g(n) = \sum_{d|n} f(d) = \sum_{d|n} f(\frac{n}{d})$ for $n \geqslant 1$, then $f(n) = \sum_{d|n} \mu(d) g(\frac{n}{d})$ for $n \geqslant 1$.*

Proof. Let $F(x) = 0$ whenever $x \notin \mathbb{N}$ and $F(n) = f(n)$ when $n \in \mathbb{N}$. Set $g(n) = G(n)$. □

Remark. In effect, the original $f(n)$ can be determined given $g(n)$ by using the inversion formula. The two sequences are said to be *Möbius transforms* of each other.

Compare this to A_n, a_n in Abel summation: If $A_n = \sum_{k \leqslant n} a_k$, then $a_n = A_n - A_{n-1}$.

For nice ordered sets, we can define sums over interesting sets, and invert the definitions.

Theorem 2.4 (Dedekind)**.** *For $n \in \mathbb{N}$, $\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1 \end{cases}$.*

Proof. Let $f(1) = 1$ and $f(n) = 0$ for $n > 1$. Then $g(n) = \sum_{d|n} f(\frac{n}{d}) = 1$ for $n \in \mathbb{N}$. Then $f(n) = \sum_{d|n} \mu(d)$. □

Exercise 2.5. Prove Dedekind implies Möbius.

Example 2.6. Let $n = 6$ and $d = 1, 2, 3, 6$. Then $\mu(1) + \mu(2) + \mu(3) + \mu(6) = 1 - 1 - 1 + 1 = 0$.

2.2 Multiplicative and completely multiplicative functions

Example 2.7. $\varphi(n)$ is a multiplicative function.

Proof. By CRT,

$$\begin{aligned} \mathbb{Z}/(m\mathbb{Z}\cap n\mathbb{Z}) &\cong \mathbb{Z}/m\mathbb{Z}\times\mathbb{Z}/n\mathbb{Z} \\ \overline{k} &\mapsto (\overline{k},\overline{k}) \\ \overline{bm_n^{-1}m+an_m^{-1}n} &\leftarrow\!\shortmid (\overline{a},\overline{b}), \end{aligned}$$

where m_n^{-1} is the inverse of m modulo n and n_m^{-1} is the inverse of n modulo m. If $\gcd(m,n)=1$, then $m\mathbb{Z}\cap n\mathbb{Z}=mn\mathbb{Z}$, given $0\leqslant a<m$ and $0\leqslant b<n$, there is a unique $0\leqslant k<mn$ so that $k\equiv a \pmod m$ and $k\equiv b \pmod n$, so $\varphi(m)\varphi(n)=\varphi(mn)$. We actually have $(\mathbb{Z}/mn\mathbb{Z})^\times\cong (Z/m\mathbb{Z})^\times\times(\mathbb{Z}/n\mathbb{Z})^\times$ since we have the units $(R\times S)^\times=R^\times\times S^\times$. ☐

Remark. The RSA encryption algorithm relies on the difficulty of factoring $n=pq$, where p,q are large distinct primes. If we know $m=\varphi(pq)$ as well as $n=pq$, $m:=\varphi(pq)=(p-1)(q-1)=n-(p+q)+1$, i.e., $p+\frac{n}{p}=n-m+1$, gives a quadratic in p, which can be easily solved. Hence knowing p,q if and only if knowing $pq,(p-1)(q-1)$.

Definition 2.8. Let f,g be multiplicative functions. Define the *Convolution* of f,g by

$$f*g(n)=\sum_{d\mid n}f(d)g\Big(\frac{n}{d}\Big)=\sum_{d_1d_2=n}f(d_1)g(d_2),\forall n\in\mathbb{N}.$$

Definition 2.9. Let $f_1,\ldots,f_k$ be multiplicative functions. Define the *n-fold Convolution* of $f_1,\ldots,f_n$ by

$$(f_1*\cdots*f_k)(n)=\sum_{d_1\cdots d_k=n}f_1(d_1)\cdots f_k(d_k),\forall n\in\mathbb{N}.$$

Remark. k-fold convolution must be associative.

Throught experiment, is there a correponding idea for $F:[1,t]\to\mathbb{C}$ and $G:[1,t]\to\mathbb{C}$, looking at functions of real intervals? Does it give any thing intersesting?

Theorem 2.10. *If f,g are multiplicative functions, then so is $f*g$. This gives us an algebra on the set of multiplicative functions.*

Proof. Let $m,n\in\mathbb{N}$ such that $\gcd(m,n)=1$. If $d\mid mn$, then $d=d_1d_2$ with $d_1\mid m$ and $d_2\mid n$. Note $\gcd(d_1,d_2)=1=\gcd(\frac{m}{d_1},\frac{n}{d_2})$. Since f,g are multiplicative function,

$$\begin{aligned}\sum_{d\mid mn}f(d)g\Big(\frac{mn}{d}\Big)&=\sum_{d_1\mid m}\sum_{d_2\mid n}f(d_1d_2)g\Big(\frac{mn}{d_1d_2}\Big)=\sum_{d_1\mid m}\sum_{d_2\mid n}f(d_1)f(d_2)g\Big(\frac{m}{d_1}\Big)g\Big(\frac{n}{d_2}\Big)\\&=\sum_{d_1\mid m}f(d_1)g\Big(\frac{m}{d_1}\Big)\sum_{d_2\mid n}f(d_2)g\Big(\frac{n}{d_2}\Big)=f*g(m)\cdot f*g(n).\end{aligned}$$

☐

Exercise 2.11. Prove or give a counterexmaple to each of the following.

(a) If f,g are completely multiplicative, then $f*g$ is multiplicative.

(b) If f, g are multiplicative, then $f * g$ is multiplicative.

(c) If f, g are completely multiplicative, then so if $f * g$.

(d) If one of f, g are completely multiplicative and the other is multiplicative, then $f*g$ is completely multiplicative.

Corollary 2.12. The set of multiplicative functions form a group under $*$. The identity is $\mathbb{1}$, where $\mathbb{1}(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1 \end{cases}$.

Question 2.13. What is the inverse of f under $*$? (Hint: Möbius/Dedekind). So we get lots of new functions.

Example 2.14. $\tau(n) = \#\text{divisors of } n = \sum_{d|n} 1 = \sum_{d|n} \mathbb{1}(d)\mathbb{1}(\frac{n}{d}) = \mathbb{1} * \mathbb{1}(n)$.

Chapter 3

Dirichlet Series

Definition 3.1. Given an arithmetic function $a : \mathbb{N} \to \mathbb{C}$, we define the *associated Dirichlet series*

$$A(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}, \ \ s \in \mathbb{C}.$$

Remark. This is defined formally for a variable s, but if we want to evaluate it at a given s, we need to worry about convergence.

Some Dirichlet series don't converge for any $s \in \mathbb{C}$: for example $\sum_{n=1}^{\infty} \frac{n!}{n^s}$ diverges no matter which $s \in \mathbb{C}$, we use for fixed s, $\frac{n!}{n^s} \to \infty$ as $n \to \infty$. However, suppose we have a series $\sum_{n=1}^{\infty} \frac{a_n}{n^s}$ and it converges at $s_0 \in \mathbb{C}$. Then $\left|\frac{a_n}{n^{s_0}}\right| \to 0$ as $n \to \infty$. Write $s_0 = \sigma_0 + i\tau_0$ with $\sigma_0, \tau_0 \in \mathbb{R}$, then

$$\left|\frac{a_n}{n^{s_0}}\right| = \left|\frac{a_n}{n^{\sigma_0} n^{i\tau_0}}\right| = \left|\frac{a_n}{n^{\sigma_0}}\right| \cdot \left|e^{-i\tau_0 \log n}\right| = \left|\frac{a_n}{n^{\sigma_0}}\right|.$$

Hence if $s = \sigma + it$ with $\sigma > \sigma_0$, then it is easy to see $\sum_{n=1}^{\infty}\left|\frac{a_n}{n^s}\right|$ converges absolutely by Direct comparison series test.

Example 3.2. Particular Dirichlet series of immense importance: introduced by Euler for integer $s \geqslant 2$ and considered for complex s by Riemann using the *Riemann zeta-function* $\zeta(s)$.

If $s \in \mathbb{C}$ with $\mathrm{Re}(s) = \sigma > 1$, we define $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$. Observe this is a valid definition for $s = \sigma + it$ with $\sigma > 1$, since $\sum_{n=1}^{\infty}\left|\frac{1}{n^s}\right| = \sum_{n=1}^{\infty} \frac{1}{n^{\sigma}}$ converges, i.e., $\sum_{n=1}^{\infty} \frac{1}{n^s}$ converges absolutely.

It diverges to ∞ at $s = 1$ since $\sum_{n=1}^{m} \frac{1}{n} \approx \int_1^m \frac{1}{n} dn = \log m \to \infty$ as $m \to \infty$. Hence $\zeta(\sigma) \to \infty$ as $\sigma \to 1^+$.

Next, let $s = 1 + it$ with $t \in \mathbb{R} \smallsetminus \{0\}$. Let $S_n = \sum_{k=1}^{n} \frac{1}{k^s}$. For, $\{x_n\} \subseteq \mathbb{C}$, $\sum_{n=1}^{\infty} x_n$ converges if and only if for $\epsilon > 0$, there exists $\ell \in \mathbb{C}$ and $N \in \mathbb{N}$ such that $|S_n - \ell| \leqslant \epsilon$ for any $n > \mathbb{N}$ if and only if there exists $N \in \mathbb{N}$ such that $\left|\sum_{k=m+1}^{n} x_n\right| = |S_n - S_m| < \epsilon$ for $n \geqslant m > N$. We'll show the series doesn't converge. Note $\frac{1}{n^s} = \frac{1}{n^{1+it}} = \frac{1}{n} e^{-it \log n} = \frac{1}{n} \cos(t \log n) - \frac{i}{n} \sin(t \log n)$. Consider just $\mathrm{Re}(\frac{1}{n^{1+it}}) = \frac{1}{n} \cos(t \log n)$. Let $2\pi r - \frac{\pi}{4} \leqslant t \log n \leqslant 2\pi r + \frac{\pi}{4}$, i.e., $\frac{2\pi r}{t} - \frac{\pi}{4t} \leqslant \log n \leqslant \frac{2\pi r}{t} + \frac{\pi}{4t}$, i.e., $\exp(\frac{2\pi r}{t} - \frac{\pi}{4t}) \leqslant n \leqslant \exp(\frac{2\pi r}{t} + \frac{\pi}{4t})$. Then $\cos(t \log n) \geqslant \frac{1}{\sqrt{2}}$ and

$$\sum_{n=\exp(\frac{2\pi r}{t} - \frac{\pi}{4t})}^{\exp(\frac{2\pi r}{t} + \frac{\pi}{4t})} \left|\frac{1}{n^{1+it}}\right| \geqslant \frac{1}{\sqrt{2} n_0} \exp\Big(\frac{2\pi r}{4t}\Big)\Big(\exp\Big(\frac{\pi}{4t}\Big) - \exp\Big(-\frac{\pi}{4t}\Big)\Big),$$

where $n_0 = \exp(\frac{2\pi r}{4t} - \frac{\pi}{4t})$. So we get this partial sum is bounded below by a constant depending only on t.

Hence $\sum_{n=1}^{\infty} \frac{1}{n^{1+it}}$ fails to converge for all $t \in \mathbb{R}$.

Theorem 3.3. *Assume $F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$ and $G(s) = \sum_{m=1}^{\infty} \frac{g(m)}{m^s}$ converge absolutely at s. Then $F(s)G(s) = \sum_{n=1}^{\infty} \frac{f*g(n)}{n^s}$ and it converges absolutely at s too.*

Proof. Since we'll be concerned with absolute convergence, we can assume s is real and $f(n), g(m) \geqslant 0$ for $m, n \in \mathbb{N}$. Since they converge absolutely, we can rearrange at will. So

$$\begin{aligned} F(s)G(s) &= \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \sum_{m=1}^{\infty} \frac{g(m)}{m^s} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{f(n)}{n^s} \frac{f(m)}{m^s} = \sum_{r=1}^{\infty} \sum_{mn=r} \frac{f(n)g(m)}{r^s} \\ &= \sum_{r=1}^{\infty} \frac{1}{r^s} \sum_{mn=r} f(n)g(m) = \sum_{r=1}^{\infty} \frac{1}{r^s} \sum_{n|r} f(n) g\Big(\frac{r}{n}\Big) = \sum_{r=1}^{\infty} \frac{f*g(r)}{r^s} = (F*G)(s). \end{aligned}$$ □

Corollary 3.4. Extend in a natural way, we get

$$(F_1 * \cdots * F_k)(s) = F_1(s) \cdots F_k(s).$$

Theorem 3.5. *Let p be fixed. For $n \in \mathbb{N}$, let $f_p(n) = \begin{cases} a_{p^k} & \text{if } n = p^k \text{ is a power of } p \\ 0 & \text{if } n \text{ is not a power of } p \\ a_1 = 1 & \text{otherwise} \end{cases}$. Then $F_p(s) = \sum_{n=1}^{\infty} \frac{f_p(n)}{n^s} = \sum_{k=1}^{\infty} \frac{f_p(p^k)}{(p^k)^s} = \sum_{k=1}^{\infty} \frac{a_{p^k}}{(p^k)^s}$ and*

$$\prod_{p<x} F_p(s) = \prod_{p<x} \sum_{n=1}^{\infty} \frac{f_p(n)}{n^s} = \sum_{n=1}^{\infty} \prod_{p<x} \frac{f_p(n)}{n^s} = \sum_{n=1}^{\infty} \frac{1}{n^s} \prod_{p<x} f_p(n) = \sum_{n=1}^{\infty} \frac{1}{n^s} \prod_{p<x, p^k || n} a_{p^k}.$$

So

$$\prod_{p} F_p(s) = \lim_{x\to\infty} \prod_{p<x} F_p(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s},$$

where $f(n) = \prod_{p^k || n} a_{p^k}$, i.e., f is the multiplicative function defined by the sequences a_{p^k}, $k = 0, 1, 2, \cdots$ and $p = 2, 3, 5, 7, \cdots$.

Example 3.6. Let $a_{p^k} = 1$ for any prime p and $k \in \mathbb{N}$. Then $F_p(s) = \sum_{k=1}^{\infty} \frac{1}{p^{ks}} = (1 - \frac{1}{p^s})^{-1}$ and $f(n) = 1$ for $n \geqslant 1$. So

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p} F_p(s) = \prod_{p} \Big(1 - \frac{1}{p^s}\Big)^{-1}.$$

Note in this example, $F_p(s) = 1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots$ is absolutely convergent for $\mathrm{Re}(s) > 0$, but $\zeta(s)$ is only absolutely convergent for $\mathrm{Re}(s) > 1$.

Remark. Can you construct sequences a_{p^k}, $k = 0, 1, 2, \cdots$ and $p = 2, 3, 5, 7, \cdots$ so that each $F_p(s)$ is absolutely convergent for $\mathrm{Re}(s) > 0$, but

(a) $F(s)$ diverges at $s = 2$, $s = 3$, $s = n$?

(b) $F(s)$ is never convergent.

The argument that a_n is fixed once x exceeds doesn't fail: it is only convergence which fails. Hence if we define $f(n) = \prod_{p^k || n} a_{p^k}$, then if $F(s)$ converges absolutely, $F(s) := \prod_p F_p(s)$.

Note $((1-\frac{1}{p^s})^{-1})^{-1} = 1 - \frac{1}{p^s} = 1 - \frac{1}{p^s} + \frac{0}{p^{2s}} + \frac{0}{p^{3s}} + \cdots$. Set $a_{p^k} = \begin{cases} 1 & \text{if } k = 0 \\ -1 & \text{if } k = 1 \\ 0 & \text{if } k > 1 \end{cases}$. Then

$$f(n) = \prod_{p^k || n} a_{p^k} = \begin{cases} 0 & \text{if } n \text{ is divisible by the square of a prime} \\ (-1)^j & j \text{ is the number of prime factors of } n \end{cases} = \mu(n).$$

So provided $\frac{1}{\zeta(s)}$ converges,

$$\frac{1}{\zeta(s)} = \prod_p \Big(1 - \frac{1}{p^s}\Big) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}.$$

Observe the Dirichlet series of $\mathbb{1}$ is 1.

Theorem 3.7. $1 * \mu = \mathbb{1}$, *where* $1(n) = 1$ *for* $n \in \mathbb{N}$.

Proof. Since $1 = \zeta(s)\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{1}{n^s} \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \sum_{n=1}^{\infty} \frac{1*\mu(n)}{n^s}$, we have $1 * \mu = \mathbb{1}$. □

Corollary 3.8. $\sum_{d|n} \mu(d) = \mathbb{1}(n)$.

Remark. $\frac{1}{\zeta(s)}$ is absolutely convergent at $\mathrm{Re}(s) > 1$ since $\sum_{n=1}^{\infty} \left| \frac{\mu(s)}{n^s} \right| = \sum_{n \text{ square free}} \frac{1}{n^{\mathrm{Re}(s)}} < \sum_{n=1}^{\infty} \frac{1}{n^{\mathrm{Re}(s)}}$, which is absolutely convergent if $\mathrm{Re}(s) > 1$.

Question 3.9. What happens when $s = 1$?

Exercise 3.10. Show by counting that $\sum_{d|n} \varphi(d) = n$.

Since $n = \sum_{d|n} \varphi(d) = \sum_{d|n} \varphi(\frac{n}{d}) = \sum_{d|n} 1(d)\varphi(\frac{n}{d}) = 1 * \varphi(n)$, we have

$$\zeta(s) \cdot \sum_{n=1}^{\infty} \frac{\varphi(n)}{n^s} = \sum_{n=1}^{\infty} \frac{1}{n^s} \cdot \sum_{n=1}^{\infty} \frac{\varphi(n)}{n^s} = \sum_{n=1}^{\infty} \frac{1 * \varphi(n)}{n^s} = \sum_{n=1}^{\infty} \frac{n}{n^s} = \sum_{n=1}^{\infty} \frac{1}{n^{s-1}} = \zeta(s-1),$$

i.e., $\sum_{n=1}^{\infty} \frac{\varphi(n)}{n^s} = \frac{\zeta(s-1)}{\zeta(s)}$. The sum of divisor function: $\sigma(n) = \sum_{d|n} d$. By definition of convolution,

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

In general, let $\sigma_k(n) = \sum_{d|n} d^k = \sum_{d|n} d^k 1(\frac{n}{d})$, then

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

Euler observed the $\zeta(s)$ diverges as $s \to 1^+$, so $\prod (1 - \frac{1}{p^s})^{-1}$ diverges as $s \to 1^+$. So there are infinitely many primes.

Euler asked: how does $\sum_p \frac{1}{p}$ behave?

3.1 The ζ function: elementary approach

Definition 3.11. For $\mathrm{Re}(s) > 1$, $\zeta(s)$ is defined by

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

Remark. There is a natural extension from $\zeta : \{2, 3, 4, \cdots\} \to \mathbb{R}$, to $\zeta : \{s \in \mathbb{C} \mid \mathrm{Re}(s) > 0\} \to \mathbb{C}$. Is there a natural way to extend ζ to larger domain?

First observation: need to deal with the problem at $s = 1$. Let's restrict ourself to real $s > 1$. Then $\frac{1}{x^s}$ is decreasing in $(0, \infty)$. So $\int_n^{n+1} \frac{1}{x^s} dx < \frac{1}{n^s} < \int_{n-1}^{n} \frac{1}{x^s} dx$ for $n \in \mathbb{N}$. Hence

$$\int_1^{\infty} \frac{1}{x^s} dx < \sum_{n=1}^{\infty} \frac{1}{n^s} = \frac{1}{1^s} + \sum_{n=2}^{\infty} \frac{1}{n^s} < 1 + \int_1^{\infty} \frac{1}{x^s} dx,$$

i.e., $\frac{1}{s-1} < \sum_{n=1}^{\infty} \frac{1}{n^s} < \frac{1}{s-1} + 1$. So the bad behaviors in $\zeta(s)$ appears to look like a singulrity $\frac{1}{s-1}$ at $s = 1$.

Question 3.12. . Can you improve this? Can you bound $\left| \frac{1}{n^s} - \int_n^{n+1} \frac{1}{x^s} ds \right|$? Can you identify errors and sum them too, giving error of the form $o(1)$? How does $\zeta(s) - \frac{1}{1-s}$ behave as $s \to 1^+$?

A different approach: Define the *Dirichlet η series*, sometimes called the *alternating ζ function*, denoted $\eta(s)$, $\zeta^*(s)$, $A(s)$ or $\alpha(s)$ by

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

By the alternating series test, for $\mathrm{Re}(s) > 0$, $\eta(s)$ is conditionally convergent.

Recall 3.13. If $\{a_n\}_{n \geqslant 1}$ such that $0 \leqslant a_n \to 0$ as $n \to \infty$ and $a_{n+1} \leqslant a_n$ for $n \in \mathbb{N}$, then $\sum_{n=1}^{\infty} (-1)^{n+1} a_n$ converges at least conditionally.

Proof. Note

$$a_1 - a_2 + a_3 - a_4 + a_5 - a_6 + \cdots = (a_1 - a_2) + (a_2 - a_3) + \cdots + (a_{2n-1} - a_{2n}) + \cdots.$$

Let $S_{2n} := a_1 - a_2 + a_3 - a_4 + \cdots + a_{2n-1} - a_{2n}$ and $S_{2n+1} := S_{2n} + a_{2n+1} \geqslant S_{2n}$. Then $S_{2n+1} = a_1 - (a_2 - a_3) - (a_4 - a_3) - \cdots - (a_{2n} - a_{2n+1}) \leqslant a_1$. So the sequence $S_2, S_4, S_6, \cdots$ is a bounded monotonic sequence, hence converges. Since $a_n \to 0$ as $n \to \infty$, $|S_{2n} - S_{2n+1}| = |a_{2n+1}| \to 0$ as $n \to \infty$. So $S_1, S_2, S_3, \cdots$ converges. ☐

Note

$$\eta(1) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \log 2 = \lim_{\epsilon \to 0} \log(2 - \epsilon) = \lim_{\epsilon \to 0} \int_0^{1-\epsilon} \frac{1}{1+x} dx.$$

For $\mathrm{Re}(s) > 0$,

$$\eta(s) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^s} = \frac{1}{1^s} - \frac{1}{2^s} + \frac{1}{3^s} - \frac{1}{4^s} + \cdots = \zeta(s) - \frac{2}{2^s}(\frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s}) = \zeta(s)(1 - 2^{s-1}).$$

So for $\mathrm{Re}(s) > 0$ with $s \neq 1$, we can define $\zeta(s)$ by

$$\zeta(s) = \frac{\eta(s)}{1 - 2^{s-1}}.$$

Let's consider for this extension of ζ to $(0,1) \cup (1,\infty)$, $\lim_{s\to 1}(s-1)\zeta(s) = \lim_{s\to 1}()$.

Proofs that there are infinitely many primes (usually) lead to lower bounds for $\pi(x)$. Recall $p_n < 2^{2^n}$, i.e., $\log_2 \log_2 p_n < n$, i.e., $\pi(x) > \log_2 \log_2 x$. Similarly get bad estimates from Euclid's proof. We'll get much more interesting estimator that Euler would have easily given.

Fact 3.14. For all $\epsilon > 0$,

(a) we have $\int_1^\infty \frac{1}{x}$ diverges, $\int_1^\infty \frac{1}{x^{1+\epsilon}}$ converges, $\int_2^\infty \frac{1}{x\log x}dx$ diverges, $\int_2^\infty \frac{1}{x(\log x)^{1+\epsilon}}dx$ converges, $\int_3^\infty \frac{1}{x\log x \log\log x}dx$ diverges, $\int_3^\infty \frac{1}{x\log x(\log\log x)^{1+\epsilon}}dx$ converges, $\cdots$;

(b) $\sum \frac{1}{n}$ diverges, $\sum \frac{1}{n^{1+\epsilon}}$ converges, $\sum \frac{1}{n\log n}$ diverges, $\sum \frac{1}{n(\log n)^{1+\epsilon}}$ converges, $\sum \frac{1}{n\log n\log\log n}$ diverges, $\sum \frac{1}{n\log n(\log\log n)^{1+\epsilon}}$ converges, $\cdots$.

Theorem 3.15 (Euler)**.** *If $s > 1$, then $\prod_p(1-\frac{1}{p^s})^{-1} = \sum_{n=1}^\infty \frac{1}{n^s}$. If finitely many primes, LHS bounded as $s \to 1^+$. Since RHS $\to \infty$ as $s \to 1^+$, LHS is a infinite product.*

Theorem 3.16 (Euler)**.** *Let $\{a_n\}_{n\geqslant 1}$ be with $0 < a_n < 1$ for $n \in \mathbb{N}$. Then $\prod_{k=1}^\infty(1+a_k)$ converges if and only if $\sum_{k=1}^\infty a_k$ converges if and only if $\prod_{k=1}^\infty(1-a_k)$ converges to > 0 if and only if $\prod_{k=1}^\infty(1-a_k)^{-1}$ converges. Hence if $\prod_p(1-\frac{1}{p})^{-1}$ diverges, then $\sum_p \frac{1}{p}$ diverges.*

Proof. Take logs and use careful approximations. □

Remark. Fix $\epsilon > 0$, supose $p_n > n\log n(\log\log n)^{1+\epsilon}$ for $n > n_0$, then $\frac{1}{p_n} < \frac{1}{n\log n(\log\log n)^{1+\epsilon}}$ for $n > n_0$, so $\sum_{n=1}^\infty \frac{1}{p_n}$ converges, a contradiction. Hence $p_n < n\log n(\log\log n)^{1+\epsilon}$ infinitely often. Note $p_n \approx n\log n$ is equivalent to $\pi(x) \approx x\frac{x}{x\log x} \approx \frac{x^2}{p_x} > \frac{x}{\log x(\log\log x)^{1+\epsilon}}$ infinitely often. This suggests that estimates for $\pi(x)$ might be around $\frac{x}{\log x}$.

We now extend previous discussion to show that $\eta(s)$ converges conditionally for $\mathrm{Re}(s) > 0$.

If $s = 1$, $\eta(s) = \sum_{n=1}^\infty \frac{(-1)^{n+1}}{n} = \sum_{k=1}^\infty(\frac{1}{2k-1} - \frac{1}{2k}) = \sum_{k=1}^\infty \frac{1}{2k(2k-1)}$, which is absolutely convergent. Assume now $s \in \mathbb{R}^{>0}$. Note

$$\sum_{n=1}^\infty \frac{(-1)^{n+1}}{n^s} = \sum_{k=1}^\infty\Big(\frac{1}{(2k-1)^s} - \frac{1}{(2k)^s}\Big) = \sum_{k=1}^\infty \frac{(2k)^s - (2k-1)^s}{(2k(2k-1))^s} = \sum_{k=1}^\infty \frac{1-(1-\frac{1}{2k})^s}{(2k-1)^s}.$$

By the Binomial theorem for non-integer exponents,

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 + \cdots + \binom{\alpha}{m}x^m + \cdots,$$

where $\binom{\alpha}{m} = \frac{\alpha(\alpha-1)\cdots(\alpha-m+1)}{m!}$. Assume $0 < s < 1$ now. Then $(s-1), (s-2), \cdots, (s-m+1)$ are all negative for $m \geqslant 2$. So $\binom{s}{m} > 0$ when $m \geqslant 2$ is even and $\binom{s}{m} < 0$ when $m \geqslant 2$ is odd. Hence

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

i.e., $1-(1-\frac{1}{2k})^s < \frac{s}{2k}$. So $\eta(s)=\sum_{k=1}^{\infty}\frac{1-(1-\frac{1}{2k})^s}{(2k-1)^s} < \sum_{k=1}^{\infty}\frac{s}{2k(2k-1)^s}$. Thus, $\eta(s)$ is conditionally convergent.

Remark. Since $(1-\frac{1}{2k})^s = 1-\frac{s}{2k}+\binom{s}{2}\frac{1}{(2k)^2}-\binom{s}{3}\frac{1}{(2k)^3}+\cdots$,

$$0<(1-\frac{1}{2k})^s-1+\frac{s}{2k}<\frac{1}{2}\frac{1}{(2k)^2}+\frac{1}{3}\frac{1}{(2k)^3}+\cdots<\frac{1}{(2k)^2}+\frac{1}{(2k)^3}+\cdots=\frac{1}{(2k)^2(1-\frac{1}{2k})}<\frac{2}{(2k)^2},$$

i.e., $\frac{s}{2k}-\frac{2}{(2k)^2}<1-(1-\frac{1}{2k})^s<\frac{s}{2k}$. So $\eta(s)=\sum_{k=1}^{\infty}\frac{1-(1-\frac{1}{2k})^s}{(2k-1)^s}=\sum_{k=1}^{\infty}(\frac{s}{2k(2k-1)^s}-\epsilon_k)$, where $0<\epsilon_k<\frac{2}{(2k)^2(2k-1)^s}$ for $k\geqslant 1$.

Question 3.17. Can you make a similarly argument work for $s\in\mathbb{C}$ with $\mathrm{Re}(s)>0$? Difficulties might arise comparing $\frac{s}{m}$ to $\frac{1}{m}$ or to 1? $\left|\binom{s}{m}\right|=\left|\frac{s(s-1)\cdots(s-(m-1))}{m!}\right|<$? Instead, more directly, write $s=\sigma+it, \sigma>0$. $\frac{1}{n^\sigma}=\frac{1}{n^\sigma n^{it}}=\frac{\cos(t\log n)-i\sin(t\log n)}{n^\sigma}$ and so $\mathrm{Re}(\eta(s))=\sum_{n=1}^{\infty}(-1)^{n+1}\frac{\cos(t\log n)}{n^\sigma}$, $\mathrm{Im}(\eta(s))=\sum_{n=1}^{\infty}(-1)^n\frac{\sin(t\log n)}{n^\sigma}$. These are no longer alternating series because

(a) the signs only alternate,

(b) the absolute values of summand are no longer monotonic.

However, the series will decompose into a sum of alternating pairs, together with "glitch" terms exponentially far apart. If we can bound the size of alternating pairs, e.g., $\frac{\cos(t\log n)}{n^\sigma}-\frac{\cos(t\log(n+1))}{(n+1)^\sigma}$ and similarly for the sin terms and bound the contribution of the terms when $\cos(t\log n)$ switches signs.

Exercise 3.18. Go back to Calculus I/II and find bounds for $\frac{\cos(t\log n)}{n^\sigma}-\frac{\cos(t\log(n+1))}{(n+1)^\sigma}$ by differentiating $\frac{\cos(t\log x)}{x^\sigma}$. Deduce that $\eta(s)=\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n^s}$ converges conditionally for all $s\in\mathbb{C}$ with $\mathrm{Re}(s)>0$.

Question 3.19. $\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}$ for $\mathrm{Re}(s)>1$, $\zeta(s)=\frac{1}{1-z^{s-1}}\eta(s)$ for $\mathrm{Re}(s)<0$ with $s\neq 1$. Both definition agrees in $\{s\in\mathbb{C}\mid \mathrm{Re}(s)>1\}$, big question, how do we extend ζ to the rest of $\mathbb{C}$?

Chapter 4

Averages of Arithmetic Functions

Number theoretic functions typically bounce around a lot. We want to be able to describe this smoothed-out behaviors. We'll do this using the concept of "average order". We'll say that two functions f and g have the same average order if $\sum_{n\leqslant x} f(n) \sim \sum_{n\leqslant x} g(n)$ as $x \to \infty$. Typically, g will be a "nice function", e.g., $g(n) = n$, $g(n) = \frac{1}{\log n}$, $g(n) = n^2$, $\cdots$.

$f(n)$ has average order n if $\sum_{n\leqslant x} f(n) \sim \frac{x^2}{2}$. Sometimes we may wish for extra asymtotic terms on RHS. For exmple, if x is an integer, $\sum_{n\leqslant x} n = \frac{x(x+1)}{2} = \frac{x^2}{2} + \frac{x}{2}$. So $\sum_{n\leqslant x} n \sim \frac{x^2}{2} + \frac{x}{2}$.

Remark. Typically a more accurate asymtotic estimate will imply extra smoothness properties of f. Some functions for which we might try to find average orders: $\tau(n), \sigma(n)$ and $\varphi(n)$.

Since τ is multiplicative, $\tau(n) = \prod_{p^k||n} \tau(p^k) = \prod_{p^k||n}(k+1) = f(n)$. So $a_{p^k} = \tau(p^k) = k+1$ and then the Dirichlet series

$$\sum_{n=1}^{\infty} \frac{\tau(n)}{n^s} = \prod_p F_p(s) = \prod_p \sum_{k=1}^{\infty} \frac{a_{p^k}}{(p^k)^s} = \prod_p \Big(1 + \frac{2}{p} + \frac{3}{p^2} + \frac{4}{p^3} + \cdots\Big).$$

Question 4.1. How big can $\tau(n)$ be? One approach: multiply by losts of **distinct** primes together. Let $n_k = \prod_{p\leqslant k} p$. Then $\tau(n_k) = \binom{k}{0} + \binom{k}{1} + \binom{k}{2} + \cdots + \binom{k}{k} = 2^k$. This will be almost the best we can do. Check $\log n_k = \sum_{p\leqslant k} \log p \sim k$ by assuming PNT and Abel summation. So n_k is about e^k and $\tau(n_k) = 2^k = e^{k\log 2} \approx n^{\log 2} = n^{0.69}$.

The average order is

$$\begin{aligned}\sum_{n\leqslant x} \tau(n) &= \sum_{d\leqslant x} \Big\lfloor \frac{x}{d} \Big\rfloor = \sum_{d\leqslant x} \frac{x}{d} + O(x) = x \sum_{d\leqslant x} \frac{1}{d} + O(x)\\ &= x(\log x + \gamma + O(1)) + O(x) = x\log x + O(x),\end{aligned}$$

since each divisor d gets counted $\lfloor \frac{x}{d} \rfloor$ times and $\frac{x}{d} - 1 \leqslant \lfloor \frac{x}{d} \rfloor \leqslant \frac{x}{d}$.

Recall/Check $\sum_{n\leqslant x} \log n \sim x\log x - x + \frac{1}{2}\log x$. So $\tau(n)$ has average order $\log n$.

When d is large, the summand is small and the error in approximating $\lfloor \frac{x}{d} \rfloor$ by $\frac{x}{d}$ is relatively large. Furthermore, there are lots of large values of $d \leqslant x$. When d is small, the summand is much large, and the relative errors are small, and there are fewer values of d.

Dirichlet observed that $\sum_{md\leqslant x} 1$ is the number of integer points which lie beneath the hyperbola $uv = x$. So

$$\begin{aligned}\sum_{n\leqslant x}\tau(n) &= 2\sum_{md\leqslant x, m\geqslant d} 1 - \sum_{d^2\leqslant x} 1 = 2\sum_{d\leqslant\sqrt{x}}\sum_{m\leqslant\frac{x}{d}} 1 - \lfloor\sqrt{x}\rfloor = 2\sum_{d\leqslant\sqrt{x}}\frac{x}{d} + O(\sqrt{x})\\ &= 2\left(x\log\sqrt{x} + x\gamma + O\Big(\frac{x}{\sqrt{x}}\Big)\right) + O(\sqrt{x}) = 2x\log\sqrt{x} + 2\gamma x + O(\sqrt{x})\\ &= x\log x + 2\gamma x + O(\sqrt{x}).\end{aligned}$$

Note $\sigma(p^k) = 1 + p + p^2 + \cdots + p^k = \frac{p^{k+1}-1}{p-1}$.

Question 4.2 (Open)**.** Does there exists odd n so that $\sigma(n) = 2n$?

Example 4.3. Note

$$\begin{aligned}\sum_{n\leqslant x}\sigma(n) &= \sum_{n\leqslant x}\sum_{d|n} d = \sum_{md\leqslant x} d = \sum_{m\leqslant x}\sum_{d\leqslant\frac{x}{m}} d = \sum_{m\leqslant x}\frac{1}{2}\Big\lfloor\frac{x}{m}\Big\rfloor\Big(\Big\lfloor\frac{x}{m}\Big\rfloor + 1\Big)\\ &= \frac{1}{2}\sum_{m\leqslant x}\Big(\frac{x}{m} - \epsilon_m\Big)\Big(\frac{x}{m} + 1 - \epsilon_m\Big) = \frac{1}{2}\sum_{m\leqslant x}\Big(\frac{x^2}{m^2} + \frac{x}{m}(1 - 2\epsilon_m) - \epsilon_m(1 - \epsilon_m)\Big)\\ &= \frac{1}{2}x^2\sum_{m\leqslant x}\frac{1}{m^2} + \frac{1}{2}xO\left(\sum_{m\leqslant x}\frac{1}{m}\right) + O(x) = \frac{1}{2}x^2\left(\zeta(2) - \sum_{m>x}\frac{1}{m^2}\right) + O(x\log x) + O(x)\\ &= \frac{1}{2}x^2\zeta(2) + O(x) + O(x\log x) = \frac{1}{2}x^2\zeta(2) + O(x\log x) \sim \frac{\pi^2 x^2}{12} + O(x\log x).\end{aligned}$$

Hence average order of $\sigma(n)$ is $\frac{\pi^2}{6}n$.

Question 4.4. How large can $\sigma(n)$ be?

Example 4.5. Since $n = \sum_{d|n}\varphi(d)$, $\varphi(n) = \sum_{d|n}\mu(d)\frac{n}{d}$ by Möbius inversion. So

$$\begin{aligned}\sum_{n\leqslant x}\varphi(n) &= \sum_{n\leqslant x}\sum_{d|n}\mu(d)\frac{n}{d} = \sum_{md\leqslant x}\mu(d)m = \sum_{d\leqslant x}\mu(d)\sum_{m\leqslant\frac{x}{d}} m = \frac{1}{2}\sum_{d\leqslant x}\mu(d)\Big\lfloor\frac{x}{d}\Big\rfloor\Big(\Big\lfloor\frac{x}{d}\Big\rfloor + 1\Big)\\ &= \frac{1}{2}\sum_{d\leqslant x}\mu(d)\Big(\frac{x^2}{d^2} + O\Big(\frac{x}{d}\Big)\Big) = \frac{1}{2}x^2\sum_{d\leqslant x}\frac{\mu(d)}{d^2} + O(x\log x).\end{aligned}$$

Now $\sum_{d=1}^{\infty}\frac{\mu(d)}{d^2}$ is absolutely convergent and is $\frac{1}{\zeta(2)} = \frac{6}{\pi^2}$. So the average order of $\varphi(n)$ is $\frac{6}{\pi^2}n$.

Interpretetion: Fix x large, Pick two integers uniformly in $[1, x]$. The probability that they are coprime is about $\frac{6}{\pi^2}$.

Remark (Upper bound on $\pi(x)$)**.** Tchebyshev(when? How long from proving these to Hadamard/de la Valee-Poussin proving $\pi(x) \sim \mathrm{Li}(x)$?) proved that there are constants A, B so that for all $x \geqslant 2$, $\frac{Ax}{\log x} < \pi(x) < \frac{Bx}{\log x}$. He also proved that if there is a contant C so that $\frac{\pi(x)}{C\frac{x}{\log x}} \to 1$, then $C = 1$.

Problem 4.6 (Open)**.** Show there are infinitely many n so that $\binom{2n}{n}$ is coprime to 205; but only finitely many n so that $\binom{2n}{n}$ is coprime 1115.

4.1 Estimate $\pi(x)$ using factorials and binomial coefficients

Definition 4.7. Let p be prime. The *p-adic valuation* or *p-adic order* is the function given $\nu_p : \mathbb{Z}_{\geqslant 0} \to \mathbb{N} \cup \{\infty\}$ defined by

$$\nu_p(n) := \{k : p^k \mid\mid n\} =: \mathrm{ord}_p(n).$$

Theorem 4.8 (Legendre's formula)**.**

$$\nu_p(n!) = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \cdots + \left\lfloor \frac{n}{p^k} \right\rfloor = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \cdots,$$

where k is such that $p^k \leqslant n < p^{k+1}$.

Remark. Note $\nu_p(n!) \leqslant \frac{n}{p} + \frac{n}{p^2} + \cdots + \frac{n}{p^k} + \cdots = \frac{n}{p-1}$. If $p < n$, $\nu_p(n!) \geqslant \left\lfloor \frac{n}{p} \right\rfloor$. If $p^2 < n$, $\nu_p(n!) \geqslant \frac{n}{p}$. By definition, $n! = \prod_{p \leqslant n} p^{\nu_p(n!)}$, i.e., $\log(n!) = \sum_{p \leqslant n} \nu_p(n!) \log p$.

Question 4.9. Can we use this identity to prove that there are infinitely primes? Let's suppose that there were only finitely many primes $p_1, \ldots, p_k \leqslant p$. Then

$$\log(n!) \leqslant \sum_{p \leqslant n} \frac{n}{p-1} \log p \leqslant \sum_{p \leqslant n} n \log p = nk \log p < np \log p.$$

But we know $\log(n!) \sim n \log n - n + \frac{1}{2} \log n + O(1)$. So $\log(n!) > np \log p$ as soon as $\log n >> p \log p$. Clealry, this naive bound is not going to be near the truth. Let's try and be smarter.

Note $\log(n!) = \sum_{p \leqslant n} \nu_p(n!) \log p$. How big is the error in estimating $\nu_p(n!)$ by $\frac{n}{p-1}$ or $\frac{n}{p}$? and can you improve the estimate?

If $\frac{n}{2} < p \leqslant n$, then $\nu_p(n!) = 1$, so $\log(n!) = \sum_{p \leqslant \frac{n}{2}} \nu_p(n) \log p + \sum_{\frac{n}{2} < p \leqslant n} \log p$.

If $\frac{n}{3} < p \leqslant \frac{n}{2}$, then $\nu_p(n!) = 2$, so $\log(n!) = \sum_{p \leqslant \frac{n}{3}} \nu_p(n) \log p + \sum_{\frac{n}{3} < p \leqslant \frac{n}{2}} 2 \log p + \sum_{\frac{n}{2} < p \leqslant n} \log p$.

In fact, so long as $p^2 > n$, $\nu_p(n) = \left\lfloor \frac{n}{p} \right\rfloor$, and $p^2 \leqslant n$ if and only if $p \leqslant \frac{n}{(\lceil n^{\frac{1}{2}} \rceil - 1) + 1} = \frac{n}{\lceil n^{\frac{1}{2}} \rceil}$, so

$$\log(n!) = \sum_{p^2 \leqslant n} \nu_p(n) \log p + \sum_{k=1}^{\lceil n^{\frac{1}{2}} \rceil - 1} \sum_{\frac{n}{k+1} < p \leqslant \frac{n}{k}} k \log p = \sum_{p^2 \leqslant n} \nu_p(n) \log p + \sum_{k=1}^{\lceil n^{\frac{1}{2}} \rceil - 1} k \sum_{\frac{n}{k+1} < p \leqslant \frac{n}{k}} \log p.$$

Assuming PNT, what is contributed from $k = 1$: $\sum_{\frac{n}{2} < p \leqslant n} \log p < \sum_{\frac{n}{2} < p \leqslant n} \log n = \log n (\pi(n) - \pi(\frac{n}{2}))$. Since $\pi(n) \sim \frac{n}{\log n}$ and $\pi(\frac{n}{2}) \sim \frac{n}{2} \frac{1}{\log \frac{n}{2}}$,

$$\pi(n) - \pi\Big(\frac{n}{2}\Big) \text{“=”} \frac{n}{\log n} - \frac{n}{2 \log \frac{n}{2}} = n \frac{2 \log \frac{n}{2} - \log n}{2 \log \frac{n}{2} \log n} = n \frac{\log n - 2 \log 2}{2 \log \frac{n}{2} \log n} \approx \frac{n}{2 \log \frac{n}{2}} \approx \frac{n}{2 \log n}.$$

Assuming PNT,

$$\sum_{p > n^{\frac{1}{2}}} \nu_p(n) \log p = \sum_{k=1}^{n^{\frac{1}{2}}} k \sum_{\frac{n}{k+1} < p \leqslant \frac{n}{k}} \log p \approx \sum_{k=1}^{n^{\frac{1}{2}}} k \left(\sum_{p \leqslant \frac{n}{k}} \log p - \sum_{p \leqslant \frac{n}{k+1}} \log p \right)$$
$$= \sum_{k=1}^{n^{\frac{1}{2}}} k \Big(\frac{n}{k} - \frac{n}{k+1} \Big) = n \sum_{k=1}^{n^{\frac{1}{2}}} \frac{1}{k+1} \approx n \log n^{\frac{1}{2}} = \frac{1}{2} n \log n.$$

Where are we making errors here? Can we bounded them?

Exercise 4.10. Write $e(n) = \pi(n) - \frac{n}{\log n}$, estimate $\sum_{\frac{n}{k+1} < p \leqslant \frac{n}{k}} k \log p$ assuming PNT, with error bounds in terms of $e(n)$.

Remark. Using $n!$ to investigate the distribution of primes:

(a) get $\pi(x) \to \infty$;

(b) where is the greatest contribution to $\log(n!) = \sum_{p \leqslant n} \nu_p(n) \log p \approx n \sum_{p \leqslant n} \frac{\log p}{p-1}$;

(1) primes are more frequent when they are small;

(2) $\frac{\log k}{k-1} \to 0$ as $k \to \infty$.

So the biggest contribution is from the smallest primes. This suggests that this approach will have difficulty being sensitive enough to behavior of large primes to estimate them. (Assuming PNT, heuriestically since we want to check that this method won't be sensitive enough to count large primes). So splitting $\sum_{p \leqslant n} \nu_p(n) \log p = \sum_{p \leqslant n^{\frac{1}{2}}} \nu_p(n) \log n + \sum_{p > n^{\frac{1}{2}}} \nu_p(n) \log p$ gives approximately equal contributions. So, to use this to estimate $\pi(x)$, we'd have to control our first half very well, which is hard, since $\nu_p(n)$ is not nicely behaved when p is small compared to p. The difficulty is that $\nu_p(n)$ is too big, and bounces around too much.

Let $n = 2m$. Consider $\binom{2m}{m} = \frac{(2m)!}{m!m!} \sim \frac{(\frac{2m}{e})^{2m}}{(\frac{m}{e})^{2m}} \frac{\sqrt{2\pi 2m}}{2\pi m} = \frac{2^{2m}}{\sqrt{\pi m}} < 4^m$ for $m \geqslant 1$. Note every prime between $m+1$ and $2m$ divides $\binom{2m}{m}$. So $\prod_{m<p<2m} p < 4^m$. Hence $m^{\pi(2m)-\pi(m)} < 4^m$, i.e., $\pi(2m) - \pi(m) < \log_m 4^m = \frac{m \log 4}{\log m} = \frac{n \log 2}{\log \frac{n}{2}} < \frac{n}{\log n}$ if $n > 2^{\text{some number}}$. So we get (with some work) $\pi(n) < \frac{2n}{\log n}$.

Remark. Note if we consider $\binom{2m-1}{m}$, we can get $\pi(n) - \pi(\frac{n}{2}) < \frac{n \log 2}{\log n - \log 2}$, $\pi(\frac{n}{2}) - \pi(\frac{n}{4}) < \frac{\frac{n}{2} \log n}{\log n - \log 4}$, $\cdots$. Some case then gives $\pi(n) < \frac{2n \log 2}{\log n}(1 + \epsilon)$.

Exercise 4.11. A lower bound on $\pi(x)$ is $\frac{n}{\log n} + \frac{2n}{\log^2 n} + \frac{3n}{\log^3 n} + \cdots$.

We have # digits in n base $p \leqslant \log_p n + 1$, # number of carries $\leqslant \log_p n$.

Exercise 4.12. $\nu_p(\binom{n}{\frac{n}{2}}) = \left\{k : p^k \,\Big|\Big|\, \binom{n}{\frac{n}{2}}\right\}$ and

$$\nu_p\left(\binom{n}{\frac{n}{2}}\right) = \nu_p(n) - 2\nu_p(\frac{n}{2}) = \#\left\{\text{times we carry computing } \frac{n}{2} + \frac{n}{2} = n \text{ in base } p\right\} \leqslant \log_p n.$$

Remark. If $p > n^{\frac{1}{2}}$, i.e., $p^2 > n$, then $\nu_p(n) = \lfloor \frac{n}{p} \rfloor$ and $\nu_p(\frac{n}{2}) = \lfloor \frac{n}{2p} \rfloor$, so $\nu_p(n) - 2\nu_p(\frac{n}{2}) = \lfloor \frac{n}{p} \rfloor - 2\lfloor \frac{n}{2p} \rfloor = 0$ or 1.

Let $\alpha_p(n) = \nu_p(n) - 2\nu_p(\frac{n}{2})$. Then $\binom{n}{\frac{n}{2}} = \prod_{p \leqslant n} p^{\alpha_p(n)}$. So

$$\sum_{p \leqslant n} \alpha_p(n) \log p = \log \binom{n}{\frac{n}{2}} \approx n \log 2 - \frac{1}{2} \log n + \log \sqrt{\frac{2}{\pi}}.$$

Also, assuming PNT,

$$\sum_{p\leqslant n^{\frac{1}{2}}} \alpha_p(n)\log p \leqslant \sum_{p\leqslant n^{\frac{1}{2}}} \log_p n \log p = \sum_{p\leqslant n^{\frac{1}{2}}} \log n = \log n \pi(n^{\frac{1}{2}}) \approx \log n \frac{n^{\frac{1}{2}}}{\log n^{\frac{1}{2}}} = 2n^{\frac{1}{2}}.$$

So this contribution is small enough that it is on the order of the conjected value in RH.

Question 4.13. When are $\nu_p(n) = 2\nu_p(\frac{n}{2})$ and $\nu_p(n) = 2\nu_p(\frac{n}{2}) + 1$?

Exercise 4.14. Let $n = 2m$. Then for $p > \sqrt{m}$, $p \mid\mid \binom{2m}{m}$ when one of the following holds (i) $p > m$ or $\frac{2m}{2} < p$, $\frac{n}{2} < p \leqslant n$ (ii) $\frac{2m}{4} < p \leqslant \frac{2m}{3}$, $\frac{n}{4} < p \leqslant \frac{n}{3}$, (iii) $\frac{2m}{6} < p \leqslant \frac{2m}{5}$, $\frac{n}{6} < p \leqslant \frac{n}{5}$, (iv) $\frac{2m}{5} < p \leqslant \frac{2m}{7}$, $\frac{n}{8} < p \leqslant \frac{n}{7}$. So

$$\log \binom{n}{\frac{n}{2}} = \sum_{p\leqslant n^{\frac{1}{2}}} \alpha_p(n)\log p + \sum_{2k<n^{\frac{1}{2}}} \sum_{\frac{n}{2k}<p\leqslant\frac{n}{2k-1}} \log p.$$

Hence the second term captures most of $n\log 2$ and there exists c so that $\pi(x) \sim \frac{cx}{\log x}$. So we can estimate $\sum_{\frac{n}{2k}<p\leqslant\frac{n}{2k-1}} \log p$ using abelian summation and deduce Chebyshev's result that c must be 1. However, trying to prove that c must exist using this method eludes Chebyshev and many others.

Books to read: The Thread, Philip Davis, Mathematician's apology, G.H.Hardy, Two cultures. C.P. Snow.

4.2 Return to ζ

Bernoulli asked what the value of $\sum_{n\geqslant 1}\frac{1}{n}$ was. Euler showed $\sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6}$. (Why "showed" not "proved"?)

If $p(x) = \prod_{i=1}^{n}(x - a_i)$ is a polynomial, then $[x^{n-1}]p(x) = -a_1 - \cdots - a_n$. Note $\sin x$ has roots at $0, \pm\pi, \pm 2\pi, \cdots$, so let's write $\sin x = \prod_{k=-\infty}^{\infty}(x - \pi k)$. This is a bad idea.

Better idea: write $p(x) = \left(\prod(1 - \frac{x}{a_i})\right)\cdot p(0)$, where a_i's are roots of $p(x)$. Since $\sin 0 = 0$ and $\lim_{x\to 0}\frac{\sin x}{x} = 1$, Euler wrote $\frac{\sin x}{x} = \prod_{k\neq 0}(1 - \frac{x}{\pi k})\cdot 1$. This is a bad idea since $\sum_{k=-\infty,k\neq 0}^{\infty} -\frac{1}{\pi k} = \sum_{k=-\infty,k\neq 0}^{\infty}\frac{1}{\pi k}$ is not absolutely convergent.

Better idea: merge the roots at $\pm\pi k$ and write $\frac{\sin x}{x} = \prod_{k=1}^{\infty}(1 - \frac{x^2}{(\pi k)^2})$. (It is not a rigorous thing to do: but it turns out that for deep reasons in Complex Analysis that for $\sin x$, this works.) Then we have an identity

$$\sin x = x\prod_{k=1}^{\infty}\Big(1 - \frac{x^2}{\pi^2 k^2}\Big).$$

RHS converges to $\sin x$ at every $x \in \mathbb{R}$. Euler assumed this, and deduced that $[x^3]\sin x = [x^2]\prod_{k=1}^{\infty}(1 - \frac{x^2}{\pi^2k^2}) = -\sum_{k=1}^{\infty}\frac{1}{\pi^2k^2}$. But $[x^3]\sin(x) = \frac{-1}{3!} = -\frac{1}{6}$. Hence $\sum_{k=1}^{\infty}\frac{1}{\pi^2k^2} = \frac{1}{6}$. So

$\zeta(2) = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$. $[x^5]\sin x = \frac{1}{5!}$. Similarly,

$$[x^5]x\prod\Big(1-\frac{x^2}{\pi^2k^2}\Big) = [x^4]\prod\Big(1-\frac{x^2}{k^2\pi^2}\Big) = \sum_{k\neq l}\frac{1}{\pi^2k^2\pi^2l^2} = \frac{1}{2}\Bigg(\sum_{k,l=1}^{\infty}\frac{1}{\pi^2k^2}\frac{1}{\pi^2l^2} - \sum_{k=1}^{\infty}\frac{1}{(\pi^2k^2)^2}\Bigg)$$

$$= \frac{1}{2}\Bigg(\Big(\sum_{k=1}^{\infty}\frac{1}{\pi^2k^2}\Big)^2 - \frac{\zeta(4)}{\pi^4}\Bigg) = \frac{1}{2}\Bigg(\Big(\frac{1}{6}\Big)^2 - \frac{\zeta(4)}{\pi^4}\Bigg).$$

But $[x^5]\sin x = \frac{1}{5!} = \frac{1}{120}$, we have $\zeta(4) = \frac{\pi^4}{90}$.

Fact 4.15. There are no non-real zeros of $\sin z$.

Proof. Write $2\sin z = e^{iz} - e^{-iz}$, so if $\sin z = 0$, $e^{iz} = e^{-iz}$. Set $z = x + iy$ with $x, y \in \mathbb{R}$, then $e^{ix-y} = e^{-ix+y}$, so $|e^{-y}| = |e^y|$, i.e., $y = 0$, hence $z \in \mathbb{R}$. □

Definition 4.16 (Euler's formula for $\sin(\pi x)$)**.**

$$\sin(\pi z) = \pi z\prod_{k=1}^{\infty}\Big(1-\frac{z^2}{k^2}\Big),\ z \in \mathbb{C}.$$

Motivated by this, our goal now will be consider ζ as a function of $s \in \mathbb{C}$ and write

$$\zeta(s) = ()\prod_{n=1}^{\infty}\Big(1+\frac{s}{2n}\Big)e^{-\frac{s}{2n}}\prod_{r\text{: nontrivial roots}}\Big(1-\frac{s}{r}\Big)e^{\frac{s}{r}}.$$

Note 4.17. (a) So far () is opaque.

(b) The "trivial" zeros of $\zeta(s)$ are all known: all negative even integers. Note we don't know how to even define $\zeta(s)$ if $\mathrm{Re}(s) \leqslant 0$.

Euler's factorization of $\sin(\pi x) = -\pi\prod_{k=-\infty}^{\infty}(k-x)$ didn't work. No convergence! Likewise, $\sin(\pi x) = \pi x\prod_{k\neq 0,k=-\infty}^{\infty}(1-\frac{x}{k})$ didn't work either, because $\sum_{k\neq 0,k=-\infty}^{\infty}\frac{x}{k}$ diverges. However, if we were to group together $(1-\frac{x}{k})$ and $(1+\frac{x}{k})$, then the convergence issues vanished.

An alternative approach: $\frac{1}{k}$ goes to zero too slowly. We want a function which has a root at $x = k$, but when we expand it out in terms of $\frac{x}{k}$, the linear terms vanish. If we consider $e^{\frac{x}{k}}$, this has no roots, but

$$\Big(1-\frac{x}{k}\Big)e^{\frac{x}{k}} = \Big(1-\frac{x}{k}\Big)\Big(1+\frac{x}{k}+\frac{x^2}{2k^2}+\frac{x^3}{6k^3}+\cdots\Big) = 1+\frac{x}{k}-\frac{x}{k}+\frac{x^2}{2k^2}-\frac{x^2}{k^2}+\frac{x^3}{6k^3}-\frac{x^3}{2k^3}+\cdots$$

$$= 1-\frac{x^2}{2k^2}-\frac{x^3}{3k^3}-\frac{x^4}{8k^4}-\cdots = 1-\frac{x^2}{2k^2}+O\Big(\frac{x^3}{k^3}\Big).$$

This factor has a root at $\frac{x}{k}$, but $\prod_{k=-\infty,k\neq 0}^{\infty}(1-\frac{x}{k})e^{\frac{x}{k}} = \prod_{k=-\infty,k\neq 0,}^{\infty}\big(1-\frac{x^2}{2k^2}+O(\frac{x^3}{k^3})\big)$, for small x at least, will be convergent. So $\pi x\prod_{k\neq 0,k=-\infty}^{\infty}(1-\frac{x}{k})e^{\frac{x}{k}}$ converges. Keep the terms $(1-\frac{x}{k})$ and $e^{\frac{x}{k}}$ paired, we can write $\sin x = \pi x\prod_{k=1}^{\infty}(1+\frac{x}{k})e^{-\frac{x}{k}}(1-\frac{x}{k})e^{\frac{x}{k}} = \pi x\prod_{k=1}^{\infty}(1-\frac{x^2}{k^2})$ as before.

In our hoped-for expression for $\zeta(s)$, the terms $\prod_{n=1}^{\infty}(1+\frac{s}{2n})e^{-\frac{s}{2n}}$ work even if now we don't have the symmetric part with $n \leqslant -1$. Provided $\sum\frac{1}{|r^2|}$ is convergent, the remaining product over the non-trivial roots $\prod(1-\frac{s}{r})e^{\frac{s}{r}}$ works, simialrly. This will be something we'll have to worry about.

Note

$$\log(\zeta(s)) = \log() + \sum_{n=1}^{\infty}\Big(\log\big(1+\frac{s}{2n}\big) - \frac{s}{2n}\Big) + \sum_r \Big(\log\big(1-\frac{s}{r}\big) + \frac{s}{r}\Big).$$

Since $\log(\zeta(s)) = \sum_p -\log(1-\frac{1}{p^s})$,

$$\frac{\zeta'(s)}{\zeta(s)} = \big(\log(\zeta(s))\big)' = \sum_p \frac{\frac{1}{p^s}\log p}{1-\frac{1}{p^s}} = \sum_p \frac{1}{p^s}\log p \sum_{k\geqslant 0}\frac{1}{p^{ks}} = \sum_p\sum_{k\geqslant 0}\frac{\log p}{p^{(k+1)s}} = \sum_p\sum_{k\geqslant 1}\frac{\log p}{p^{ks}}.$$

We'll want to relate $\pi(x)$ to the behavior of this Dirichlet series.

Note if we take $\log f$, where f is a complex function of z, we get bad behavior when $f(z) = 0$ and $f(z) = \infty$. If $f(z) = (z-r)^\alpha$, near $z = r$, $\frac{f'(z)}{f(z)}$ behaves like $\frac{\alpha(z-r)^{\alpha-1}}{z-r} = \frac{\alpha}{z-r}$. So if r is either a pole or a root of some order, possibly even fractional, it converted to a simple pole of $\frac{f'(z)}{f(z)}$.

Our goal is to understand $\pi(x) = \sum_{p\leqslant x} 1$. However, we'll see that it is sometimes easier to work with related functions: for example,

$$\Pi(x) = \sum_{p^k\leqslant x} 1 = \sum_{p\leqslant x} 1 + \sum_{p\leqslant x^{\frac{1}{2}}} 1 + \sum_{p\leqslant x^{\frac{1}{3}}} 1 + \cdots = \pi(x) + \pi(x^{\frac{1}{2}}) + \pi(x^{\frac{1}{3}}) + \pi(x^{\frac{1}{4}}) + \cdots = \sum_{r\geqslant 1}\pi(x^{\frac{1}{r}}).$$

Then good approximation for $\Pi(x)$ can give approximations for $\pi(x)$ via

$$\pi(x) = \sum_{r\geqslant 1}\pi(x^{\frac{1}{r}})\mu(r) = \pi(x) - \pi(x^{\frac{1}{2}}) - \pi(x^{\frac{1}{3}}) - \pi(x^{\frac{1}{5}}) + \pi(x^{\frac{1}{6}}) - \cdots.$$

Note that these apparently infinite series for Π in terms of π, and π in terms of Π actually only have at most $\log_2 x$ terms for any fixed $x \geqslant 1$. If $r > \log_2 x$, $x^{\frac{1}{r}} < x^{\frac{1}{\log_2 x}} = 2$. We'll also want to transform between estimate for $\sum_{p\leqslant x}\log p, \sum_{p^k\leqslant x}\log p$, and estimate for $\Pi(x)$, $\pi(x)$.

Definition 4.18. Define the *von Mangoldt function* $\Lambda : \mathbb{N} \to \mathbb{R}$ by

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \text{ is a prime power} \\ 0 & \text{otherwise} \end{cases}.$$

Exercise 4.19. Now let $a_m = \Lambda(m)$ and $b_m = \frac{1}{\log m}$. Write $\psi(x) = \sum_{p^k\leqslant x}\log p = \sum_{m\leqslant x}\Lambda(m)$. Express $\Pi(x)$ in terms of $\psi(x)$ using summation by parts.

Theorem 4.20.

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

Then by Möbius inversion,

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

Proof. It follows from the fundamental theorem of arithmetic. □

4.3 Gamma Function

Definition 4.21. Define the *gamma function* by

$$\Gamma(s) = \int_0^\infty e^{-t} t^{s-1} dt, \ \operatorname{Re}(s) > 0.$$

Remark. This integral converges because

(a) $\int_0^1 \left| e^{-t} t^{s-1} \right| dt < \int_0^1 \left| t^{s-1} \right| dt = \int_0^1 t^{\operatorname{Re}(s)-1} dt < \infty$;

(b) $\int_1^\infty e^{-t} t^{s-1} dt$ converges: for any $t > 0$, $\frac{t^n}{n!} < e^t$, so $e^{-t} < \frac{n!}{t^n}$, now choose $n_s > \operatorname{Re}(s)$, then $\int_1^\infty \left| e^{-t} t^{s-1} \right| dt \leqslant \int_1^\infty \frac{n_s!}{t^{n_s}} \left| t^{s-1} \right| dt = n_s! \int_1^\infty t^{-(n_s+1-\operatorname{Re}(s))} dt = \frac{n_s!}{n_s - \operatorname{Re}(s)} < \infty$. The second point actually shows that $\int_1^\infty e^{-t} t^{s-1} dt$ converges for all $s \in \mathbb{C}$!

Fact 4.22. $\Gamma(n+1) = n!$ for $n \in \mathbb{Z}_{\geqslant 0}$.

Theorem 4.23 (Functional equation)**.** $\Gamma(s+1) = s\Gamma(s)$ *for* $\operatorname{Re}(s) > 0$.

Proof. Use integration by parts. ☐

We could also now define $\Gamma(s)$ for $s \in \mathbb{C} \setminus \{0, -1, -2, -3, -4, \cdots\}$ via $\Gamma(s) = \int_0^\infty e^{-t} t^{s-1} dt$ if $\operatorname{Re}(s) > 0$ and $\Gamma(s) = \frac{\Gamma(s+1)}{s} = \frac{\Gamma(s+2)}{s(s+1)} = \cdots = \frac{\Gamma(s+k)}{s(s+1)\cdots(s+k-1)}$ for $\operatorname{Re}(s) > k!$. This only breaks at non-positive integers.

Instead, let's consider $\int_0^1 e^{-t} t^{s-1} dt$ more carefully. Provided $\operatorname{Re}(s) > 0$, we have

$$\int_0^1 \sum_{k=0}^\infty \left| (-1)^k \frac{t^k}{k!} t^{s-1} \right| dt = \int_0^1 \sum_{k=0}^\infty \frac{t^k}{k!} \left| t^{s-1} \right| dt = \int_0^1 e^t \left| t^{s-1} \right| dt < e \int_0^1 t^{\operatorname{Re}(s)-1} dt < \infty.$$

So by Fubini/Tonelli theorems,

$$\int_0^1 e^{-t} t^{s-1} dt = \int_0^1 \sum_{k=0}^\infty (-1)^k \frac{t^k}{k!} t^{s-1} dt = \sum_{k=0}^\infty \frac{(-1)^k}{k!} \int_0^1 t^{k+s-1} dt = \sum_{k=0}^\infty \frac{(-1)^k}{k!} \frac{1}{k+s}.$$

Hence for $\operatorname{Re}(s) > 0$,

$$\Gamma(s) = \sum_{k=0}^\infty \frac{(-1)^k}{k!} \frac{1}{k+s} + \int_1^\infty e^{-t} t^{s-1} dt.$$

This formula works for all $s \in \mathbb{C} \setminus \{0, -1, -2, -3, \cdots\}$ and in fact, if we define $\Gamma(s)$ this way, then $\Gamma(s) - \sum_{k=0}^\infty \frac{(-1)^k}{k!} \frac{1}{k+s}$ is analytic around $s = -k$.

Exercise 4.24. Show that with the new definition of $\Gamma(s)$ for all $s \neq 0, -1, -2, -3, \cdots$, $\Gamma(s+1) = s\Gamma(s)$. So the two extensions to $\mathbb{C} \setminus \{0, -1, -2, \cdots\}$ agree.

Why are we interested in $\Gamma(s)$? $\sin(s)$? Our underlying goal is to studying $\pi(x)$? We will see that $\pi(x)$ can be expressed as a function over the roots of $\zeta(s)$.

Lemma 4.25 (RH)**.** Riemann Hypothesis is: zeros of ζ in $0 \leqslant \operatorname{Re}(s) \leqslant 1$ all lie on line $s = \frac{1}{2} + it$.

Remark. Implication for $\pi(x)$ is this: if RH is true, then $|\pi(x) - \mathrm{Li}(x)| = O(x^{\frac{1}{2}+\epsilon})$. If all roots have $\mathrm{Re}(s) < \frac{1}{2} + \alpha$, then for any $\epsilon > 0$, $|\pi(x) - \mathrm{Li}(x)| < O(x^{\frac{1}{2}+\alpha+\epsilon})$. If there is a root with $\mathrm{Re}(s) = \frac{1}{2} + \alpha$, then $|\pi(x) - \mathrm{Li}(x)|$ will get as big as $x^{\frac{1}{2}+\alpha-\epsilon}$.

We've managed to define ζ for $\mathrm{Re}(s) > 1$ via $\zeta(s) = \sum_{n\geqslant 1} \frac{1}{n^s}$. We've extended the definition to $\mathrm{Re}(s) > 0, s \neq 1$ via $\eta(s) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n^s} = (1 - 2^{s-1})\zeta(s)$.

Question 4.26. What should $\zeta(0)$ be? What is $\lim_{s\to 0^+} \eta(s)$? If $s = \sigma + it$, what is $\lim_{\sigma\to 0^+} \eta(\sigma + it)$? Can we extend the definition of $\zeta(s)$ to $\mathrm{Re}(s) = 0$?

We'll see a better way to extend the definition of ζ.

Recall 4.27 (Analytic Continuation)**.** Given a complex analytic function f on $\mathcal{D} \subseteq \mathbb{C}$, pick $z_0 \in \mathcal{D}$, then in $B_r(z_0) \subseteq \mathcal{D}$, we have a convergent power series

$$\widetilde{f}(z) = \sum_{k=0}^{\infty} (z - z_0)^k \frac{f^{(k)}(z_0)}{k!}.$$

Then $\widetilde{f} : B_{R_{z_0}}(z_0) \to \mathbb{C}$ is an *analytic continuation* of f at z_0, where the *radius of convergence* of the power series is

$$R_{z_0} = \sup\{r > 0 \mid \exists F : B_r(z_0) \to \mathbb{C} \text{ an analytic continuation of } f \text{ at } z_0\}.$$

Then we can extend the function f to $\mathcal{D} \cup \{z \in \mathbb{C} \bigm| |z - z_0| < R_{z_0}\}$.

Example 4.28. Let $f(z) = 1 + z + z^2 + z^3 + \cdots = \frac{1}{1-z}$ if $|z| < 1$. Then $f(z)$ is analytic in $|z| < 1$. But at $z_0 = -\frac{1}{2}$, say, the function $f(z)$ for z near z_0 is given by a power series $f(z) = \sum_{k=0}^{\infty} a_k \frac{(z+\frac{1}{2})^k}{k!}$, which has radius of convergence $\frac{3}{2}$.

We'll show

(a) $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin(\pi s)}$,

(b) $\zeta(s)$ satisfies the functional equation: *the reflection formula*

$$\Gamma\Big(\frac{s}{2}\Big)\zeta(s)\pi^{-\frac{s}{2}} = \Gamma\Big(\frac{1-s}{2}\Big)\zeta(1-s)\pi^{-\frac{1-s}{2}},\ s \in \mathbb{C}.$$

This gives $\zeta(s)$ for $\mathrm{Re}(s) \leqslant \frac{1}{2}$ in terms of $\zeta(s)$ for $\mathrm{Re}(s) \geqslant \frac{1}{2}$ defining $\zeta(s)$ for all of $\mathbb{C}$ except $s = 1$. At $s = 0$, $\Gamma(0)\zeta(0)\pi^{-\frac{1}{2}} = \Gamma(\frac{1}{2})\zeta(1)\pi^{\frac{1}{2}}$ saying nothing about $\zeta(0)$, where $\Gamma(0)$ and $\zeta(1)$ are poles.

Question 4.29. What is $1-1+1-1+1-\cdots$? What is $1+1+1+1+1+\cdots$? What is $1+2+3+4+5+\cdots$? What is $1-2+3-4+5-\cdots$?

4.4 Understand why $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin(\pi s)}$

First, note since $\Gamma(s)$ has pole at $0, -1, -2, \cdots$, $\Gamma(1-s)$ has pole at $1, 2, 3, \cdots$. Note $\sin(\pi s)$ has zeros at $\pm k$, $\frac{\pi}{\sin(\pi s)}$ has poles at $\pm k$.

Second: At $s = \frac{1}{2}$, $\frac{\pi}{\sin(\frac{\pi}{2})} = \pi$. Note $\Gamma(\frac{1}{2}) = \int_0^\infty e^{-t}t^{-\frac{1}{2}}dt$, substituting $t = u^2$, then $dt = 2udu = 2\sqrt{t}du$, i.e., $t^{-\frac{1}{2}}dt = 2du$, so $\Gamma(\frac{1}{2}) = 2\int_0^\infty e^{-u^2}du = \int_{-\infty}^\infty e^{-u^2}du = \sqrt{\pi}$. So at $s = \frac{1}{2}$, the formula is correct.

Hence the formula seems reasonable at least.

Outline to prove Euler's formula $\sin(\pi x) = \pi x\prod_{k=1}^\infty(1-\frac{x^2}{k^2})$. Define $I_n(x) = \int_0^{\frac{\pi}{2}}\cos(xt)\cos^n(t)dt$ for $n \geqslant 0$. Then $I_0(x) = \frac{1}{x}\sin(\frac{\pi x}{2})$ with $I_0(0) = \frac{\pi}{2}$. So by integrating by parts, we have $n(n-1)I_{n-2}(x) = (n^2-x^2)I_n(x)$. Then $\frac{n(n-1)}{n(n-1)}\frac{I_{n-2}(x)}{I_{n-2}(0)} = \frac{n^2-x^2}{n^2}\frac{I_n(x)}{I_n(0)}$, i.e., $\frac{I_{n-2}(x)}{I_{n-2}(0)} = (1-\frac{x^2}{n^2})\frac{I_n(x)}{I_n(0)}$. Hence by induction,

$$\frac{I_0(x)}{I_0(0)} = \prod_{k=1}^n\Big(1-\frac{x^2}{(2k)^2}\Big)\frac{I_{2n}(x)}{I_{2n}(0)},\ n \geqslant 1.$$

Since$1-\cos(xt) \leqslant \frac{x^2t^2}{2}$, check

$$|I_{2n}(0) - I_{2n}(x)| = \int_0^{\frac{\pi}{2}}(1-\cos(xt))\cos^{2n}(t)dt \leqslant \int_0^{\frac{\pi}{2}}\frac{x^2t^2\cos^{2n}(t)}{2}dt \leqslant \frac{x^2}{2}\frac{\pi^2}{4}\int_0^{\frac{\pi}{2}}\cos^{2n}(t)dt \to 0,$$

which implies $\frac{I_{2n}(x)}{I_{2n}(0)} \to 1$ as $n \to \infty$. Moreover, since $\prod_{k=1}^\infty(1-\frac{x^2}{(2k)^2})$ converges for each $x \in \mathbb{C}$, we have $\frac{\frac{1}{x}\sin(\pi\frac{x}{2})}{\frac{\pi}{2}} = \frac{I_0(x)}{I_0(0)} = \prod_{k=1}^\infty(1-\frac{x^2}{(2k)^2})$ for $x \in \mathbb{R}$, i.e., $\sin(\frac{\pi x}{2}) = \frac{\pi x}{2}\prod_{k=1}^\infty(1-\frac{x^2}{(2k)^2})$ for $x \in \mathbb{R}$. Thus, $\sin(\pi x) = \pi x\prod_{k=1}^\infty(1-\frac{x^2}{k^2})$ for $x \in \mathbb{R}$.

Remark. For $x \in \mathbb{C}$, need some work with the integer to prove it directly, or muttes something about analytic continuation which agree on a disk).

Suppose we know f is an analytic function in $|z| < 1+\epsilon$ with $\epsilon > 0$ and we know f on the circle $|z| = 1$. Then Cauchy's integral formula gives

$$f(z_0) = \frac{1}{2\pi i}\int_{|z|=1}\frac{f(z)}{z-z_0}dz.$$

So knowledge of f at $|z| = 1$ gives f for $|z| < 1$, provided that $f(z)$ is analytic on an open subset of $\mathbb{C}$ containing $|z| \leqslant 1$.

Chapter 5

Roadmap

We want to understand $\zeta(s)$ for $\mathrm{Re}(s) \leqslant 0$, so we can have it defined for all $s \in \mathbb{C} \smallsetminus \{1\}$. To do this, we want to prove the reflection formula

$$\Gamma\Big(\frac{s}{2}\Big)\zeta(s)\pi^{-\frac{s}{2}} = \Gamma\Big(\frac{1-s}{2}\Big)\zeta(1-s)\pi^{-\frac{1-s}{2}},\ 0 < \mathrm{Re}(s) < 1.$$

Then since $\zeta(s)$ is analytic in $\mathrm{Re}(s) > 0, s \neq 1$, we'll define $\zeta(s)$ by the reflection formula for $\mathrm{Re}(s) \leqslant 0$ (except $s = 0$ since it is the reflection point of $s = 1$). The resulting function will be the (unique) analytic continuation to $\mathbb{C} \smallsetminus \{0, 1\}$. To define at 0, we consider $\lim_{s\to 0^+} \eta(s) = \lim_{s\to 0^+} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^s}$. The result wil be analytic at 0 as well. To do this, we need to show $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin(\pi s)}$. To show it, we've proved Euler's product for $\sin(x)$: we want to show $\Gamma(s)\Gamma(1-s)$ has the approximate product form to get $\Gamma \neq 0$. To do this, we'll show

$$\Gamma(x) = \lim_{n\to\infty} \frac{(n-1)!n^x}{(x)_n},$$

where $(x)_n = x(x+1)\cdots(x+n-1)$.

Theorem 5.1. *There is a unique F satisfying $F(0) = 1$, $F(x+1) = xF(x)$, $\lim_{n\to\infty} \frac{F(x+n)}{n^x F(n)} = 1$. Furthermore, $F(x) = \lim_{n\to\infty} \frac{(n-1)!n^x}{(x)_n}$. So, if we prove $\lim_{n\to\infty} \frac{\Gamma(x+n)}{n^x\Gamma(n)} = 1$, we'll deduce $\Gamma = F$.*

Proof. Assume $F(x)$ satisfies the conditions. Then $F(x+n) = (x+n-1)(x+n-2)\cdots xF(x) = (x)_nF(x)$. So $\frac{F(x+n)}{n^xF(x)} = \frac{(x)_n}{n^x}$. Hence $\lim_{n\to\infty} \frac{(x)_n}{n^x}\frac{F(x)}{F(n)} = \lim_{n\to\infty} \frac{F(x+n)}{n^xF(n)} = 1$. So $F(x) = \lim_{n\to\infty} \frac{F(n)n^x}{(x)_n} = \lim_{n\to\infty} \frac{(n-1)!n^x}{(x)_n}$. This shows that if the limit exists, F is unique. We'll show Γ satisfies $\lim_{n\to\infty} \frac{\Gamma(x+n)}{n^x\Gamma(n)} = 1$ for $0 < x < 1$, hence the limit does exist and Γ satisfies $\Gamma(x) = \lim_{n\to\infty} \frac{(n-1)!n^x}{(x)_n}$.

Indeed, let $J = \frac{\Gamma(x+n)}{n^x(n-1)!} = \frac{1}{n^x(n-1)!}\int_0^\infty e^{-t}t^{x+n-1}dt$. Replace t by ny, $dt = ndy$, we obtain

$$J = \frac{1}{n^x(n-1)!}\int_0^\infty e^{-ny}y^{x+n-1}n^{x+n-1}ndy = \frac{n^n}{(n-1)!}\int_0^\infty e^{-ny}y^{x+n-1}dy.$$

We want to show this converges to 1 as $n \to \infty$. Since $0 < x < 1$,

$$\frac{n^n}{(n-1)!}\Big(\int_0^1 e^{-ny} y^n dy + \int_1^\infty e^{-ny} y^{n-1} dy\Big) < J < \frac{n^n}{(n-1)!}\Big(\int_0^1 e^{-ny} y^{n-1} dy + \int_1^\infty e^{-ny} y^n dy\Big).$$

Note

$$\begin{aligned}\int_0^1 e^{-ny} y^n dy + \int_1^\infty e^{-ny} y^{n-1} dy &= \int_0^1 y^n d\Big(-\frac{1}{n}e^{-ny}\Big) + \int_1^\infty e^{-ny} y^{n-1} dy \\ &= -\frac{1}{n}e^{-ny}y^n\Big|_0^1 + \int_1^\infty e^{-ny} y^{n-1} dy \\ &= -\frac{e^{-n}}{n} + \int_0^\infty e^{-ny} y^{n-1} dy = \frac{(n-1)!}{n^n} - \frac{e^{-n}}{n},\end{aligned}$$

and

$$\begin{aligned}\int_0^1 e^{-ny} y^{n-1} dy + \int_1^\infty e^{-ny} y^n dy &= \int_0^1 e^{-ny} y^{n-1} dy + \int_1^\infty y^n d\Big(-\frac{1}{n}e^{-ny}\Big) \\ &= \int_0^1 e^{-ny} y^{n-1} dy - \frac{1}{n}e^{-ny}y^n\Big|_1^\infty + \int_1^\infty y^{n-1} e^{-ny} dy \\ &= \frac{e^{-n}}{n} + \int_0^\infty y^{n-1} e^{-ny} dy = \frac{(n-1)!}{n} + \frac{e^{-n}}{n}.\end{aligned}$$

So we get $\frac{n^n}{(n-1)!}(\frac{(n-1)!}{n^n} - \frac{e^{-n}}{n}) < J < \frac{n^n}{(n-1)!}(\frac{(n-1)!}{n^n} + \frac{e^{-n}}{n})$, i.e., $1 - (\frac{e}{n})^n \frac{1}{n!} < J < 1 + (\frac{n}{e})^n \frac{1}{n!}$. Since $(\frac{n}{e})^n \frac{1}{n!} \sim \frac{1}{\sqrt{2\pi n}}$, $\lim_{n\to\infty} J = 1$. □

Remark. We can also do the above proof through considering the difference between upper and lower bounds by integrating $\int_0^1 e^{-ny}(y^{n-1} - y^n)dy = \frac{1}{n}e^{-ny}y^n\Big|_0^1 = \frac{1}{n}e^{-n}$.

Intuitively, how big is $\int_0^1 e^{-ny}(y^{n-1} - y^n)dy$? Approximate the integral near its maximum value: let $0 = \frac{d}{dy}(e^{-ny}(y^{n-1} - y^n)) = e^{-ny}y^{n-2}(ny^2 - 2ny + (n-1))$, then $y = 1 \pm \frac{1}{\sqrt{n}}$. How big is the integrand at $y_1 = 1 - \frac{1}{\sqrt{n}}$? Note

$$e^{-ny_1} y_1^{n-1}(1 - y_1) = e^{-n+\sqrt{n}}\Big(1 - \frac{1}{\sqrt{n}}\Big)^{n-1}\frac{1}{\sqrt{n}} \approx e^{-n}e^{\sqrt{n}}e^{-\sqrt{n}}\frac{1}{\sqrt{n}} \approx \frac{e^{-n}}{\sqrt{n}},$$

where

$$\begin{aligned}y_1^{n-1} &= \exp\bigg((n-1)\log\Big(1 - \frac{1}{\sqrt{n}}\Big)\bigg) = \exp\bigg((n-1)\Big(\frac{1}{\sqrt{n}} + \frac{1}{2n} + \frac{1}{3n^{\frac{3}{2}}} + \cdots\Big)\bigg) \\ &= \exp\bigg(\sqrt{n} + \frac{1}{2} + \Big(\frac{1}{3} - 1\Big)\frac{1}{\sqrt{n}} + O\Big(\frac{1}{n}\Big)\bigg) \approx \exp\Big(\sqrt{n} + \frac{1}{2}\Big).\end{aligned}$$

Now expand the integrand about $y = 1 - \frac{1}{\sqrt{n}}$, it looks approximately normal, with width about $\frac{1}{\sqrt{n}}$, so we expect integral to be about $\frac{c}{\sqrt{n}}\frac{e^{-n}}{\sqrt{n}} = \frac{ce^{-n}}{n}$. Similarly, for $\int_1^\infty e^{-n}(y^n - y^{n-1})dy$.

So now apply the results to get

$$\begin{aligned}\Gamma(x)\Gamma(1-x) &= \lim_{n\to\infty} \frac{(n-1)!n^x}{(x)_n}\frac{(n-1)!n^{1-x}}{(1-x)_n} = \frac{(n-1)!^2 n}{x(1+x)\cdots(n-1+x)(1-x)\cdots(n-x)} \\ &= \lim_{n\to\infty} \frac{(n-1)!^2 n}{x(1+x)(1-x)(2+x)(2-x)\cdots(n-1+x)(n-1+x)(n-x)} \\ &= \lim_{n\to\infty} \frac{(n-1)!^2 n}{x(1-x^2)(2^2-x^2)\cdots((n-1)^2-x^2)(n-x)} \\ &= \lim_{n\to\infty} \frac{1}{x(1-\frac{x^2}{1})(1-\frac{x^2}{2^2})\cdots(1-\frac{x^2}{(n-1)^2})(1-\frac{x}{n})} = \frac{1}{x}\prod_{k=1}^{\infty}\frac{1}{1-\frac{x^2}{k^2}}.\end{aligned}$$

So

$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin(\pi x)},\ x \in \mathbb{R}.$$

Since $\Gamma(x)$ has simple poles at $0, -1, -2, \cdots$ and $\Gamma(1-x)$ has simple poles at $1, 2, 3, \cdots$, the poles of $\Gamma(x)\Gamma(1-x)$ are at precisely the poles of $\frac{\pi}{\sin(\pi x)}$. Consequently, $\Gamma(x)$ cannot be zero. Indeed, if $\Gamma(x) = 0$, then $x \notin \mathbb{Z}$, also since $\frac{1}{\sin(\pi x)} \neq 0$, $\Gamma(1-x)$ has a pole there, but we know the poles of $\Gamma(1-x)$ correspond to zeros of $\sin(\pi x)$, a contradiction. Hence $\Gamma(x) \neq 0$ for any $x \in \mathbb{C}$.

$\Gamma(s) = \int_0^\infty e^{-t}t^{s-1}dt$. So by a change of variable, $\int_0^\infty e^{-nt}t^{s-1}dt = \frac{1}{n^s}\Gamma(s)$. Hence

$$\sum_{n=1}^{\infty}\frac{1}{n^s}\Gamma(s) = \sum_{n=1}^{\infty}\int_0^\infty e^{-nt}t^{s-1}dt.$$

Remark. Integrands are positive if s is real, but more care is needed otherwise.

For $\mathrm{Re}(s) > 1$,

$$\zeta(s)\Gamma(s) = \int_0^\infty \sum_{n=1}^{\infty} e^{-nt}t^{s-1}dt = \int_0^\infty t^{s-1}\frac{e^{-t}}{1-e^{-t}}dt = \int_0^\infty \frac{t^{s-1}}{e^t-1}dt.$$

Exercise 5.2. Prove carefully that $\int_0^1 \frac{t^{s-1}}{e^t-1}dt$ converges for $\mathrm{Re}(s) > 1$, and $\int_1^\infty \frac{t^{s-1}}{e^t-1}dt$ converges for all $s \in \mathbb{C}$.

Fact 5.3.

$$\frac{t}{e^t-1} = \frac{t}{t+\frac{t^2}{2!}+\frac{t^3}{3!}+\cdots} = \frac{1}{1+\frac{t}{2!}+\frac{t^2}{3!}+\cdots} = \sum_{k=1}^{\infty}\frac{B_k t^k}{k!} = 1 - \frac{t}{2} + \sum_{k=1}^{\infty}\frac{B_{2k}t^{2k}}{(2k)!},$$

where B_{2k}'s are the *Bernoulli number*, which converges in the disk $|t| < 2\pi$.

Since the zeros of $e^z - 1$ are at $z = 2\pi ki, k \in \mathbb{Z}$, the poles of $\frac{z}{e^z-1}$ are at $2\pi ki, k = 1, 2, \cdots$. So the power series converges for $|t| < 2\pi$.

Note

$$\int_0^1 \frac{t^{s-1}}{e^t-1}dt = \int_0^1\left(t^{s-2} - \frac{t^{s-1}}{2} + \sum_{k=1}^{\infty}\frac{B_{2k}t^{2k+s-2}}{(2k)!}\right)dt = \frac{1}{s-1} - \frac{1}{2s} + \sum_{k\geqslant 1}\frac{B_{2k}}{(2k)!}\frac{1}{2k+s-1},$$

which is valid for all $s \in \mathbb{C}, s \neq -2k+1, 0, 1$ (for all $s \in \mathbb{C}$ if we allow poles). So we can define

$$\zeta(s)\Gamma(s) = \frac{1}{s-1} - \frac{1}{2s} + \sum_{k \geqslant 1} \frac{B_{2k}}{(2k)!} \frac{1}{2k+s-1} + \int_1^\infty \frac{t^{s-1}}{e^t-1} dt.$$

Since we have a valid expression for $\Gamma(s), s \in \mathbb{C} \setminus \{0, -1, -2, -3, \cdots\}$ which is non-zero, we can use this to define an analytic function $\zeta(s)$ with a simple pole at $s = 1$.

Remark. Note we don't have a good handle on the value of $\zeta(-2k)\Gamma(-2k)$.

By the following proposition, we see that does give an analytic (and hence the analytic) continuation of ζ to $\mathbb{C} \setminus \{1\}$ or of $(s-1)\zeta(s)$ to $\mathbb{C}$.

Proposition 5.4. Suppose $f(z) = \frac{a}{z-z_0} + f_1(z)$ and $g(z) = \frac{b}{z-z_0} + g_1(z)$, where $f_1(z)$ is analytic near z_0 and $g_1(z)$ is analytic near z_0. Then

$$\lim_{z \to z_0} \frac{f(z)}{g(z)} = \lim_{z \to z_0} \frac{\frac{a}{z-z_0} + f_1(z)}{\frac{b}{z-z_0} + g_1(z)} = \lim_{z \to z_0} \frac{a + (z-z_0)f_1(z)}{b + (z-z_0)g_1(z)} = \frac{a + 0f_1(z_0)}{b + 0g_1(z_0)} = \frac{a}{b}.$$

Note 5.5. The pole at 0 in this expression comes from $\Gamma(s)$ since 0 is a simple pole of $\Gamma(s)$ and $\zeta(s)\Gamma(s)$.

Since $\Gamma(s)$ has poles at even negative integers but $\Gamma(s)\zeta(s)$ does not, we must have $\zeta(-2k) = 0$ for $k = 1, 2, 3, \cdots$.

Theorem 5.6. $\zeta(0) = -\frac{1}{2} = B_1$ *and*

$$\zeta(-n) = -\frac{B_{n+1}}{(n+1)!}, \forall n \geqslant 1.$$

Proof. By the expression of $\zeta(s)\Gamma(s)$, it has a pole $\frac{B_{2k}}{(2k)!}\frac{1}{2k+s-1}$ in the vicinity of $s = -(2k-1)$. Since $\Gamma(s)$ has a pole $\frac{(-1)^{2k-1}}{2k-1+s} = \frac{-1}{2k-1+s}$ at $s = -(2k-1)$, we know that

$$\zeta(-(2k-1)) = \frac{\zeta \cdot \Gamma(-(2k-1))}{\Gamma(-(2k-1))} = \frac{\frac{B_{2k}}{(2k)!}\frac{1}{2k+s-1}}{\frac{-1}{2k-1+s}} = -\frac{B_{2k}}{(2k)!}.$$

Moreover, $\zeta(-2k) = 0 = -\frac{0}{(2k+1)!} = -\frac{B_{2k+1}}{(2k+1)!}$ for $k \geqslant 1$.

Similarly, $\zeta(0) = \frac{-\frac{1}{2}}{\frac{(-1)^0}{0!}} = -\frac{1}{2} = B_1$. ☐

We know $\zeta(s)$ diverges at $s = 1$, and we've seen Euler's evaluation of $\zeta(2k), k \geqslant 1$. $\zeta(2) = \frac{\pi^2}{6}$, $\zeta(4) = \frac{\pi^4}{90}$, $\zeta(2k) = \frac{r}{s}\pi^{2k}$ with $\frac{r}{s} \in \mathbb{Q}$. The expression involving Bernoulli numbers:

$$\zeta(2k) = (-1)^{k+1}\frac{B_{2k}(2\pi)^{2k}}{2(2k)!}.$$

So $\zeta(6) = \frac{\pi^6}{945}$, $\zeta(8) = \frac{\pi^8}{9450}$, $\zeta(10) = \frac{\pi^{10}}{93555}$, $\zeta(12) = \frac{691\pi^{12}}{638512875}$, $\zeta(14) = \frac{2\pi^{14}}{18243225}$. So $\zeta(2k)$ is irrational, and indeed, transcendental!

1970s, $\zeta(3)$ was shown by Aperg to be irrational. Believed by all to be transcendental. Believed by all not to be a rational multiple of π^3.

Question 5.7. What about $\zeta(5), \zeta(7), \zeta(9), \zeta(11)$? Zudelin (2001) showed that at least one of these four values must be irrational.

Digression: 33 is the sum of 3 cubes. Is 42?

Our gosl is to relate $\zeta(s)$ and its zeros to $\pi(x)$ and related functions.

Fact 5.8. If we approximate $f(x)$ on $[a, b]$ by fourier series, if $f(x)$ has a discontinuity at c, then our approximation looks like average of the end points.

If we want a function which has a nice approximation by something like Fourier series, it makes sense to have jumps average at discontinuities. So, for example, we'll define

$$\Pi(x) = \frac{1}{2}\big(\#\{n < x \mid n \text{ is prime}\} + \#\{n \leqslant x \mid n \text{ is prime}\}\big).$$

So $\pi(\frac{1}{2}) = 0$, $\pi(2) = \frac{1}{2}$, $\pi(2\frac{1}{2}) = 1$, $\pi(3) = 1\frac{1}{2}$, $\pi(4) = 2$ and so on.

Fact 5.9.

$$\Pi(x) = \sum_{p<x} 1 + \frac{1}{2}\mathbb{1}_{\{x \text{ is prime}\}}.$$

Definition 5.10. Define the *Jacobi Theta function* by

$$\theta(t) = \sum_{k=-\infty}^{\infty} e^{-\pi k^2 t}.$$

Remark. θ converges for $t > 0$. For small t, this looks like a discrete analog of the normal integral. For large t, it looks like $1 + 2e^{-\pi t} \approx 1$.

Theorem 5.11.

$$\theta\Big(\frac{1}{t}\Big) = \sqrt{t}\theta(t).$$

Remark. Is this believable? If t is big, $\theta(t) \approx 1$, $\theta(\frac{1}{t})$ is a sum of terms for which $\frac{\pi k^2}{t}$ is not too big, that is, k^2 is not too much bigger than t, that is, k is on the order of $\sqrt{t}$. If we carefully approximate $\theta(\frac{1}{t})$ by a normal integral, if t is big, $\theta(t) \approx \sqrt{t}$. Remarkably, it is not just approximation true, it is exact!

Fact 5.12.

$$\left(\frac{1}{10^5}\sum_{k=-\infty}^{\infty} e^{-\frac{k^2}{10^{10}}}\right)^2 \neq \pi.$$

But it is correct to 42 billion digits!

Proof. Set $-\frac{k^2}{10^{10}} = -\pi k^2 t$, i.e., $\pi t = \frac{1}{10^{10}}$, i.e, $t = \frac{1}{10^{10}\pi}$. Since $\frac{1}{t} = 10^{10}\pi$ is large,

$$10^{-5}\theta(t) = 10^{-5}\frac{1}{\sqrt{t}}\theta\Big(\frac{1}{t}\Big) = 10^{-5}10^5\sqrt{\pi}\theta\Big(\frac{1}{t}\Big) \approx \sqrt{\pi}\big(1 + 2e^{-\frac{\pi}{t}}\big) \approx \sqrt{\pi}\big(1 + 2e^{-10^{10}}\big).$$

So

$$\left(\frac{1}{10^5}\sum_{k=-100}^{100} e^{-\frac{k^2}{10^{10}}}\right)^2 \approx \pi\big(1 + 4e^{-10^{10}}\big),$$

which turns out to mean we get π to about 42 billion digits. $\square$

Fact 5.13.

$$P(\text{100 heads and 100 tails tossing a fair coin 200 times}) \approx \frac{1}{10\sqrt{\pi}} = \frac{1}{\sqrt{100\pi}}.$$

5.1 Mini Polemic

$\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}dt$ is the wrong way to think about this! dt is the wrong differential! Better: use $\frac{dt}{t}$, $\Gamma(x) = \int_0^\infty e^{-t}t^x \frac{dt}{t}$. Reason: consider the change of variable $u = t^\alpha$. Then $du = \alpha t^{\alpha-1}dt = \alpha t^\alpha \frac{dt}{t} = \alpha u \frac{dt}{t}$. So the change of variables $u = t^\alpha$ introduces a change of differentials $\frac{dt}{t} = \frac{1}{\alpha}\frac{du}{u}$.

We'll want to consider for $t > 0$, $f(t) = \sum_{k=1}^\infty e^{-\pi k^2 t} = \frac{1}{2}(\theta(t) - 1)$ or $\theta(t) = 2f(t) + 1$. Note

$$f(t) = \frac{1}{2}\left(\frac{1}{\sqrt{t}}\theta\Big(\frac{1}{t}\Big) - 1\right) = \frac{1}{2}\left(\frac{1}{\sqrt{t}}\Big(2f\Big(\frac{1}{t}\Big) + 1\Big) - 1\right) = \frac{1}{\sqrt{t}}f\Big(\frac{1}{t}\Big) + \frac{1}{2\sqrt{t}} - \frac{1}{2}.$$

Let $R(s) = \pi^{-\frac{s}{2}}\Gamma(\frac{s}{2})\zeta(s)$. Start with $\Gamma(\frac{s}{2}) = \int_0^\infty e^{-t}t^{\frac{s}{2}}\frac{dt}{t}$, so $\frac{1}{n^s}\pi^{-\frac{s}{2}}\Gamma(\frac{s}{2}) = \int_0^\infty e^{-\pi n^2 t}t^{\frac{s}{2}}\frac{dt}{t}$, hence for $\mathrm{Re}(s) > 1$,

$$\begin{aligned}
R(s) = \pi^{-\frac{s}{2}}\Gamma\Big(\frac{s}{2}\Big)\zeta(s) &= \sum_{n=1}^\infty \frac{1}{n^s}\pi^{-\frac{s}{2}}\Gamma\Big(\frac{s}{2}\Big) = \int_0^\infty \left(\sum_{n=1}^\infty e^{-\pi n^2 t}\right)t^{\frac{s}{2}}\frac{dt}{t} = \int_0^\infty f(t)t^{\frac{s}{2}}\frac{dt}{t} \\
&= \int_0^1 f(t)t^{\frac{s}{2}}\frac{dt}{t} + \int_1^\infty f(t)t^{\frac{s}{2}}\frac{dt}{t} = \int_1^\infty f\Big(\frac{1}{t}\Big)t^{-\frac{s}{2}}\frac{dt}{t} + \int_1^\infty f(t)t^{\frac{s}{2}}\frac{dt}{t} \\
&= \int_1^\infty \left(-\frac{1}{2} + \frac{\sqrt{t}}{2} + \sqrt{t}f(t)\right)t^{-\frac{s}{2}}\frac{dt}{t} + \int_1^\infty f(t)t^{\frac{s}{2}}\frac{dt}{t} = \frac{1}{s(s-1)} + \int_1^\infty f(t)\big(t^{\frac{1-s}{2}} + t^{\frac{s}{2}}\big)\frac{dt}{t}.
\end{aligned}$$

So by symmetry $R(s) = R(1-s)$ for $\mathrm{Re}(s) > 1$ or $\mathrm{Re}(s) < 0$. Hence

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

Definition 5.14. Define *Riemann Xi function* by

$$\xi(s) = s(1-s)R(s).$$

Then $\xi(s) = \xi(1-s)$ and $\xi(s)$ is an entire function.

5.2 Zeros of ζ

Where can ζ be zero? We know $s = -2, -4, -6, \cdots$ are "the trivial zeros" of ζ. No other zeros coming from poles of Γ. Can we have $\zeta(s) = 0$ with $\mathrm{Re}(s) > 1$? No: indeed, if $\zeta(s) = 0$, then $\frac{1}{\zeta(s)}$ is a pole, but $\frac{1}{\zeta(s)} = \sum_{n=1}^\infty \frac{\mu(n)}{n^s}$ converges absolutely if $\mathrm{Re}(s) > 1$, so $\frac{1}{\zeta(s)}$ has no poles with $\mathrm{Re}(s) > 1$, i.e., $\zeta(s)$ has no zeros with $\mathrm{Re}(s) > 1$. Hence, other than $-2, -4, -6, \cdots$, the reflection formula tells $\zeta(s)$ has no zeros with $\mathrm{Re}(s) < 0$. Thus, all non-trial zeros of $\zeta(s)$ have $0 \leqslant \mathrm{Re}(s) \leqslant 1$. Furthermore, all non-trivial zeros ρ either have $\rho = \frac{1}{2} + it$, or come in pairs $\rho = \sigma + it$ with $\sigma > \frac{1}{2}$, $\rho' = 1 - \sigma + it$ with $1 - \sigma < \frac{1}{2}$.

Short term goal: factor $\zeta(s)$ as a product of zeros plus necessary extra terms, and to then consider $\frac{\zeta'(s)}{\zeta(s)} = \frac{d}{ds}(\log\zeta(s))$, which is naturally a sum over zeros of $\zeta(s)$.

Theorem 5.15 (Weierstrass). *If $f(z)$ is an entire function, and the zeros of f (with multiplicity) are $\alpha_1, \alpha_2, \cdots$ with $\alpha_n \to \infty$ as $n \to \infty$, then there exists a polynomial $g(z)$ and an integer k so that $f(z) = z^k e^{g(z)} \prod_{\alpha_j} E_{m_j}(\frac{z}{\alpha_j})$, where for each root α_j, we have an integer m_j, and $E_m(\frac{z}{\alpha}) =$*

$$\left(1 - \tfrac{z}{\alpha}\right) \exp\Big(\underbrace{\frac{z}{\alpha} + \frac{z^2}{2\alpha^2} + \frac{z^3}{3\alpha^3} + \cdots + \frac{z^m}{m\alpha^m}}_{\text{truncation at } \frac{z^m}{m} \text{ of } \log(1-\frac{z}{\alpha})}\Big).$$

Example 5.16.

$$\sin(\pi z) = \pi z \prod_{k=-\infty, k\neq 0}^{\infty} \Big(1 - \frac{z}{k}\Big) = \pi z \prod_{k=-\infty, k\neq 0}^{\infty} \Big(1 - \frac{z}{k}\Big) e^{\frac{z}{k}} = z e^{\log \pi} \prod_{k=-\infty, k\neq 0}^{\infty} \Big(1 - \frac{z}{k}\Big) \exp\Big(\frac{z}{k}\Big).$$

When we try to write $\sin(\pi z) = \pi z \prod_{k=-\infty,k\neq 0}^{\infty}(1 - \frac{z}{k})$, we had a product whose corresponding series was only conditionally convergent. Introducing the $e^{\frac{z}{k}}$ term means $(1 - \frac{z}{k})e^{\frac{z}{k}} = 1 - \frac{z^2}{2k^2} + O(\frac{z^3}{k^3})$.

In general, we need $\exp(\frac{z}{\alpha} + \frac{z^2}{2\alpha^2} + \frac{z^3}{3\alpha^3} + \cdots + \frac{z^m}{m\alpha^m})$ as a correction factor.

Example 5.17.

$$\frac{1}{\Gamma(s+1)} = e^{\gamma s} \prod_{n=1}^{\infty} \Big(1 + \frac{s}{n}\Big) e^{-\frac{s}{n}}.$$

Example 5.18. Since we've seen that $(s-1)\zeta(s)$ is entire, by Weierstrass, (plus some work)

$$(s-1)\zeta(s) = \frac{1}{2}\Big(\frac{2\pi}{e}\Big)^s \prod_{n=1}^{\infty}\Big(1 + \frac{s}{2n}\Big)e^{-\frac{s}{2n}} \prod_{\rho}\Big(1 - \frac{s}{\rho}\Big)e^{\frac{s}{\rho}},$$

where $\frac{1}{2}(\frac{2\pi}{e})^s = \exp(-\log 2 + s\log(\frac{2\pi}{e})) = \exp(g(s))$, $\prod_{n=1}^{\infty}(1 + \frac{s}{2n})e^{-\frac{s}{2n}}$ is the product over the trivial zeros $-2, -4, -6, \cdots$, and $\prod_\rho (1 - \frac{s}{\rho})e^{\frac{s}{\rho}}$ is the product over the zeros with $\rho = \sigma + it, 0 \leqslant \sigma \leqslant 1$, the nontrivial zeros.

Since $(s-1)\zeta(s)$ is entire, it has a power series expansion valid in $\mathbb{C}$, which we could obtain by multiplying out the terms in the two infinite products. Also, $\frac{1}{\Gamma(s)} = \frac{s}{\Gamma(s+1)} = s e^{\gamma s} \prod_{n=1}^{\infty}(1 + \frac{s}{n})e^{-\frac{s}{n}}$, so $\frac{1}{\Gamma(\frac{s}{2})} = \frac{s}{2} e^{\gamma \frac{s}{2}} \prod_{n=1}^{\infty}(1 + \frac{s}{2n})e^{-\frac{s}{2n}}$, hence

$$(s-1)\zeta(s) = \frac{(\frac{2\pi}{e})^s}{s e^{\gamma \frac{s}{2}}} \cdot \frac{1}{\Gamma(\frac{s}{2})} \cdot \prod_{\rho}\Big(1 - \frac{s}{\rho}\Big)e^{\frac{s}{\rho}}.$$

Note there is a statement here about covergence. It implies that the roots ρ cannot grow too slowly: in particular, $(1 - \frac{z}{\rho})e^{\frac{z}{\rho}} = 1 - \frac{z^2}{2\rho^2} + O(\frac{z^3}{\rho^3})$, so for fixed z, in order to have convergence, we need $\sum_\rho \frac{1}{|\rho^2|}$ converges. So $|\rho_n|$ should grow substantially faster than $n^{\frac{1}{2}}$.

Let $N_T = \#\text{roots in } [0,1] \times [-T, T]$, then we cannot have N_T as big as T^2.

Definition 5.19. Define the *(second) Chebyshev ψ-function* by

$$\psi(x) = \sum_{p^k \leqslant x} \log p = \sum_{n \leqslant x} \Lambda(n).$$

Remark. This formula relates $\psi(x)$ to the roots ρ of $\zeta(s)$.

Since $(s-1)\zeta(s)$ is entire and $(s-1)\zeta(s)$ is real for real s, we have if ρ is a root of ζ in the *critical strip* $\mathbb{H} = \{0 < \mathrm{Re}(s) < 1\}$, so is $\bar{\rho}$. So we can write the product as

$$\prod_{\rho\in\mathbb{H}}\Big(1-\frac{s}{\rho}\Big)\Big(1-\frac{s}{\bar{\rho}}\Big)e^{\frac{s}{\rho}+\frac{s}{\bar{\rho}}} = \prod_{\rho\in\mathbb{H}}\left(1-\Big(\frac{s}{\rho}+\frac{s}{\bar{\rho}}\Big)+\frac{s^2}{\rho\bar{\rho}}\right)e^{\frac{s}{\rho}+\frac{s}{\bar{\rho}}},$$

where $s(\frac{1}{\rho}+\frac{1}{\bar{\rho}}) = s\frac{\rho+\bar{\rho}}{\rho\bar{\rho}}$. If $\rho = \frac{1}{2}+it$, i.e., it is a root as predicted by RH, then $\rho+\bar{\rho} = 1$, so $\frac{s}{\rho}+\frac{s}{\bar{\rho}} = \frac{s}{\rho\bar{\rho}}$, ($\rho\bar{\rho} = \frac{1}{4}+t^2$) hence

$$\prod_{\rho\in\mathbb{H}}\Big(1-\frac{s}{\rho}\Big)\Big(1-\frac{s}{\bar{\rho}}\Big)e^{\frac{s}{\rho}+\frac{s}{\bar{\rho}}} = \prod_{\rho\in\mathbb{H}}\left(1-\frac{s}{\rho\bar{\rho}}+\frac{s^2}{\rho\bar{\rho}}\right)e^{\frac{s}{\rho\bar{\rho}}}.$$

Remark. If $\rho = \sigma+it$ with $\sigma \neq \frac{1}{2}$, then the value $\rho' = (\frac{1}{2}-\sigma)+it$ is also a root by reflection formula, as are $\bar{\rho}$ and $\overline{\rho'}$.

Note

$$\frac{d}{ds}\big(\log((s-1)\zeta(s))\big) = \frac{d}{ds}\left(\log\bigg(\frac{1}{2}\Big(\frac{2\pi}{e}\Big)^s\prod_{n=1}^{\infty}\Big(1+\frac{s}{2n}\Big)e^{-\frac{s}{2n}}\prod_{\rho}\Big(1-\frac{s}{\rho}\Big)e^{\frac{s}{\rho}}\bigg)\right),$$

i.e.,

$$\begin{aligned}\frac{1}{s-1}+\frac{\zeta'(s)}{\zeta(s)} &= \log\Big(\frac{2\pi}{e}\Big)+\sum_{n=1}^{\infty}\Big(\frac{1}{2n}\frac{1}{1+\frac{s}{2n}}-\frac{1}{2n}\Big)+\sum_{\rho}\Big(-\frac{1}{\rho}\frac{1}{1-\frac{s}{\rho}}+\frac{1}{\rho}\Big)\\ &= \log(2\pi)-1+\sum_{n=1}^{\infty}\Big(\frac{1}{s+2n}-\frac{1}{2n}\Big)+\sum_{\rho}\Big(\frac{1}{s-\rho}+\frac{1}{\rho}\Big),\end{aligned}$$

i.e.,

$$\frac{\zeta'(s)}{\zeta(s)} = \log(2\pi)-\frac{s}{s-1}+\sum_{n=1}^{\infty}\Big(\frac{1}{s+2n}-\frac{1}{2n}\Big)+\sum_{\rho}\Big(\frac{1}{s-\rho}+\frac{1}{\rho}\Big).$$

But if $\mathrm{Re}(s) > 1$, $\zeta(s) = \prod_p(1-\frac{1}{p^s})^{-1}$, and so

$$\begin{aligned}\frac{\zeta'(s)}{\zeta(s)} &= \frac{d}{ds}\big(\log(\zeta(s))\big) = \sum_p\frac{d}{ds}\Big(-\log\Big(1-\frac{1}{p^s}\Big)\Big) = \sum_p\frac{1}{1-\frac{1}{p^s}}\frac{d}{ds}(e^{-s\log p}) = -\sum_p\frac{p^{-s}\log p}{1-p^{-s}}\\ &= -\sum_p(\log p)\Big(\frac{1}{1-p^{-s}}-1\Big) = -\sum_p\log p\sum_{k=1}^{\infty}\frac{1}{(p^s)^k} = -\sum_{p^k}\frac{\log p}{p^{ks}} = -\sum_n\frac{\Lambda(n)}{n^s}.\end{aligned}$$

Earlier we wrote $\psi(x) = \sum_{n\leqslant x}\Lambda(n)$. This is fine if x is not a prime power. We need to smooth out the jumps.

Definition 5.20. Define

$$\Psi(x) = \frac{1}{2}\left(\sum_{n<x}\Lambda(n)+\sum_{n\leqslant x}\Lambda(n)\right) = \sum_{n<x}\Lambda(n)+\frac{1}{2}\Lambda(x).$$

For now all summations include half of the final term. e.g., $\Psi(x) = \sum_{n<x} \Lambda(n) = \sum_{p^k<x} \log p$. Let

$$\Pi(x) := \sum_{p^k<x} \frac{1}{k} = \sum_{p<x} 1 + \frac{1}{2} \sum_{p<x^{\frac{1}{2}}} 1 + \frac{1}{3} \sum_{p<x^{\frac{1}{3}}} 1 + \cdots = \sum_{k=1}^{\infty} \frac{1}{k} \pi(x^{\frac{1}{k}}) = \pi(x) + \sum_{k=2}^{\lfloor \log_2 x \rfloor} \frac{1}{k} \pi(x^{\frac{1}{k}}).$$

By Chebyshev, there are constants $C_1, C_2 > 0$ so that for all $x \geqslant 2$, $C_1 \frac{x}{\log x} < \pi(x) < C_2 \frac{x}{\log x}$. So

$$\begin{aligned} |\Pi(x) - \pi(x)| = \sum_{k=2}^{\lfloor \log_2 x \rfloor} \frac{1}{k} \pi(x^{\frac{1}{k}}) &< \frac{1}{2}\pi(x^{\frac{1}{2}}) + \pi(x^{\frac{1}{3}}) \left(\frac{1}{3} + \cdots + \frac{1}{\lfloor \log_2 x \rfloor} \right) \\ &< \frac{1}{2}\pi(x^{\frac{1}{2}}) + \pi(x^{\frac{1}{3}}) \log(\log_2 x) < \frac{1}{2} C_2 \frac{x^{\frac{1}{2}}}{\frac{1}{2}\log x} + C_2 \frac{x^{\frac{1}{3}}}{\frac{1}{3}\log x} \log(\log_2 x). \end{aligned}$$

Hence $|\Pi(x) - \pi(x)| = O(\frac{x^{\frac{1}{2}}}{\log x})$. This is smaller than the difference between $\pi(x)$ and $\mathrm{Li}(x)$ even if the Riemann Hypothesis is true. So RH has same implication for $|\Pi(x) - \mathrm{Li}(x)|$, as it does for $|\pi(x) - \mathrm{Li}(x)|$.

Note by Mobius inversion,

$$\pi(x) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \Pi(x^{\frac{1}{n}}) = \Pi(x) - \Pi(x^{\frac{1}{2}}) - \Pi(x^{\frac{1}{3}}) - \Pi(x^{\frac{1}{5}}) + \Pi(x^{\frac{1}{6}}) + \cdots.$$

We now wish to relate Ψ and Π. If we let $a_n = \Lambda(n)$, $b_n = \frac{1}{\log n}$, then $a_n b_n = \begin{cases} 0 & \text{if } n \neq p^k \\ \frac{1}{k} & \text{if } n = p^k \end{cases}$, so $\sum_{n<x} a_n b_n = \Pi(x)$ (again half the term for x). Hence (taking care to check how Abelian summations behave with "half the x-term") we can show

$$\begin{aligned} \Pi(x) &= A_x b_{\lfloor x \rfloor} + \sum_{k<x} A_k (b_k - b_{k+1}) = \frac{\Psi(x)}{\log x} + \sum_{k<x} \Psi(k) \left(\frac{1}{\log k} - \frac{1}{\log(k+1)} \right) \\ &= \frac{\Psi(x)}{\log x} + \sum_{k<x} \frac{\Psi(k)}{\log k \log(k+1)} \log \frac{k+1}{k} \approx \frac{\Psi(x)}{\log x} + \sum_{k<x} \frac{1}{\log k \log(k+1)} = \frac{\Psi(x)}{\log x} + O\left(\frac{x}{(\log x)^2} \right). \end{aligned}$$

So if we show $\Psi(x) \sim x$, then $\Pi(x) \sim \frac{x}{\log x}$, so $\pi(x) \sim \frac{x}{\log x}$ and $\psi(x) \sim x$.

So our local goal is to Prove von Mangoldt's formula

$$\Psi(x) = x - \sum_{\rho} \frac{x^{\rho}}{\rho} - \frac{1}{2} \log(1 - \frac{1}{x^2}) - \log(2\pi),$$

and show the contribution from $\sum_{\rho} \frac{x^{\rho}}{\rho}$ is $o(x)$. This will prove the Prime Number Theorem.

Definition 5.21. The *Mellin transform* $M : f \to Mf$ is given by

$$Mf = \int_1^{\infty} f(x) x^{-s} \frac{dx}{x}.$$

Remark. M is linear, so we can compute $M\Psi$ through $\frac{1}{\rho} M x^{\rho}, -\frac{1}{2} M(\log(1 - \frac{1}{x^2})), M(\log(2\pi))$.

Theorem 5.22.

$$\Psi(x) = x - \sum_\rho \frac{x^\rho}{\rho} - \frac{1}{2}\log\Bigl(1-\frac{1}{x^2}\Bigr) - \log(2\pi).$$

Proof. How to get to $sM\Psi$? First $n^{-s} = s\int_n^\infty x^{-s}\frac{dx}{x}$, so

$$\begin{aligned}\frac{\zeta'(s)}{\zeta(s)} &= -\sum_{n\geqslant 1}\frac{\Lambda(n)}{n^s} = -s\sum_{n\geqslant 1}\Lambda(n)\int_n^\infty x^{-s}\frac{dx}{x}\\ &= -s\int_1^\infty \sum_{n\leqslant x}\Lambda(n)x^{-s}\frac{dx}{x} = -s\int_1^\infty \Psi(x)x^{-s}\frac{dx}{x} = -sM\Psi,\end{aligned}$$

where the third from last doesn't include half x-term, the Penultimate does include, but this doesn't change the integral.

Observe for $k\in\mathbb{Z}$, $Mx^k = \int_1^\infty x^k x^{-s}\frac{dx}{x} = \frac{1}{s-k}$ provided $\mathrm{Re}(s) > k$. So for all $s \neq k$, the complex function $\frac{1}{s-k}$ is defined. Since $\log(1-\frac{1}{x^2}) = \sum_{n=1}^\infty \frac{1}{nx^{2n}}$, we have

$$\begin{aligned}-sM\left(x - \sum_\rho \frac{x^\rho}{\rho} - \frac{1}{2}\log\Bigl(1-\frac{1}{x^2}\Bigr) - \log(2\pi)\right) &= -\frac{s}{s-1} + s\sum_\rho \frac{1}{\rho}\frac{1}{s-\rho} + \frac{s}{2}\sum_{n=1}^\infty \frac{1}{n}\frac{1}{s+2n} + \log(2\pi)\\ &= \log(2\pi) - \frac{s}{s-1} + \sum_{n=1}^\infty \frac{s}{2n(s+2n)} + \sum_\rho \frac{s}{\rho(s-\rho)}\end{aligned}$$ □

But we know

$$\frac{\zeta'(s)}{\zeta(s)} = \log(2\pi) - \frac{s}{1-s} + \sum \frac{-s}{2n(s+2n)} + \sum \frac{s}{\rho(s-\rho)},$$

where sign difference is from the half of the final term.

Exercise 5.23. $M\log(1-\frac{1}{x^2})$ reasonably behaved as an integral, in the vivinity of $x=1$?

Exercise 5.24. Do we need to worry about ρ being complex in sMx^ρ?

Definition 5.25. *Von Mangoldt's explicit formula*:

$$\Psi(x) = x - \lim_{T\to\infty}\sum_{|x|<T}\frac{x^\rho}{\rho} - \frac{1}{2}\log\Bigl(1-\frac{1}{x^2}\Bigr) - \log(2\pi).$$

Remark. We've just shown both sides have the same Mellin transform, hence are the same except at points of discontinuity.

Key term in Von Mangoldt's explicit formula is the sum $\sum_\rho \frac{x^\rho}{\rho}$.

Nontrial roots of ζ come in conjugate pairs $\rho, \bar\rho$. If $\rho = \sigma + it$, then

$$\begin{aligned}\frac{x^\rho}{\rho} + \frac{x^{\bar\rho}}{\bar\rho} &= x^\sigma\left(\frac{x^{it}}{\sigma+it} + \frac{x^{-it}}{\sigma-it}\right) = \frac{x^\sigma}{\rho\bar\rho}\Bigl(e^{it\log x}(\sigma-it) + e^{-it\log x}(\sigma+it)\Bigr)\\ &= \frac{x^\sigma}{\rho\bar\rho}2\bigl(\sigma\cos(t\log x) + t\sin(t\log x)\bigr).\end{aligned}$$

If we now write $\sigma + it = |\rho|e^{i\alpha}$ so that $\cos\alpha = \frac{\sigma}{|\rho|}$ and $\sin\alpha = \frac{t}{|\rho|}$, then

$$\frac{x^\rho}{\rho} + \frac{x^{\bar\rho}}{\bar\rho} = \frac{2x^\sigma}{|\rho|}\big(\cos\alpha\cos(t\log x) + \sin\alpha\sin(t\log\alpha)\big) = \frac{2x^\sigma}{|\rho|}\cos(t\log x - \alpha).$$

If we suppose RH is true, i.e., for all ρ, $\sigma = \frac{1}{2}$, then the contribution to Von Mangoldt explicit formula is $x^{\frac{1}{2}}\sum_{\mathrm{Im}(\rho)>0}\frac{2}{|\rho|}\cos(t\log x - \alpha)$. If this sum is such that it behaves reasonable well (which is true), then the summand is a reasonable function of α, so the summand changes fairly smoothly. If we can show (which is true) for all $\epsilon > 0$, $f(x) := \sum\frac{2}{|\rho|}\cos(t\log x - \alpha) = O(x^\epsilon)$, then we'll get

$$\Psi(x) = x - x^{\frac{1}{2}}f(x) + O\Big(\frac{1}{x^2}\Big) + O(1) = x + O\big(x^{\frac{1}{2}+\epsilon}\big).$$

Hence $\pi(x) = \frac{x}{\log x} + O(x^{\frac{1}{2}+\epsilon})$.

Converse is true too: if we take a root ρ with $\sigma > \frac{1}{2}$, we'll get a contribution to the error on the order of x^σ. Whenever $t\log x - \alpha$ is an integer multiple of 2π, we'd get a contribution $\frac{2}{|\rho|}x^\sigma$. So the contribution would in periodic in $\log x$.

Theorem 5.26 (Abel summation via integrals)**.** *Suppose $f:\mathbb{R}\to\mathbb{C}$ has a continuous derivatives on the interval $[y,x]$ on $\mathbb{R}$ with $y < x$. Let $r \in \mathbb{N}_0$. Then*

$$\sum_{y<r\leqslant x} a(r)f(r) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)dt,$$

where as before $A(t) = \sum_{y<r\leqslant t} a(r)$.

In particular, if $y = 1$,

$$\sum_{1\leqslant r\leqslant x} a(r)f(r) = A(x)f(x) - \int_1^x A(t)f'(t)dt.$$

Remark. $A(t)$ is continuous except when t is an integer (and indeed is constant on $(n, n+1)$.)

Proposition 5.27. Let f have continuous derivatives on $[1,x]$, e.g., $f(x) = \log(x)$. (Check endpoints in the preceding.)

(a) $\sum_{1\leqslant r\leqslant x} a(r)f(r) = A(x)f(x) - \int_1^x A(t)f'(t)dt.$

(b) $\sum_{1\leqslant r\leqslant x} a(r)(f(x) - f(r)) = \int_1^x A(t)f'(t)dt.$

(c) If f has continuous derivative on $[2,x]$ and $a(1) = 0$, then $\sum_{2\leqslant r\leqslant x} a(r)f(r) = A(x)f(x) - \int_2^x A(t)f'(t)dt.$

Example 5.28. (a) $\sum_{1\leqslant r\leqslant x}\frac{a(r)}{r} = \frac{A(x)}{x} + \int_1^x\frac{A(t)}{t^2}dt.$

(b) $\sum_{y<r\leqslant x} a(r)\log r = A(x)\log x - A(y)\log y - \int_y^x\frac{A(t)}{t}dt.$

(c) $\sum_{1\leqslant r\leqslant x}(x-r)a(r) = \int_1^x A(t)dt$, $f(x) = x$ on $[1,x]$.

(d) $\sum_{1\leqslant r\leqslant x} ra(r) = xA(x) - \int_1^x A(t)dt.$

(e) If $a(1) = 0$, $\sum_{2\leqslant r\leqslant x} \frac{a(r)}{\log r} = \frac{A(x)}{\log x} + \int_2^x \frac{A(t)}{t\log^2 t}dt$.

We can get upper bounds on growth of some sums.

Example 5.29. Suppose $a(r)$ is such that for some $c > 0$, $A(t) \leqslant ct$ for all t. Then

$$\sum_{1\leqslant r\leqslant x} \frac{a(r)}{r} = \frac{A(x)}{x} + \int_1^x \frac{A(t)}{t^2}dt \leqslant c + c\log x = c(1+\log x).$$

Morel: if we replace $a(r)$ by its "average behavior", we get the same growth in $\sum \frac{a(r)}{r}$.

Example 5.30. Suppose there is a constant $c > 0$,

(a) $A(t) \leqslant c$,

(b) $f(t) \to 0$ as $t \to \infty$,

(c) $I = \int_1^\infty |f'(t)|dt$ converges.

e.g., if $f(t) \downarrow 0$. Then $\sum_{r=1}^\infty a(r)f(r)$ converges and is at most cI.

For $f(r) = \frac{1}{r}$, if $A(t) \leqslant c$ for all t, then $\sum_{r=1}^\infty \frac{a(r)}{r} = \int_1^\infty \frac{A(t)}{t^2}dt$, i.e., the sum converges.

Exercise 5.31. Suppose $|a(r)| \leqslant 1$ and $A(n) \leqslant c$ for all n. Show $\sum_{r=1}^\infty \frac{a(r)}{r} \leqslant 1 + \log c$. Hint: $A(t) \leqslant t$ for $1 \leqslant t \leqslant c$.

Exercise 5.32. (a) Prove that if $B(n) \to B$ as $n \to \infty$, then $\frac{1}{n}(B(1)+\cdots+B(n)) \to B$ as $n \to \infty$.

(b) Suppose $\sum_{r=1}^\infty \frac{a(r)}{r}$ converges. Write $b(r) = \frac{a(r)}{r}$, $B(n)$ the summatory function. Express $B(r)$ in terms of $A(r)$, and deduce $\frac{A(n)}{n^2} \to 0$ as $n \to \infty$.

Lemma 5.33. $\sum_{r=1}^\infty \frac{\sin(r\theta)}{r}$ converges.

Proof. If $e^{i\theta} \neq \pm 1$,

$$s(n) := \sum_{r=1}^n \sin(r\theta) = \frac{1}{2i}\left(\sum_{r=1}^n e^{ir\theta} - \sum_{r=1}^n e^{-ir\theta}\right) = \frac{1}{2i}\left(\frac{e^{i\theta}(1-e^{in\theta})}{1-e^{i\theta}} - \frac{e^{-i\theta}(1-e^{-in\theta})}{1-e^{-i\theta}}\right).$$

If $e^{i\theta} = \pm 1$, then $\sin(r\theta) = 0$. So $|s(n)| < \frac{1}{|1-e^{i\theta}|} + \frac{1}{|1-e^{-i\theta}|} < \frac{2}{|1-e^{i\theta}|}$. So for any fixed θ, $s(n)$ is bounded. Hence $\sum_{r=1}^\infty \frac{\sin(r\theta)}{r}$ converges (conditionally) by previous Exercise. □

Now we use Abel summation to relate the value of Dirichlet series $\sum_{n=1}^\infty \frac{a(n)}{n^s}$ to Mellin transforms via

$$\sum_{n\leqslant x} \frac{a(n)}{n^s} = \frac{A(x)}{x^s} + s\int_1^x \frac{A(t)}{t^s}\frac{dt}{t}.$$

So if $\frac{A(x)}{x^s} \to 0$, then the Dirichlet series $\sum_{n=1}^\infty \frac{a(n)}{n^s}$ converges if and only if $s\int_1^\infty \frac{A(t)}{t^s}\frac{dt}{t}$ converges.

The Euler-Maclourin summation formula works in a different way compared to abel summation. We wish to evaluate $\sum_{a<n\leqslant b} f(n)$, where $f(x)$ is a very well behaved function of x. We wish to relate $\sum_{a<n\leqslant b} f(n)$ to $\int_a^b f(t)dt$, where $a,b\in\mathbb{Z}$. Note

$$f(n)=\int_{n-1}^{n} f(t)dt-\int_{n-1}^{n}(f(t)-f(n))dt=\int_{n-1}^{n} f(t)dt-(f(t)-f(n))(t-c_n)|_{n-1}^{n}+\int_{n-1}^{n} f'(t)(t-c_n)dt.$$

where c_n is a constant for $n\in\mathbb{N}$. If we choose $c_n=n-\frac{1}{2}$, then on the interval $[n-1,n]$, $t-c_n=t-n+\frac{1}{2}=\{t\}-\frac{1}{2}$, where $\{t\}$ is the *fractional part* of t. So

$$\begin{aligned}\sum_{a<n\leqslant b} f(n) &= \int_a^b f(t)dt-\sum_{a<n\leqslant b}(f(t)-f(n))(\{t\}-1/2)|_{n-1}^{n}+\int_a^b f'(t)(\{t\}-1/2)dt\\ &=\int_a^b f(t)dt+\sum_{a<n\leqslant b}(f(n-1)-f(n))(\{n-1\}-1/2)+\int_a^b f'(t)(\{t\}-1/2)dt\\ &=\int_a^b f(t)dt+\frac{1}{2}\sum_{a<n\leqslant b}(f(n)-f(n-1))+\int_a^b f'(t)(\{t\}-1/2)dt\\ &=\int_a^b f(t)dt+\frac{1}{2}(f(b)-f(a))+\int_a^b f'(t)(\{t\}-1/2)dt.\end{aligned}$$

Let's write $b_1(t)=\{t\}-\frac{1}{2}$ on $[0,1]$, which is not a polynomial on $[0,1]$, but a perfectly nice function. Iteratively define $b_r(t)$ by $b_r'(t)=rb_{r-1}(t)$ (so $b_r(t)$ is defined up to a constant of integration) and $\int_0^1 b_r(t)dt$ (this gives us the constant of integration). We then extend these functions to $B_r(t)$ by $B_r(t)=b_r(\{t\})$ if $t\in\mathbb{R}\setminus\mathbb{Z}$.

If we choose $c_n=n$, $t-c_n=t-n=\{t\}$, then we have the following theorem:

Theorem 5.34.

$$\sum_{n=2}^{\infty} f(n)=\int_1^N f(x)dx+\int_1^N \{x\}f'(x)dx.$$

Theorem 5.35. *For any integers $a\leqslant b, k\geqslant 0$ and any function $f\in\mathcal{C}^{k+1}[a,b]$, we have*

$$\sum_{a<n\leqslant b} f(n)=\int_a^b f(t)dt+\sum_{r=0}^{k}\frac{(-1)^{r+1}B_{r+1}(0)}{(r+1)!}\left(f^{(r)}(b)-f^{(r)}(a)\right)+\frac{(-1)^k}{(k+1)!}\int_a^b B_{k+1}(t)f^{(k+1)}(t)dt,$$

where the $B_{r+1}(0)$ are the Bernoulli numbers and $B_{k+1}(t)$ are Bernoulli polynomials.

Estimating

$$\log(b!)=\sum_{1<n\leqslant b}\log n=\int_1^b \log t dt+\sum_{r=0}^{k}\frac{B_{r+1}(0)}{(r+1)!}(r-1)!\Big(\frac{1}{b^r}-\frac{1}{a^r}\Big)+\frac{1}{k+1}\int_a^b B_{k+1}(t)\frac{1}{t^{k+1}}dt.$$

This easily gives the (divergent) asymptotic series for $\log(b!)$.

Lemma 5.36. $3+4\cos\theta+\cos 2\theta\geqslant 0$ for all θ.

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

Proposition 5.37. Suppose we have a Dirichlet series $f(s) = \sum_{n\geqslant 1} \frac{a(n)}{n^s}$ with $a_n \geqslant 0$ for $n \geqslant 1$, and suppose the Dirichlet series converges for all $\mathrm{Re}(s) > \sigma_0$. Then for $\sigma > \sigma_0$, $3f(\sigma) + 4\,\mathrm{Re}\big(f(\sigma + it)\big) + \mathrm{Re}(f(\sigma + 2it)) \geqslant 0$.

Proof. It is enough to show $\mathrm{Re}\big(\frac{3a(n)}{n^\sigma} + \frac{4a(n)}{n^{\sigma+it}} + \frac{a(n)}{n^{\sigma+2it}}\big) \geqslant 0$. Note

$$\mathrm{Re}\left(\frac{a(n)}{n^\sigma}\left(3 + \frac{4}{n^{it}} + \frac{1}{n^{2it}}\right)\right) = \frac{a(n)}{n^\sigma}\big(3 + 4\cos(t\log n) + \cos(2t\log n)\big) \geqslant 0. \qquad \square$$

Corollary 5.38. Fix $t \in \mathbb{R}$. For $\sigma > 1$, define

$$H(\sigma) = \zeta(\sigma)^3 |\zeta(\sigma + it)|^4 |\zeta(\sigma + 2it)|.$$

Then $H(\sigma) \geqslant 1$.

Proof. Consider $\log(H(\sigma))$. For $\mathrm{Re}(s) \geqslant 1$, $f(s) := \log(\zeta(s)) = -\sum_p \log(1 - \frac{1}{p^s}) = \sum_{p^k} \frac{1}{kp^{ks}} = \sum_{n\geqslant 1} \frac{a(n)}{n^s}$, where $a(n) = \begin{cases} \frac{1}{k} & \text{if } n = p^k \\ 0 & \text{o.w.} \end{cases} \geqslant 0$. Also, $\log|z| = \mathrm{Re}(\log z)$, so the result follows from the previous proposition with $\sigma_0 = 1$. $\square$

Corollary 5.39. Since ζ is continuous,

$$\lim_{\sigma\to 1^+} \zeta(\sigma)^3 |\zeta(\sigma + it)|^4 |\zeta(\sigma + 2it)| \geqslant 1.$$

Theorem 5.40 (Weierstrass's Theorem for Series)**.** . *Assume $f_1, f_2, \cdots$ are holomorphic in an open set $\mathcal{D}$ and $\sum_{n=1}^\infty f_n(z)$ converges uniformly on every closed and bounded subset of $\mathcal{D}$. Then*

(a) $F(z) = \sum_{n=1}^\infty f_n(z)$ is holomorphic on $\mathcal{D}$,

(b) For all $k \geqslant 1$, $\sum_{n=1}^\infty f_i^{(k)}(z)$ converges on $\mathcal{D}$, and converge uniformly on every closed and bounded subset of $\mathcal{D}$ with limit $F^{(k)}(z)$. (So the series can be differentiated term by term.)

Example 5.41. Note $\frac{\{x\}}{x^{s+1}}$ is not continuous in x given s, s. Define $f_n(s) = \int_n^{n+1} \frac{\{x\}}{x^{s+1}}dx$ for $n \geqslant 1$. Then we can show for any $\delta > 0$,

(a) $\int_1^\infty \frac{\{x\}}{x^{1+s}}dx = \sum_{n=1}^\infty f_n(s)$ converges uniformly on $\mathrm{Re}(s) > \delta$,

(b) for $n \geqslant 1$, f_n is holomorphic on $\mathrm{Re}(s) > \delta$ with derivative

$$\frac{d}{ds} f_n(s) = \int_n^{n+1} \frac{\{x\}}{x^{s+1}}(-\log x)dx.$$

So

$$\frac{d}{ds}\int_1^\infty \frac{\{x\}}{x^{s+1}}dx = \frac{d}{ds}\sum_{n=1}^\infty f_n(s) = \sum_{n=1}^\infty \frac{d}{ds}f_n(s) = -\sum_{n=1}^\infty \int_n^{n+1} \frac{\{x\}}{x^{s+1}}\log x dx = -\int_{n=1}^\infty \frac{\{x\}}{x^{s+1}}\log x dx.$$

Theorem 5.42. *The analytic continuation of ζ to $\mathrm{Re}(s) > 0$ is given by*

$$\zeta(s) = \frac{1}{s-1} + \frac{1}{2} - s\int_1^\infty \frac{x - \lfloor x \rfloor - \frac{1}{2}}{x^{s+1}}dx = 1 + \frac{1}{s-1} - s\int_1^\infty \frac{\{x\}}{x^{s+1}}dx,$$

where x^{-s-1} takes the principal value $e^{-\log(x)(s+1)}$.

Proof. By previous theorem, $\sum_{n=2}^\infty f(n) = \int_1^N f(x)dx + \int_1^N \{x\}f'(x)dx$. So

$$\sum_{n=1}^N \frac{1}{n^s} = 1 + \int_1^N \frac{1}{x^s}dx - s\int_1^N \frac{\{x\}}{x^{s+1}}dx = 1 + \frac{1}{s-1} - \frac{N^{1-s}}{s-1} - s\int_1^N \frac{\{x\}}{x^{s+1}}dx.$$

If $\mathrm{Re}(s) > 1$, $\left|N^{1-s}\right| = N^{1-\sigma} \to \infty$ as $N \to \infty$, while by previous example, $\int_1^N \frac{\{x\}}{x^{s+1}}$ converges as $N \to \infty$ since $\left|\frac{\{x\}}{x^{s+1}}\right| \leqslant \frac{1}{x^{\sigma+1}}$. So $\zeta(s) = 1 + \frac{1}{s-1} - \lim_{N\to\infty} s\int_1^N \frac{\{x\}}{x^{s+1}}dx$. □

Corollary 5.43.

$$\zeta(s) = \sum_{n=1}^N \frac{1}{n^s} + \frac{N^{1-s}}{s-1} - s\int_N^\infty \frac{\{x\}}{x^{s+1}}dx.$$

Corollary 5.44. $\zeta(s) \neq 0$ for all s with $\mathrm{Re}(s) \geqslant 1$.

Proof. We've known $\zeta(s) \neq 0$ for $\mathrm{Re}(s) > 1$. Suppose $\zeta(s) = 0$ for some $s = 1 + it$ with $t \neq 0$. Since $\zeta(s)$ is **analytic** at $s = 1+it$, we have $\lim_{\sigma\to 1^+} \frac{\zeta(\sigma+it)}{\sigma-1} = \lim_{\sigma\to 1^+} \frac{\zeta(\sigma+it)-\zeta(1+it)}{\sigma-1} = \zeta'(1+it)$. Also, since $(\sigma-1)\zeta(\sigma) \to 1$ as $\sigma \to 1^+$, $(\sigma > 1)$

$$\begin{aligned}\lim_{\sigma\to 1^+} |\zeta'(1+it)|^4 |\zeta(\sigma+2it)| &= \lim_{\sigma\to 1^+}\left|\frac{\zeta(\sigma+it)}{\sigma-1}\right|^4 |\zeta(\sigma+2it)| \\ &= \lim_{\sigma\to 1^+} \frac{1}{\sigma-1}\frac{1}{\big((\sigma-1)\zeta(\sigma)\big)^3}\zeta(\sigma)^3 |\zeta(\sigma+it)|^4 |\zeta(\sigma+2it)| \geqslant \frac{1}{\sigma-1}.\end{aligned}$$

So as $\sigma \to 1^+$, $\zeta(\sigma+2it) \to \infty$, a contradiction since $\zeta(s)$ has no poles except $s = 1$ for $\sigma \geqslant 1$. Hence our assumption that $\zeta(1+it) = 0$ must be false, and hence $\zeta(s) \neq 0$ for all s with $\mathrm{Re}(s) \geqslant 1$. □

Question 5.45. Since we know that for $t > 0$, $|\zeta(1+it)| > 0$, can we obtain upper bounds on $|\zeta(s)|, s = \sigma + it, \sigma > 1$? Note $|\zeta(s) - 1| = \sum_{n\geqslant 2} \frac{1}{|n^s|} \leqslant \frac{1}{2^\sigma} + \int_2^\infty \frac{1}{x^\sigma}dx$.

Assume $t \geqslant 0$.

Let $r^*(s) = -s\int_1^\infty \frac{x - \lfloor x\rfloor - \frac{1}{2}}{x^{s+1}}dx$. Check $|r^*(s)| \leqslant \frac{|s|}{2\sigma}$ using $\left|\frac{1}{x^s}\right| = \frac{1}{x^\sigma}$. Hence for $|s-1| > 1$, $|\zeta(s)| < \frac{3}{2} + \frac{|s|}{2\sigma} < \frac{3}{2} + \frac{\sigma}{2\sigma} + \frac{t}{2\sigma} = 2 + \frac{t}{2\sigma}$. This is linear in t: not a very good bound. If we assume t is a little bigger, say $t \geqslant 2$, then we can obtain better bounds.

Once we know that ζ has no zeros with $s = 1 + it$, we'll also know (since it has no poles except at $s = 1$ and no zeros when $\mathrm{Re}(s) \geqslant 1$) that $\frac{\zeta'(s)}{\zeta(s)} = \frac{d}{ds}\log(\zeta(s))$ has a pole at 1 and no other poles when $\mathrm{Re}(s) \geqslant 1$. Hence $\zeta(s) - \frac{\zeta'(s)}{\zeta(s)}$ is well-behaved for $\mathrm{Re}(s) \geqslant 1$. When we consider the corresponding Dirichlet series, this will lead us to $\psi(x) \sim x$ and hence to $\pi(x) \sim \frac{x}{\log x}$ and hence PNT.

Assume now $\sigma \geqslant 1, t \geqslant 2, s = \sigma + it$.

Theorem 5.46. *For $\sigma \geqslant 1$ and $t \geqslant 2$, $|\zeta(\sigma+it)| \leqslant 4+\log t$.*

Proof. Prove for $\sigma > 1$, and deduce for $\sigma = 1$ by continuity (since $t \geqslant 2$, we avoid $s = 1$). $\zeta(s) = \sum_{n=1}^{N} \frac{1}{n^s} + \frac{N^{1-s}}{s-1} + r_n(s)$, where $r_n(s) = -s\int_N^\infty \frac{\{x\}}{x^{s+1}}dx$. So $|r_n(s)| \leqslant \frac{|s|}{\sigma N^\sigma}$.

So for fixed $\sigma > 1, t \geqslant 2$, put $N = \lfloor t \rfloor$, $\left|\sum_{n=1}^{N} \frac{1}{n^s}\right| \leqslant \sum_{n=1}^{N} \frac{1}{n} < 1+\log N \leqslant 1+\log t$, $\left|\frac{N^{1-s}}{s-1}\right| \leqslant \frac{1}{|s-1|} < \frac{1}{t} \leqslant \frac{1}{2}$, $|r_n(s)| \leqslant \frac{|s|}{\sigma N^\sigma} \leqslant (1+\frac{t}{\sigma})\frac{1}{N^\sigma} \leqslant \frac{1+t}{N} \leqslant \frac{N+2}{N} \leqslant 2$. So $|\zeta(s)| < 1+\log t+\frac{1}{2}+2 < 4+\log t$. □

Exercise 5.47. When $s = 1+it$, $t > 0$ fixed, $\zeta(s) = \sum_{n=1}^{N} \frac{1}{n^s} + \frac{e^{-it\log N}}{it} + O(\frac{1}{N})$. So the partial sum asymptotically approach to a circle of radius $\frac{1}{t}$ around $\zeta(1+it)$.

Also want upper bounds on $\zeta'(s)$. By previous example,

$$\zeta'(s) = \sum_{n=1}^{N} \frac{\log n}{n^s} - \frac{N^{1-s}}{s-1}\log N - \frac{N^{1-s}}{(s-1)^2} - \int_N^\infty \frac{\{x\}}{x^{s+1}}dx + s\int_N^\infty \frac{\{x\}\log x}{x^{s+1}}dx.$$

Let $I_1(s) = \int_N^\infty \frac{\{x\}}{x^{s+1}}dx$ and $I_2(s) = \int_N^\infty \frac{\{x\}\log x}{x^{s+1}}dx$. Then for fixed $\sigma > 1, t \geqslant 2$, setting $N = \lfloor t \rfloor$. Then $\left|\sum_{n=1}^{N} \frac{\log n}{n^s}\right| \leqslant \left|\sum_{n=1}^{N} \frac{\log n}{n}\right| \leqslant \frac{1}{2}(\log N)^2 + \frac{1}{8} \leqslant \frac{1}{2}\log^2 t + \frac{1}{8}$, $\left|\frac{N^{1-s}}{s-1}\log N\right| \leqslant \frac{\log N}{t} \leqslant \frac{\log t}{t} \leqslant \frac{1}{e}+\frac{1}{2}$, $\left|\frac{N^{1-s}}{(s-1)^2}\right| \leqslant \frac{1}{t^2} < \frac{1}{4}$, $|I_1(s)| \leqslant \int_1^\infty \frac{1}{x^2}dx = \frac{1}{N} \leqslant \frac{1}{2}$, $|sI_2(s)| \leqslant |s|\int_N^\infty \frac{\log x}{x^{\sigma+1}}dx \leqslant |s|\frac{\log N}{\sigma N^\sigma} + \frac{1}{\sigma^2 N^\sigma} \leqslant (1+\frac{t}{\sigma})\frac{\log t+1}{N} < (1+\frac{t}{\sigma})(\log N+1) < 2(\log t+3)$? Putting it together,

$$|\zeta'(s)| \leqslant \left|\sum_{n=1}^{N} \frac{\log n}{n^s}\right| + \left|\frac{N^{1-s}}{s-1}\log N\right| + \left|\frac{N^{1-s}}{(s-1)^2}\right| + |I_1(s)| + |sI_2(s)| \leqslant \frac{1}{2}(\log t+3)^2.$$

So we have upper bounds for $\zeta(s), \zeta'(s)$. But we want upper bounds on $\frac{\zeta'(s)}{\zeta(s)}$ too. So we'll need to improve the proof that $\zeta(s) \neq 0$ for $s = 1+it$ to obtain lower bounds on $|\zeta(1+it)|$ or equivalently, upper bounds on $\left|\frac{1}{\zeta(1+it)}\right|$ so that we get $|\zeta(s)| < P_1(\log t)$, $\left|\frac{\zeta'(s)}{\zeta(s)}\right| \leqslant P_2(\log t)$ for polynomials $P_1()$ and $P_2()$, then we can consider $\left|\zeta(s) - \frac{\zeta'(s)}{\zeta(s)}\right| < P_3(\log t)$, where $\zeta(s) - \frac{\zeta'(s)}{\zeta(s)}$ has no poles for $\sigma \geqslant 1$. We'll then interpret this in terms of Dirichlet series $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ and $\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}$.

We'll show that since $\left|\zeta(s) - \frac{\zeta'(s)}{\zeta(s)}\right| < P_3(\log t)$, the function $A(x) = \sum_{n \leqslant x}(1-\Lambda(n))$ will have $A(x) = o(x)$, which will imply $\Psi(x) = \sum_{n\leqslant x}\Lambda(n)$ will satisfy $\Psi(x) \sim x$. This will imply the Prime Number Theorem, $\pi(x) \sim \int_2^x \frac{1}{\log t}dt$.

Theorem 5.48. *Suppose for all $\sigma \geqslant 1$ and $t \geqslant t_0$, we have bounds $M_1(t), M_2(t) \geqslant 1$ so that $|\zeta(\sigma+2it)| < M_1(t)$ and $|\zeta'(\sigma+it)| < M_2(t)$. Then $\left|\frac{1}{\zeta(\sigma+it)}\right| \leqslant 2^5 M_1(t)M_2(t)^3$. In particular, with $t_0 = 2$, $M_1(t) = 4+\log t < 5+\log t$ and $M_2(t) = \frac{1}{2}(3+\log t)^2$, we get $\left|\frac{1}{\zeta(\sigma+it)}\right| \leqslant 4(5+\log t)^7$. So $\left|\frac{\zeta'(s)}{\zeta(s)}\right| \leqslant c(5+\log t)^9$.*

Proof. We'll prove for $\sigma > 1$, and deduce for $\sigma = 1$ by continuity. If $\sigma > \frac{5}{4}$, the result is easy: $\left|\frac{1}{\zeta(s)}\right| < \zeta(\sigma) < \frac{\sigma}{\sigma-1}$, where the second inequality is from

$$\sum_{n=1}^{\infty} \frac{1}{n^\sigma} < 1 + \int_1^\infty \frac{1}{x^\sigma} dx = 1 + \frac{1}{\sigma - 1} = \frac{\sigma}{\sigma - 1} < \frac{5/4}{1/4} = 5 < 2^5 \leqslant 2^5 M_1(t) M_2(t).$$

So we can assume $1 < \sigma < \frac{5}{4}$. Hence $\frac{\sigma}{\sigma-1} < \frac{5}{4}\frac{1}{\sigma-1} < 2^{1/3}\frac{1}{\sigma-1}$. So for $\sigma \in (1, 5/4)$, the inequality $\zeta(\sigma)^3 |\zeta(\sigma+it)|^4 |\zeta(\sigma+2it)| \geqslant 1$ implies $(\frac{2^1/3}{\sigma-1})|\zeta(\sigma+it)|^4 M_1(t) \geqslant 1$. Hence $|\sigma + it| \geqslant \frac{(\sigma-1)^{3/4}}{2^{1/4} M_1(t)^{1/4}} =: f(\sigma)$. Now find η so that $f(\eta) = 2(\eta-1)M_2(t)$ so that $\eta - 1 = \frac{1}{2^5 M_1(t) M_2(t)} \leqslant \frac{1}{2^5}$ (so $\eta < \frac{5}{4}$) (Note $\eta - 1 = \frac{f(\eta)}{M_2(t)} = \frac{(\eta-1)^{3/4}}{2^{1/4} M_1(t)^{1/4} M_1(t)^{1/4}}$, $2^4(\eta-1)^4 M_2(t)^4 = \frac{f(\eta)^4}{M_2(t)^4} = \frac{(\eta-1)^3}{2M_1(t)}$). Then we have two cases to consider: if $\eta \leqslant \sigma$, then $f(\sigma) \geqslant f(\eta) = \frac{1}{2^5 M_1(t) M_2(t)^4}$; if $\eta > \sigma$, then $\zeta(\eta+it) - \zeta(\sigma+it) = \int_0^\eta \zeta'(x+it)dx$, so $|\zeta(\eta+it) - \zeta(\sigma+it)| \leqslant (\eta-\sigma)M_2(t) \leqslant (\eta-1)M_2(t)$. So by the triangle inequality,

$$|\zeta(\sigma+it)| \geqslant |\zeta(\eta+it)| - (\eta-1)M_2(t) = 2(\eta-1)M_2(t) - (\eta-1)M_2(t) \geqslant \frac{1}{2^5 M_1(t) M_2(t)^5}.$$

Completing the proof of the upper bound for $\frac{1}{|\zeta(s)|}$ and hence the upper bound for $\left|\frac{\zeta'(s)}{\zeta(s)}\right|$. □

How does this average behaviors of a Dirichlet series affect the analytic behaviors and vice versa? Suppose $A(x) = \sum_{n \leqslant x} a_n$ ($\frac{1}{2}$ when $x \in \mathbb{Z}$). How can we obtain information about $A(x)$ from $f(s) = \sum_{n \geqslant 1} \frac{a_n}{n^s}$? Some generating functions and complex analysis. If we could find a generating function $G(z) = \sum_{n=1}^\infty a_n z^n$ and if $G(z)$ converges inside $|z| < R$, then for $0 < r < R$ we'd have $a_n = \frac{1}{2\pi i}\oint_{|z|=r} \frac{G(z)}{z^{n+1}} dz$. So $A_n = \sum_{k \leqslant n} a_k$, $A_n = [z^n]\frac{G(z)}{1-z}$. So if $r < \min(1, R)$, $A_n = \frac{1}{2\pi i}\oint_{|z|=r} \frac{G(z)}{(1-z)z^{n+1}} dz$.

Exercise 5.49. Go back through notes and find where we used integrals to transform z^n into $\frac{1}{n^s}$.

Review: ζ has a simple pole at $s = 1$ and no zeros with $\operatorname{Re}(s) \geqslant 1$; $|\zeta(1+it)| < P_1(\log|t|)$ for $|t| \geqslant 2$; $\frac{\zeta'}{\zeta}$ has a simple pole at $s = 1$ and no other poles for $\operatorname{Re}(s) \geqslant 1$; since $\zeta'(s) \leqslant P_2(\log|t|), |t| \geqslant 2, s = 1 + it$ and $\left|\frac{1}{\zeta(s)}\right| \leqslant P_3(\log|t|), s = 1+it, |t| \geqslant 2$, we have $\left|\frac{\zeta'(s)}{\zeta(s)}\right| < P_4(\log|t|)$ for $|t| \geqslant 2$.

So by making the polynomials have even degree (which we can do) when adding positive constants, we can extend $\left|\frac{\zeta'(s)}{\zeta(s)}\right| \leqslant P(\log|t|)$ for all $s = 1+it, t \neq 0$. Hence $\left|\zeta(s) - \frac{\zeta'(s)}{\zeta(s)}\right| \leqslant P_5(\log|t|)$ for all $s = 1 + it$.

Dirichlet series

$$\zeta(s) - \frac{\zeta'(s)}{\zeta(s)} = \sum_{n \geqslant 1} \frac{1 - \Lambda(n)}{n^s}.$$

We want to conclude that if $f(s)$ has no poles for $\operatorname{Re}(s) \geqslant 1$ and $f(s) = \sum_{n \geqslant 1} \frac{a_n}{n^s}$ and (some condition about $|f(s)|$ not growing too fast on $s = 1+it$), then $A(x) = \sum_{n \leqslant x} a_n$ has nice behaviors, say $A(x) = o(x)$. This would prove PNT! Then we'd have $\sum_{n \leqslant x}(1 - \Lambda(n)) = o(x)$. So $\sum_{n \leqslant x} \Lambda(n) = \sum_{n \leqslant x} 1 - o(x) = \lfloor x \rfloor - o(x) = x + o(x)$. Hence $\Psi(x) \sim x$. Thus $\Pi(x) \sim \operatorname{Li}(x)$ by Riemann summation, so $\pi(x) \sim \operatorname{Li}(x)$ by Mobius inversion.

Note 5.50. We don't know that a_n is small: it can be relative large: $|a_n| = \log n - 1$ when n is prime. This means we're not going to be able to use any sorts of convergence results, especially since the series we're subtracting $\zeta(s)$ and $\frac{\zeta'(s)}{\zeta(s)}$ both fail to converge at $s = 1$. We're going to need cancellations in the summands of $A(x)$ to help us out. So suppose we have $f(s) = \sum_{n\geqslant 1} \frac{a_n}{n^s}$, $A(x) = \sum_{n\leqslant x} a_n$. What information about $A(x)$ can we obtain from information about $f(s)$?

Some tools:

(a) Cauchy's residue theorem: Suppose $f(s)$ has a simple pole at s_0 and $f(s) - \frac{r_{-1}}{s-s_0} = r_0 + (s - s_0)r_1(s)$ and $r_1(s)$ is analytic inside a domain $\mathcal{D}$. Then provided Γ lies inside $\mathcal{D}$ (so in particular it encloses no other poles of $f(s)$!) and encloses s_0, then

$$\frac{1}{2\pi i}\oint_\Gamma f(s)ds = r_{-1}.$$

(b) If Γ is a curve of length L and if $|f(s)| \leqslant M$ for s on Γ, then $\left|\oint_\Gamma f(s)ds\right| \leqslant ML$.

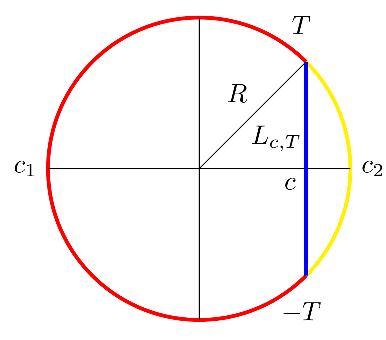
(c) Principal values: suppose $f(s)$ is such that for each $T > 0$, $\int_{c-iT}^{c+iT} f(s)ds = \int_{c-iT}^{c+iT} f(c + iT)idt$ exists, integral along the line $c + it, |t| \leqslant T$, and if $\lim_{T\to\infty} \int_{c-iT}^{c+iT} f(s)ds$ exists, we'll write

$$\int_{L_c} f(s)ds = \lim_{T\to\infty}\int_{c-iT}^{c+iT} f(s)ds$$

for the principal value of this integral.

Note it doesn't imply the integral converges: for that we'd need distinct upper and lower bounds for T, T'. Hence we'll take advantage, perhaps, of cancellation in the integral.

(d) (Shifted) Heaviside function $E(x) = \begin{cases} 1 & \text{if } x \geqslant 1 \\ 0 & \text{if } x < 1 \end{cases}$. Fact. If $x > 0$ and $c > 0$, then $\frac{1}{2\pi i}\int_{L_c} \frac{x^s}{s^2}ds = E(x)\log x$. Why $\log x$? $x^s = e^{s\log x} = \sum_{k=0}^{\infty} \frac{(s\log x)^k}{k!}$. Let c_1, c_2 be the portions of the circle of radius $R = \sqrt{c^2 + T^2}$ to either side of $L_{c,T}$.



So

$$\frac{1}{2\pi i}\oint_{c_1\cup c_2}\frac{x^s}{s^2}ds - \frac{1}{2\pi i}\oint_{c_1\cup c_2}\sum_{k=0}^{\infty} s^{k-2}\frac{(\log x)^k}{k!}ds = [s^1]e^{s\log x} = \log x.$$

So if $\int_{c_1} \frac{x^s}{s^2} ds \to 0$ as $R \to \infty$, then $\frac{1}{2\pi i}\int_{L_c} \frac{x^s}{s^2} ds = \log x$. Want to show this holds if $x \geqslant 1$. Conversely, if $\lim_{R\to\infty} \oint_{c_2} \frac{x^s}{s^2} ds = 0$, then since $L_{c,T} \cup c_2$ doesn't enclose 0, $\int_{L_{c,T}\cup c_2} \frac{x^s}{s^2} ds = 0$, so $\frac{1}{2\pi i}\oint_{L^c} \frac{x^s}{s^2} ds = 0$. Want to show this holds if $x < 1$.

Theorem 5.51. *$\frac{1}{2\pi i}\int_{L_c} \frac{x^s}{s^2} ds = E(x)\log x$.*

Proof. We want to show

(a) if $x \geqslant 1$, $\int_{c_1} \frac{x^s}{s^2} ds = 0$ as $R \to \infty$;

(b) if $x < 1$, $\int_{c_2} \frac{x^s}{s^2} ds \to 0$ as $R \to \infty$.

Assume $x \geqslant 1$. Then on c_1, $\mathrm{Re}(s) < c$, so $|x|^s < |x^c|$, then $\left|\frac{x^s}{s^2}\right| < \frac{x^c}{R^2}$ on c_1, so $\left|\frac{1}{2\pi i}\int_{c_1} \frac{x^s}{s^2} ds\right| \leqslant \frac{1}{2\pi} x^c \frac{2\pi R}{R^2} = \frac{x^c}{R} \to 0$ as $R \to \infty$, hence $\frac{1}{2\pi i}\int_{L_c} \frac{x^s}{s^2} ds = \log x$.

If $x \in (0,1)$, consider $c_2 \cup L_{c,T}$. On c_2, $\mathrm{Re}(c) > c$, so $x^{\mathrm{Re}(s)} < x^c$, so $\left|\frac{1}{2\pi i}\int_{c_2} \frac{x^s}{s^2} ds\right| \leqslant \frac{1}{2\pi} x^c \frac{2\pi R}{R^2} = \frac{x^c}{R} \to 0$ as $R \to \infty$. Since $\int_{c_2 \cup L_{c,T}} \frac{x^s}{s^2} ds = 0$ and $\int_{c_2} \frac{x^s}{s^2} ds \to 0$ as $R \to \infty$, we have $\int_{L_c} \frac{x^s}{s^2} ds \to 0$. □

Likewise, using some contours c_1, c_2, we can prove

Proposition 5.52. If $x > 0, c > 1$, $\frac{1}{2\pi i}\int_{L_c} \frac{x^s}{s(s-1)} ds = (x-1)E(x)$, $\frac{1}{2\pi i}\int_{c_1\cup c_2} (\frac{x^s}{s-1} - \frac{x^s}{s}) ds = x^1 - x^0 = x - 1$.

Now if we let $f(s) = \sum_{n\geqslant 1} \frac{a_n}{n^s}$ and consider $x^s f(s) = \sum_{n\geqslant 1} a_n (\frac{x}{n})^s$, then $\frac{1}{2\pi i}\int_{L_c} \frac{x^s f(s)}{s(s-1)} ds$ "ought to be"

$$\sum_{n\geqslant 1} \frac{a_n}{2\pi i}\int_{L_c} \frac{(\frac{x}{n})^s}{s(s-1)} ds = \sum_{n\geqslant 1} a_n \Big(\frac{x}{n} - 1\Big) E\Big(\frac{x}{n}\Big) = \sum_{n\leqslant x} a_n \Big(\frac{x}{n} - 1\Big) = x \sum_{n\leqslant x} \frac{a_n}{n} - \sum_{n\leqslant x} a_n ?$$

Let's try to get there. Suppose $f(s) = \sum_{n\geqslant 1} \frac{a_n}{n^s}$ is convergent for $\mathrm{Re}(s) > 1$. $A(x) = \sum_{n\leqslant x} a_n$. Then for any $c > 1$ and $x > 1$,

$$\frac{1}{2\pi i}\int_{L_c} \frac{x^s}{s(s-1)} f(s) ds = \sum_{n\leqslant x} a_n \Big(\frac{1}{n} - \frac{1}{x}\Big) = \int_1^x \frac{A(y)}{y^2} dy.$$

Proof in a moment.

Theorem 5.53. *Suppose $f(s) = \sum_{n\geqslant 1} \frac{a_n}{n^s}$ is absolutely convergent for $\mathrm{Re}(s) > 1$. $A(x) = \sum_{n\leqslant x} a_n$. Then for $c > 1, c > 1$,*

$$\frac{1}{2\pi i}\int_{L_c} \frac{x^s}{s(s-1)} f(s) ds = \sum_{n\leqslant x} a_n (\frac{1}{n} - \frac{1}{x}) = \int_1^x \frac{A(y)}{y^2} dy.$$

Proof. Write $x^s f(x) = G(s) + H(s)$, where $G(s) = \sum_{n\leqslant x} a_n (\frac{x}{n})^s$ and $H(s) = \sum_{n>x} a_n (\frac{x}{n})^s$. Since $c > 1$, $H(c)$ converges, say to M. For $n > x$ and $\mathrm{Re}(s) \leqslant c$, $\left|(\frac{x}{n})^s\right| \leqslant \left|\frac{x}{n}\right|^c$, so $absH(s) \leqslant M$. Since $G(s)$ is a finite sum, $\frac{1}{2\pi i}\int_{L_c} \frac{G(s)}{s(s-1)} ds = \sum_{n\leqslant x} a_n (\frac{x}{n} - 1)$ $(= \sum_{n\geqslant 1} a_n (\frac{x}{n} - 1) E(\frac{x}{n}))$. $\left|\frac{1}{2\pi i}\int_{L_c} \frac{H(s)}{s(s-1)} ds\right|$

is bounded, so the integral exists, and since $H(s)$ is analytic in $\mathrm{Re}(s) > 1$, the integral is 0 (by considering $\frac{1}{2\pi i}\int_{c_2}\frac{H(s)}{s(s-1)}ds = 0$, so $\int_{L_c}\frac{H(s)}{s(s-1)}ds = 0$). Hence

$$\frac{1}{x}\left(\frac{1}{2\pi i}\int_{L_c}\frac{G(s)}{s(s-1)}ds + \frac{1}{2\pi i}\int_{L_c}\frac{H(s)}{s(s-1)}ds\right) = \sum_{n\leqslant x} a_n(\frac{1}{n}-\frac{1}{x}) = \int_1^x \frac{A(y)}{y^2}dy. \qquad \square$$

Theorem 5.54 (Riemann-Lebesgue)**.** *Suppose $\varphi : \mathbb{R} \to \mathbb{C}$ has a continuous derivative, and that $\int_{-\infty}^{\infty}|\varphi(t)|dt$ converges. Let $F(\lambda) = \int_{-\infty}^{\infty} e^{i\lambda t}\varphi(t)dt$ (which exists for all λ since $\left|e^{i\lambda t}\right| = 1$). $\frac{1}{2\pi i}F(-\lambda)$ is the Fourier transform $\varphi(t)$. Then $F(\lambda) \to 0$ as $\lambda \to \infty$. Furthermore, if $\int_{-\infty}^{\infty}\varphi'(t)dt$ converges and $\left|\int_{-\infty}^{\infty}\varphi'(t)dt\right| = I_1$ also converges, then $F(\lambda) = -\frac{1}{i\lambda}\int_{-\infty}^{\infty}e^{i\lambda t}\varphi'(t)dt$ and $|F(\lambda)| \leqslant \frac{I_1}{\lambda}$ giving a rate of convergence of $F(\lambda)$ to 0 as $\lambda \to \infty$.*

Proof. Let $\epsilon > 0$. Pick (fix) T such that $\int_{-\infty}^{-T}|\varphi(t)|dt < \frac{\epsilon}{3}$ and $\int_T^{\infty}|\varphi(t)|dt < \frac{\epsilon}{3}$. Let $F_T(\lambda) = \int_{-T}^{T} e^{i\lambda t}dt = \frac{1}{i\lambda}e^{i\lambda t}\varphi(t)|_{-T}^{T} - \frac{1}{i\lambda}\int_{-T}^{T}e^{i\lambda T}\varphi'(t)dt$ by integrating by parts. Now φ' is continuous, and hence bounded on $[-T,T]$, say $|\varphi(t)| \leqslant M$ for $t \in [-T,T]$, so $|F_T(\lambda)| \leqslant \frac{1}{\lambda}|\varphi(T)| + \frac{1}{\lambda}|\varphi(-T)| + \frac{2}{\lambda}MT = \frac{1}{\lambda}(|\varphi(T)| + |\varphi(-T)| + 2MT)$. Choose λ sufficiently large so that $|F_T(\lambda)| \leqslant \frac{\epsilon}{3}$. Then $|F(\lambda)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$. $\square$

Theorem 5.55 (Fundamental)**.** *$f(s) = \sum_{n\geqslant 1}\frac{a_n}{n^s}$ converges absolutely for $\mathrm{Re}(s) > 1$ to a function analytic on a region including $\mathrm{Re}(s) \geqslant 1$ except at $s = 1$, where it has at most a simple pole, so*

$$f(s) = \frac{\alpha}{s-1} + \alpha_0 + (s-1)h(s),$$

where $h(s)$ is analytic in a region containing $\mathrm{Re}(s) \geqslant 1$, and hence there is a function $P(t)$ and $t_0 \geqslant 1$, so that $|f(\sigma + it)| \leqslant P(t)$ for all $\sigma \geqslant 1, t \geqslant t_0$ and $\int_1^{\infty}\frac{P(t)}{t^2}dt$ converges, then $\int_1^{\infty}\frac{A(x)-\alpha x}{x^2}dx$ converges to $\alpha' = \alpha_0 - \alpha$, $A(x) = \sum_{n\leqslant x} a_n$.

Proof. Let $f(s) = \frac{\alpha}{s-1} + \alpha_0 + (s-1)h(s)$ and $\varphi(s) = \frac{h(s)}{s}$. Then φ, h are both analytic on $\mathrm{Re}(s) \geqslant 1$ and $\frac{1}{s-1} = \frac{s}{s-1} - 1$. So $(s-1)h(s) = f(s) - \frac{\alpha}{s-1} - \alpha_0 = f(s) - \frac{\alpha s}{s-1} - \alpha'$. Hence $\varphi(s) = \frac{f(s)}{s(s-1)} - \frac{\alpha}{(s-1)^2} - \frac{\alpha'}{s(s-1)}$, i.e., $s(s-1)\varphi(s) = f(s) - \frac{s\alpha}{s-1} - \alpha'$. So $|s(s-1)\varphi(s)| \leqslant P(t) + |\alpha| + |\alpha'| \leqslant P_1(t)$. $\left|\frac{s}{s-1}\right| = \left|\frac{\sigma+it}{(\sigma-1)+it}\right| = \frac{|(\sigma+it)((\sigma-1)-it)|}{(\sigma-1)^2+t^2}$. Assume this is fixed. Then $|s(s-1)| = |s||s-1| > t^2$. So $|\varphi(s)| \leqslant \frac{P_1(t)}{t^2}$ and we'll need to check $\int_1^{\infty}\frac{P_1(t)}{t^2}dt$ converges, hence so does $\int_1^{\infty}|\varphi(\sigma+it)|dt$ and hence so does $\int_{-\infty}^{\infty}|\varphi(\sigma+it)|dt$. Now for $x > 1$ and $c \geqslant 1$, define $I(x,c) = \frac{1}{2\pi i}\int_{L_c}x^{s-1}\varphi(s)ds$. Then for $c > 1$,

$$\begin{aligned} I(x,c) &= \frac{1}{2\pi i}\int_{L_c}\frac{x^{s-1}}{s(s-1)}f(s)ds - \frac{\alpha}{2\pi i}\int_{L_c}\frac{x^{s-1}}{(s-1)^2}ds - \frac{\alpha'}{2\pi i}\int_{L_c}\frac{x^{s-1}}{s(s-1)}ds \\ &= \int_1^x \frac{A(y)}{y^2}dy - \alpha\log x - \alpha'(1-\frac{1}{x}). \end{aligned}$$

Note this is independent of c. Now we need to show $I(x,1) = \lim_{c\to 1^+ I(x,c)} = I(x,1)$. Check this: use the fact that tails are small, and the functions are uniformly continuous on $[-T,T]$. Hence for $x > 1$,

$$I(x,1) = \int_1^x \frac{A(y)}{y^2}dy - \alpha\log x - \alpha'(1-\frac{1}{x}) = \int_1^x \frac{A(y)-\alpha y}{y^2}dy - \alpha' + \frac{\alpha'}{x}.$$

So if we show $\lim_{x\to\infty} I(x,1) = 0$, then we'll have $\lim_{x\to\infty} \int_1^\infty \frac{A(y)-\alpha y}{y^2} dy = \alpha'$. But

$$I(x,1) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} x^{it} \varphi(1+it)dt = \frac{1}{2\pi i} \int_{-\infty}^{\infty} e^{i\lambda t} \varphi(1+it)dt,$$

where $\lambda = \log x$ and since $\int_{-\infty}^{\infty} |\varphi(1+it)| < \infty$, by the Riemann-Lebesgue Theorem, $I(x,1) \to 0$ as $x \to \infty$. So we have Bunch o'stuff implies $\int_1^\infty \frac{A(x)-\alpha x}{x^2} dx$ converges. Now we need to show that the convergence of this integral, for nice sequences a_n, implies the converges of $\frac{A(x)}{x}$ to α as $x \to \infty$. □

Proposition 5.56. Suppose $A(x)$ is a real-valued function on $[1,\infty)$ and α is such that $\int_1^\infty \frac{A(x)-\alpha x}{x^2} dx$ converges. Suppose further that either

(i) $A(x)$ is non-negative and weakly increasing, or

(ii) $A(x) = B(x) - (B(x) - A(x))$ and $B(x)$ and $B(x) - A(x)$ are non-negative and weakly increasing and for some $\beta \in \mathbb{R}$, $\int_1^\infty \frac{B(x)-\beta x}{x^2} dx$ converges.

Then $\frac{A(x)}{x} \to \alpha$ as $x \to \infty$.

Proof. First, we show (i) implies (ii): If $\int_1^\infty \frac{A(x)-\alpha x}{x^2} dx$ and $\int_1^\infty \frac{B(x)-\beta(x)}{x^2} dx$ converge, then so does $\int_1^\infty \frac{B(x)-A(x)-(\beta-\alpha)x}{x^2} dx$, so if $B, B-A$ satisfy (i), $\frac{B(x)}{x} \to \beta$, $\frac{B(x)-A(x)}{x} \to \beta - \alpha$, then $\frac{A(x)}{x} \to \alpha$.

Observe also since $A(x)$ is non-negative that if $\frac{A(x)}{x} \to \alpha$, there are only two possibilities: $\alpha = 0$ or $\alpha > 0$.

Case 1. $\alpha > 0$: by replacing $A(x)$ by $\frac{A(x)}{\alpha}$, i.e., a_n is replaced by $\frac{a_n}{\alpha}$, we may assume $\alpha = 1$. So take $A(x)$ weakly increasing, $A(x) \geqslant 0$, and suppose $\int_1^\infty \frac{A(x)-x}{x^2} dx$ converges. Want to show is that if $0 < \epsilon < \frac{1}{2}$, we can find $x_0(\epsilon)$ so that if $x > x_0$, $\left|\frac{A(x)}{x} - 1\right| < \epsilon$. Part (a): find x_1 so that if $x > x_1$, $\frac{A(x)}{x} < 1+\epsilon$. Part (b): find x_2 so that $x > x_2$, $\frac{A(x)}{x} > 1-\epsilon$. Since $\int_1^\infty \frac{A(x)-x}{x^2} dx$ converges, for every $\delta > 0$, we can find $M(\delta)$ so that if $M_1 > M_0 > M$, $\left|\int_{M_1}^{M_2} \frac{A(x)-x}{x^2} dx\right| < \delta$. Suppose $\frac{A(M_0)}{M_0} > 1+\epsilon$. Since $A(x)$ is weakly increasing, $A(M_1) \geqslant A(M_0)$, so

$$\int_{M_0}^{M_1} \frac{A(x)-x}{x^2} dx \geqslant \int_{M_0}^{M_1} \frac{A(M_0)-x}{x^2} dx = \int_{M_0}^{M_1} \frac{A(M_0)}{x^2} dx - \int_{M_0}^{M_1} \frac{x}{x^2} dx = A(M_0)(\frac{1}{M_0} - \frac{1}{M_1}) - \log \frac{M_1}{M_0}.$$

Now fix $M_1 = M_0(1+\epsilon)$, so

$$A(M_0)(\frac{1}{M_0} - \frac{1}{M_1}) - A(M_0)(\frac{1}{M_0} - \frac{1}{M_0(1+\epsilon)}) = \frac{A(M_0)}{M_0} \frac{\epsilon}{1+\epsilon} > \frac{1+\epsilon}{1+\epsilon}\epsilon = \epsilon, \log \frac{M_1}{M_0} = \log(1+\epsilon).$$

So $\int_{M_0}^{M_1} \frac{A(x)-x}{x^2} dx \geqslant \epsilon - \log(1+\epsilon)$. Now since $\epsilon < \frac{1}{2}$, $\epsilon - \log(1+\epsilon) > \frac{\epsilon^2}{3}$ say. In other words, we can't make the tail small. Lets tidy it up. Given $\epsilon > 0$, with $\epsilon < \frac{1}{2}$, let $\delta = \epsilon - \log(1+\epsilon) > \frac{\epsilon^2}{3} > 0$. Now let $M = M(\delta)$ be such that if $M_1 > M_0 > M$, $\int_{M_0}^{M_1} \frac{A(x)-x}{x^2} dx < \delta$. Then we must have $\frac{A(M_0)}{M_0} \leqslant 1+\epsilon$. Let $M_1 = M_0(1+\epsilon)$, since otherwise, a contradiction. Check similarly that for all $\epsilon > 0$, there exists x_2 so that if $x > x_2$, $\frac{A(x)}{x} > 1-\epsilon$.

Case 2. $\alpha = 0$: then $\int_1^\infty \frac{A(x)}{x^2} dx$ converges, then for every $\epsilon > 0$, there exists M so that if $x_0 > M$, $\int_{x_0}^\infty \frac{A(x)}{x} < \epsilon$. (*) But if $\frac{A(x_0)}{x_0} > \epsilon$, since $A(x)$ is weakly increasing, $\int_{x_0}^\infty \frac{A(x)}{x^2} \geqslant \int_{x_0}^\infty \frac{A(x_0)}{x^2} = \frac{A(x_0)}{x_0} > \epsilon$, contradicting (*). Hence $\frac{A(x_0)}{x_0} < \epsilon$, completing the proof. □

This puts the final nail in place in the proof of the PNT! Indeed, we have $\Psi(x) = \sum_{n \leqslant x} \Lambda(n)$. We've seen the Dirichlet series $\frac{\zeta'(s)}{\zeta(s)}$. We've seen that $\frac{\zeta'(1+it)}{\zeta(1+it)} < P(t)$ for $|t| > 2$, where P is a polynomial in $\log t$. Setting $f(s) = \frac{\zeta'(s)}{\zeta(s)}$, we get

(a) $f(s) = \sum_{n \geqslant 1} \frac{\Lambda(n)}{n^s}$ absolutely convergent for $\mathrm{Re}(s)$.

(b) $f(s) = \frac{1}{s-1} + c + (s-1)h(s)$, $h(s)$ is analytic in a region containing $\mathrm{Re}(s) \geqslant 1$.

(c) $\frac{\zeta'(1+it)}{\zeta(1+it)} < P(t)$.

Hence by the fundamental theorem, $\int_1^\infty \frac{\Psi(x)-x}{x^2} dx$ converges. Since $\Lambda(n) \geqslant 0$, $\Psi(x)$ is weakly increasing, non-negative, hence $\int_1^\infty \frac{\Psi(x)-x}{x^2} dx$ converges, so $\frac{\Psi(x)}{x} \to 1$ as $x \to \infty$, which is equivalent to $\pi(x) \sim \mathrm{Li}(x)$ by Abel summation.

Remark. Hadamard, De la Valée-Poussin in 1896, both proved PNT in much the same way. For decades, wondered "are all the complex analysis tools necessary"? 1948, Atle Selberg communicated some ideas to Paul Turan, who communicated then to Paul Erdös.

Primes act somewhat randomly. Often even if we can't prove what a sequence behavior must be we can determine what it ought to be by heuristic, probability argument. For example, the twin prime conjecture: infinitely many p so that p_1, p_2 are prime. The Goldbach conjecture: for all even $n \geqslant 4$, n is the sum of two primes. The twin Goldbach conjecture: for all sufficiently large n, $6n$ can be expressed as $6n = p + (6n - p)$ and $6n = (p+2) + (6n - p - 2)$ with $p, p+2, p+2, 6n-p-2$ are all prime. (When $p \geqslant 3$, $p = 6n-1$ or $p = 6n+1$).

Why should we expect these to be true? Heuristically, the "probability" that a number n is prime is about $\frac{1}{\log n}$ by PNT.

The probability and odd n is prime is $\frac{2}{\log n}$ since all even numbers are not prime. If we also rule out odd multiple of 3, we get more like $\frac{3}{\log n}$. So probability $6n-1, 6n+1$ are both prime is about $\frac{3}{\log(6n-1)} \frac{3}{\log(6n+1)} \approx \frac{9}{(\log n)^2}$. So $\sum_{n \leqslant x} \mathbb{1}_{\{6n \pm 1 \text{ prime}\}}$ should behave like $1 + \sum_{n \geqslant 2, n \leqslant x} \frac{9}{(\log n)^2}$.

In studying $\pi(x)$ we determined that $\sum_{p \text{ prime}, p \leqslant x} \frac{1}{p} \sim \log\log x$. How does $\sum_{6n \pm 1 \text{ prime}, n \leqslant x} \frac{1}{n}$ behave? Heuristically,

$$\sum_{p \text{ prime}, p \leqslant x} \frac{1}{p} \approx \sum_{2 \leqslant n \leqslant x} \frac{1}{n \log n} \approx \int_2^x \frac{1}{t \log t} dt = \log\log t|_2^x \sim \log\log x,$$

$$\sum_{6n\pm 1 \text{ prime}, n \leqslant x} \frac{1}{n} \text{``}\approx\text{''} \sum_{2 \leqslant n \leqslant x} \frac{9}{n(\log n)^2} \approx \int_2^x \frac{1}{t(\log t)^2} dt,$$

which is bounded as $x \to \infty$.

Brun invented sieve methods to prove that $\sum_{6n \pm 1 \text{ prime}, n \leqslant x} \frac{1}{n}$ converges. How many twin primes $\leqslant x$ should we see? Should look like $\frac{cx}{(\log x)^2}$.

AIM: Riemann Hypothesis, Statement equivalent to RH.

Turning to Goldbach: $2n = p + q$ has $g(n)$ solutions, where $g(n) > 0$ for $n \geqslant 2$ and $g(n) \sim \frac{cn}{(\log n)^2} + L(n)$.

Twin prime conjection:

$$\pi_2(n) \sim \frac{c' n}{(\log n)^2}.$$

How many twin-prime pair representations of $6n$ do we get? $2n = p + q = (p+2) + (q-2)$, $p, q, p+2, q-2$ are all prime. Need $p, p+2$ prime (so except for 3,5). So $p = 6k-1$, $p+2 = 6k+1$ and $q = 6l+1$, $q-2 = 6l-1$. So $p+q = 6k-1+6l+1 = 6(k+l)$. Hence $2n \equiv 0 \pmod 6$.

Exercise 5.57. Heuristically, how $tg(6n)$ should grow? Also, how much spread in the graph should we see?

Chapter 6

Generating Functionology

Let $n \in \mathbb{Z}_{\geqslant 0}$.

Definition 6.1. The *power series* associated with this will have a varible z and will be denoted by

$$f(z) = a_0 + a_1 z + \cdots + a_n z^n + \cdots = \sum_{n \geqslant 0} a_k z^k.$$

Remark. Under certain circumstances, $f(z)$ will have nice properties, e.g., when evaluating $f(z)$ at $z = 1$ or regarding f as a function of a complex variable or a rational function, etc.

Example 6.2. Let $a_n = b^n$ for any $n \geqslant 0$, where b is a constant. Then $f(z) = \sum_{n \geqslant 0} b^n z^n$. So $bzf(z) = \sum_{n \geqslant 0} b^{n+1} z^{n+1} = \sum_{n \geqslant 1} b^n z^n = f(z) - 1$, i.e., $f(z) = \frac{1}{1-bz}$. So we have a nice representation for $f(z)$ as a quotient of polynomial.

Example 6.3. Let $a_0 = 1$ and $a_n = ba_{n-1}$ for any $n \geqslant 1$, where b is a constant. Then $f(z) = 1 + \sum_{n \geqslant 1} ba_{n-1} z^n = 1 + bz \sum_{n \geqslant 1} a_{n-1} z^{n-1} = 1 + \sum_{n \geqslant 0} a_n z^n = 1 + bzf(z)$, i.e., $f(z) = \frac{1}{1-bz}$.

Example 6.4. Let $a_0 = 1$, $a_1 = 2$ and $a_n = a_{n-1} + a_{n-2}$ for any $n \geqslant 2$.

(a) Approch 1: $f(z) = 1 + 2z + \sum_{n \geqslant 2}(a_{n-1} + a_{n-2}) z^n = 1 + 2z + \sum_{n \geqslant 2} a_{n-1} z^n + \sum_{n \geqslant 2} a_{n-2} z^n = 1 + 2z + z \sum_{n \geqslant 1} a_n z^n + z^2 \sum_{n \geqslant 0} a_n z^n = 1 + 2z + z(f(z) - 1) + z^2 f(z)$, i.e., $f(z) = \frac{1+z}{1-z-z^2}$.

(b) Approch 2: Since $a_n z^n = a_{n-1} z^n + a_{n-2} z^n$ for any $n \geqslant 2$, we have $\sum_{n \geqslant 2} a_n z^n = z \sum_{n \geqslant 1} a_n z^n + z^2 \sum_{n \geqslant 0} a_n z^n$, i.e., $f(z)$, i.e., $f(z) - 1 - 2z = z(f(z) - 1) + z^2 f(z)$, i.e., $f(z) = \frac{1+z}{1-z-z^2}$.

In $\mathbb{C}$, we can write $(1 - z - z^2) = (1 - \alpha z)(1 - \beta z)$, where $\alpha = \frac{1+\sqrt{5}}{2}$ and $\beta = \frac{1-\sqrt{5}}{2}$. So we can write $\frac{1+z}{1-z-z^2} = \frac{A}{1-\alpha z} + \frac{B}{1-\beta z} = A \sum_{n \geqslant 0} \alpha^n z^n + B \sum_{n \geqslant 0} \beta^n z^n$ for some $A, B \in \mathbb{C}$. Thus, $a_n = A\alpha^n + B\beta^n = A\left(\frac{1+\sqrt{5}}{2}\right)^n + B\left(\frac{1-\sqrt{5}}{2}\right)^n$ for any $n \geqslant 0$. Since $B\left(\frac{1-\sqrt{5}}{2}\right)^n \to 0$ as $n \to \infty$, we have the n^{th} Fibonacci number is the nearest integer to $A\left(\frac{1+\sqrt{5}}{2}\right)^n$ when n is large enough.

Example 6.5. Let $a_0 = 1$, $a_1 = 2$ and $a_n = a_{n-1} + a_{n-2}$ for any $n \geqslant 2$. Set $\mu_1 = \begin{bmatrix} a_1 \\ a_0 \end{bmatrix}$ and $\mu_n = \begin{bmatrix} a_n \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} a_{n-1} + a_{n-2} \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ a_{n-2} \end{bmatrix}$ for any $n \geqslant 2$. So $\mu_n = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{n-1} \mu_1$.

Remark. What can generating functions do for us?

(a) Sometimes allows us to find an exact value for a count of some type of object.

(b) When that's not possible, perhaps can we obtain a recurrence? E.g., how many set partitions of $\{1, \ldots, n\}$ are there?

(c) Find averages of values and other statistical properties of a sequence.

(d) Find asymptotics for growth rate of sequences.

(e) Prove nice properties of the sequence, e.g., monotonic log concave unimodal.

(f) Prove identities, e.g., $\sum_{j=0}^{n} \binom{n}{j}^2 = \binom{2n}{n}$.

Remark (Four Colors Theorem)**.**

Theorem 6.6 (Tutte)**.**

Theorem 6.7.

Corollary 6.8.

Definition 6.9. Dirichlet generating function: $\{a_n\}_{n \geqslant 1} \longleftrightarrow f(s) = \sum_{n \geqslant 1} \frac{a_n}{n^s}$.

Definition 6.10. (a) An ordinary generating function (ogf): $\{a_n\}_{n \geqslant 0} \longleftrightarrow f(x) = \sum_{n \geqslant 0} a_n x^n$. Purely formal series.
Useful for counting unlabelled objects, e.g., binary sequence with restrictions, partitions of integers.

(b) An exponential generating function (egf): $\{a_n\}_{n \geqslant 0} \longleftrightarrow f(x) = \sum_{n \geqslant 0} \frac{a_n x^n}{n!}$.
Useful for counting labelled objects, e.g., set partition: labelled trees, hands of cards.

Remark. Generally, egf's tend to be more useful when the number of objects grows faster than exponentially in n.

Example 6.11. (a) $a_n = \frac{1}{n!}$ for any $n \geqslant 0$, ogf $f(x) = e^x = \sum_{n \geqslant 0} \frac{x^n}{n!}$.

(b) $a_n = 1$ for any $n \geqslant 0$, egf $f(x) = e^x = \sum_{n \geqslant 0} \frac{x^n}{n!}$.

Remark. Note $e^x = \sum_{n \geqslant 0} \frac{x^n}{n!}$ for all $x \in \mathbb{C}$. Let $x > 0$. Then $\frac{x^n}{n!} < e^x$, i.e., $n! > \frac{x^n}{e^x}$. Let $x = n$, $n! > (\frac{n}{e})^n$. The truth is $n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$.
Can we improve this? Maximize $x^n e^{-x}$. Since $\frac{d}{dx}(x^n e^{-x}) = (nx^{n-1} - x^n)e^{-x}$, it is maximized at $x = n$.
Can we do a better job of comparing e^x and the terms around $\frac{x^n}{n!}$? Compare $\frac{x^{n+k}}{(n+k)!}$ to $\frac{x^n}{n!}$ at $x = n$ for $k = -l, \ldots, l$, where l is something like $n^{1/2}$. This method works very well. For many combinatorial sequence it gives a lower bound for an ogf from reality by $c\sqrt{n}$.

Remark (Notation, due to Knuth)**.** If $f(x) = \sum_{n \geqslant 0} a_n x^n$, then

$$[x^n]f(x) = a_n, [x^n] = \text{“the coefficient of } x^n \text{ in”}.$$

If $g(x) = \sum_{n \geqslant 0} b_n \frac{x^n}{n!}$, then $[\frac{x^n}{n!}]g(x) = b_n$.

Remark. Some useful series to know: $\sum_{n\geqslant 0} x^n = 1+x+x^2+\cdots = \frac{1}{1-x}$. Euler introduces that we can operate $\sum_{n\geqslant 0} a_n x^n$ by $x\frac{d}{dx}$: $x\frac{d}{dx}\sum_{n\geqslant 0} a_n x^n = \sum_{n\geqslant 1} n a_n x^n$. So $\sum_{n\geqslant 1} n x^n = x\frac{d}{dx}\frac{1}{1-x} = \frac{x}{(1-x)^2}$. Moreover, we can integrate it to get $-\log(1-x) = \int_0^x \frac{1}{1-t}dx = \int_0^x \sum_{n\geqslant 0} t^n dt = \sum_{n\geqslant 0}\int_0^x t^n dt = \sum_{n\geqslant 0}\frac{x^{n+1}}{n+1} = \sum_{n\geqslant 1}\frac{x^n}{n}$.

Theorem 6.12 (Binomial Theorem)**.** *We have the following versions.*

(a) Version A. $\sum_{k=0}^n \binom{n}{k} x^k = (1+x)^n$*, where* $\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}$ *This formula works even if* n *is not a nonnegative integer.*

(b) Version B. If $n \notin \mathbb{N}$*, then* $(1+x)^n = \sum_k \binom{n}{k} x^k$*. Useful convention: sum over all* k *for which the summand is defined.* $(1+x)^{-1} = \sum_k \binom{-1}{k} x^k$*, where* $\binom{-1}{k} = \frac{(-1)(-2)\cdots(-k)}{k!} = (-1)^k$*. So we have* $\frac{1}{1+x} = (1+x)^{-1} = \sum_{k\geqslant 0}(-1)^k x^k$*. Then* $\frac{1}{1-x} = (1-x)^{-1} = \sum_{k\geqslant 0}(-1)^k(-x)^k = \sum_{k\geqslant 0} x^k$*. Let* $m \in \mathbb{N}$*. Then* $\frac{1}{(1-x)^m} = (1-x)^{-m} = \sum_{k\geqslant 0}\binom{-m}{k}(-1)^k x^k$*, where*

$$\binom{-m}{k} = \frac{(-m)(-m-1)\cdots(-m-k+1)}{k!} = (-1)^k\frac{m\cdots(m+k-1)}{k!} = (-1)^k\binom{m+k-1}{m-1}.$$

So $\frac{1}{(1-x)^m} = (1-x)^{-m} = \sum_{k\geqslant 0}\binom{m+k-1}{m-1}x^k$*.*

(c) Version C. $\frac{1}{(1-x)^{1/2}} = (1-x)^{-1/2} = \sum_{k\geqslant 0}(-1)^k\binom{-1/2}{k}x^k$*, where*

$$\binom{-\frac{1}{2}}{k} = \frac{(-\frac{1}{2})\cdots(-1/2-k+1)}{k!} = (-1)^k\frac{1\cdot 3\cdots(2k-1)}{2^k k!} = (-1)^k\frac{2k!}{2^k k! 2\cdot 4\cdot\cdots\cdot 2k} = \frac{(-1)^k}{4^k}\binom{2k}{k}.$$

So $\frac{1}{(1-x)^{1/2}} = (1-x)^{-1/2} = \sum_{k\geqslant 0}\frac{1}{4^k}\binom{2k}{k}x^k$ *and then* $(1-4x)^{-1/2} = \sum_{k\geqslant 0}\binom{2k}{k}x^k$*.*

(d) Version C. Let z *be a variable.* $(1+x)^z = \sum_{k\geqslant 0}\binom{z}{k}x^k$*. This is a bivariate generating function and its terms are monomials of the form* $a_{ij}x^i z^j$*, where* $j \geqslant i$*. We will revisit these coefficients* a_{ij}*'s later.*

Remark. $e^x = \sum_{k\geqslant 0}\frac{x^k}{k!}$. Why is the first binomial theorem true? Note if $\binom{n}{k}$ denotes the number of k-subsets of $\{1,\dots,n\}$, then $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$. Obverse that $\binom{0}{0} = \frac{0!}{0!0!}$ is true and that $f(n,k) = \frac{n!}{k!(n-k)!}$ satisfies $f(n,k) = f(n-1,k) + f(n-1,k-1)$. Hence by induction on n, $f(n,k) = \binom{n}{k}$. Also, note $\#\mathcal{P}(\{1,\dots,n\}) = 2^n$. So $\sum_{k=0}^n\binom{n}{k}x^k$ is the generating function for subsets of $\{1,\dots,n\}$, where k marks the size of the subset $(1+x)^n = \underbrace{(1+x)\cdots(1+x)}_{n \text{ times}}$ corresponding to picking a subset of $\{1,\dots,n\}$ one element at a time. Hence $(1+x)^n = \sum_{k=0}^n\binom{n}{k}x^k$.
Note that

$$\sum_{n=0}^{\infty}\sum_{k=0}^{n}\binom{n}{k}x^k y^n = \sum_{n=0}^{\infty}(1+x)^n y^n = \sum_{n=0}^{\infty}((1+x)y)^n = \frac{1}{1-(1+x)y} = \frac{1}{1-y-xy}$$
$$= \frac{1}{\left(1-\frac{xy}{1-y}\right)(1-y)} = \frac{1}{1-y}\sum_{k\geqslant 0}x^k\left(\frac{y}{1-y}\right)^k = \sum_{k\geqslant 0}x^k\frac{y^k}{(1-y)^{k+1}}.$$

Since x^k term dominates, $\binom{n}{k} = [y^n x^k]\frac{1}{1-y-xy} = [y^n]\frac{y^k}{(1-y)^{k+1}} = [y^{n-k}]\frac{1}{(1-y)^{k+1}}$. This implies the binomial theorem for negative exponents that we derived eariler.

6.1 Set partions and Permutation

Definition 6.13. A set *partion* of $\{1,\dots,n\}$ into k parts is a set $\{S_1,\dots,S_k\}$ such that $\emptyset \neq S_j \subseteq \{1,\dots,n\}$, $S_i \cap S_j = \emptyset$ for any $1 \leqslant i \neq j \leqslant k$ and $\bigsqcup_{j=1}^{k} S_j = \{1,\dots,n\}$.

Remark (Notation)**.** Kruth, Wif and only if, etc. : $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}$. Stanley: $S(n,k)$.

Definition 6.14. The numbers $S(n,k)$ are called Stirling numbers of the second kind.

Remark. facts

(a) $S(0,0) = 1$.

(b) $S(n,0) = 0$ for any $n \in \mathbb{N}$.

(c) $S(n,1) = \begin{cases} 1 & \text{if } n \geqslant 1 \\ 0 & \text{if } n = 0 \end{cases}$.

(d) $S(n,n) = 1$ for any $n \in \mathbb{Z}^+$.

(e) $S(n,n+t) = 0$ for any $n \in \mathbb{Z}^+$ and $t \in \mathbb{N}$.

Remark. How can we compute a table of $S(n,k)$? Consider a set partition of $\{1,\dots,n\}$ into k parts. Either n is a part of by itself, or n is a part of size $\geqslant 2$. So $S(n,k) = S(n-1,k-1) + kS(n-1,k)$. Then we can compute $S(n,k)$ for all k.

$$\begin{array}{lll} S(0,0)=1 & & \\ S(1,0)=0 & S(1,1)=1 & \\ S(2,0)=0 & S(2,1)=1 & S(2,2)=1 \\ & S(3,1)=1 & S(3,2)=3 \\ & & S(3,3)=1. \end{array}$$

Definition 6.15. Define the *Bell number* $B_0 = 1$, $B_n = \sum_{k=0}^{n} S(n,k)$ for any $n \in \mathbb{N}$.

Remark. Is there a formula for B_n? How fast does B_n grow?
Is there a nice form for the ogf for B_n?
Is there a nice form for the egf for B_n?
Is there a nice way for to compute B_n? E.g., a recurrence.

Remark. What about generating functions for $S(n,k)$? $A_n(y) = \sum_k S(n,k)y^k$, $B_k(x) = \sum_n S(n,k)x^n$ and $C(x,y) = \sum_{n,k} S(n,k)x^n y^k$.

(a) $B_k(x) = x\sum_n S(n-1,k-1)x^{n-1} + kx\sum_n S(n-1,k)x^{n-1}$, i.e., $B_k(x) = xB_{k-1}(x) + kxB_k(x)$. So $B_k(x)(1-kx) = xB_{k-1}(x)$. Then $B_0(x) = 1$, $B_1(x) = \frac{1}{1-x}$, $B_2(x) = \frac{x^2}{(1-x)(1-2x)}$, $\cdots$, $B_k(x) = \frac{x^k}{(1-x)\cdots(1-kx)}$. Use partial fractions to expand RHS as $\sum_{j=1}^{k} \frac{\alpha_j}{1-jx}$. Fix $1 \leqslant r \leqslant k$, multiply by $(1-rx)$ and evaluate at $x = \frac{1}{r}$, $B_k(x)$ becomes

$$\frac{(1/r)^k}{(1-1/r)\cdots(1-(r-1)/r)(1-(r+1)/r)\cdots(1-k/r)} = \frac{1}{r}\frac{1}{(r-1)\cdots(1)(-1)\cdots(-(k-r))}.$$

RHS is α_r. So $\alpha_r = \frac{(-1)^{k-r}}{r!(k-r)!} = \frac{(-1)^{k-r}}{k!}\binom{k}{r}$. Hence $S(n,k) = [x^n]B_k(x) = \sum_{j=1}^{k}[x^n]\frac{\alpha_j}{1-jx} = \sum_{j=1}^{n}\alpha_j j^n = \sum_{j=1}^{n}\frac{(-1)^{k-j}}{k!}\binom{k}{j}j^n$. When $k > n$, $S(n,k) = 0$.
$k!S(n,k) = \sum_{j=1}^{k}(-1)^{k-j}\binom{k}{j}j^n$ is the number of onto functions from $\{1,\ldots,n\}$ to $\{1,\ldots,k\}$.

(b) $A_n(y) = \sum_k S(n,k)y^k = y\sum_k S(n-1,k-1)y^{k-1} + \sum_{k=0}^{n} kS(n-1,k)y^k = yA_{n-1}(y) + y\frac{d}{dy}A_{n-1}(y)$, i.e., $A_n(y) = y(1+D_y)A_{n-1}(y)$.
Define

$$b_n := \sum_{k=0}^{\infty} S_{n,k} = \sum_{k=0}^{\infty}\frac{1}{k!}\sum_{j=1}^{k}(-1)^{k-j}\binom{k}{j}j^n = \sum_{j=0}^{\infty}\frac{j^n}{j!}\sum_{k=j}^{\infty}\frac{(-1)^{k-j}}{(k-j)!} = \frac{1}{e}\sum_{j=0}^{\infty}\frac{j^n}{j!}.$$

So we try to find a nice expression for $B(x) = \sum_{n\geqslant 0}\frac{b_n x^n}{n!}$.

$$\begin{aligned} B(x) - 1 &= \frac{1}{e}\sum_{n\geqslant 1}\sum_{j=0}^{\infty}\frac{j^n}{n!j!}x^n = \frac{1}{e}\sum_{j\geqslant 1}\frac{1}{j!}\sum_{n\geqslant 1}\frac{(jx)^n}{n!} = \frac{1}{e}\sum_{j\geqslant 1}\frac{1}{j!}(e^{jx}-1) \\ &= \frac{1}{e}\sum_{j\geqslant 1}\frac{(e^x)^j}{j!} - \frac{1}{e}\sum_{j\geqslant 1}\frac{1}{j!} = \frac{1}{e}(e^{e^x}-1) - \frac{1}{e}(e-1) = e^{e^x-1} - 1. \end{aligned}$$

So $B(x) = e^{e^x-1}$.
Why e^{e^x-1} is more natural than e^{e^x}. Note $e^{e^x-1} = \sum_{j=0}^{\infty}\frac{(e^x-1)^j}{j!} = \sum_{j=0}^{\infty}\frac{(x(1+\frac{x}{2!}+\frac{x^2}{3!}+\cdots))^j}{j!}$. So the j^{th} term in the expansion has all powers of $x \geqslant j$. To compute $[\frac{x^n}{n!}]\sum_{j=0}^{\infty}\frac{(x(1+\frac{x}{2!}+\frac{x^2}{3!}+\cdots))^j}{j!}$, we only have to consider finitely many different j's and only finite partitions of $(1+\frac{x}{2!}+\frac{x^2}{3!}+\cdots)^j$.
If f is a generating function, then we will be able to consider $e^{f(x)}$ as a generating function if and only if $f(0) = 0$. So $\sum_{n\geqslant 0} b_n\frac{x^n}{n!} = B(x) = e^{e^x-1}$. We want to obtain from this expression a recurrence for b_n. Apply $x\frac{d}{dx}$ to the log of both sides, we have $\sum_{n\geqslant 0} nb_n\frac{x^n}{n!} = xe^x e^{e^x-1} = xe^x B(x)$. So $nb_n = [\frac{x^n}{n!}]xe^x B(x) = n[\frac{x^{n-1}}{(n-1)!}]e^x B(x)$. Hence if $n > 0$, $b_n = [\frac{x^{n-1}}{(n-1)!}]e^x B(x) = [\frac{x^{n-1}}{(n-1)!}]e^x e^{e^x-1}$. Since $e^x = \sum_{i\geqslant 0} 1\frac{x^i}{i!}$ and $e^{e^x-1} = \sum_{j\geqslant 0} b_j\frac{x^j}{j!}$, we have $e^x e^{e^x-1} = \sum_{n=0}^{\infty}\frac{x^n}{n!}\sum_{k=0}^{n}\binom{n}{k}b_k$. So $[\frac{x^{n-1}}{(n-k)!}]e^x e^{e^x-1} = \sum_{k=0}^{n-1}\binom{n-1}{k}b_k$. Thus, $b_n = \sum_{k=0}^{n-1}\binom{n-1}{k}b_k$. Then $b_0 = 1$, $b(1) = 1, \cdots$.
How fast does b_n grow? Note $\frac{b_n}{n!}x^n < e^{e^x-1}$ for all $x \in \mathbb{R}$, i.e., $\frac{b_n}{n!} < \frac{e^{e^x}-1}{x^n}$. Let $\frac{d}{dx}\frac{e^{e^x-1}}{x^n} = \frac{e^{e^x-1}(xe^{e^x}-n)}{x^{n+1}} = 0$. Since $e^{e^x-1} > 0$ for any $x \in \mathbb{R}$, we have $xe^x - n = 0$, i.e., $x \approx \log n - \log\log n \approx \log n$. Clearly, $\frac{e^{e^x-1}}{x^n}$ obtains its minimal around $\log n$. So $\frac{b_n}{n!} < \frac{e^{e^{\log n}-1}}{(\log n)^n}$, i.e., $b_n < n!\frac{e^{e^{\log n}-1}}{(\log n)^n} = n!\frac{e^{n-1}}{(\log n)^n} = \frac{1}{e}\left(\frac{e}{\log n}\right)^n n! \approx \frac{1}{e}\left(\frac{n}{\log n}\right)^n\sqrt{2\pi n}$.

Remark. (a)

$$\begin{aligned} \left(\sum_{i=0}^{\infty}c_i x^i\right)\left(\sum_{j=0}^{\infty}d_j x^j\right) &= \sum_{n=0}^{\infty}\sum_{k=0}^{n}c_k x^k d_{n-k}x^{n-k} = \sum_{n=0}^{\infty}\sum_{k=0}^{n}c_k d_{n-k}x^n \\ &= \sum_{n=0}^{\infty}(c_0 d_n + c_1 d_{n-1} + \cdots + c_n d_0)x^n. \end{aligned}$$

(b)

$$\left(\sum_{i=0}^{\infty} a_i \frac{x^i}{i!}\right)\left(\sum_{j=0}^{\infty} b_j \frac{x^j}{j!}\right) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} a_k \frac{x^k}{k!} b_{n-k}\frac{x^{n-k}}{(n-k)!} = \sum_{n=0}^{\infty}\frac{x^n}{n!}\sum_{k=0}^{n}\binom{n}{k} a_k b_{n-k}.$$

Example 6.16. Can we give a combinatorial explanation of $b_n = \sum_{k=0}^{n-1}\binom{n-1}{k} b_k$?

Example 6.17. Using the $x\frac{d}{dx}$ operator, we can get a useful information. Let $f(x) = \sum_{n\geqslant 0} a_n x^n$ and $g(x) = \sum_{n\geqslant 0} b_n x^n$.

(a) Consider $f(x) = e^{g(x)}$. Assume $g(0) = 0$. Then $\log f = g$, so $x\frac{d}{dx}\log f(x) = x\frac{d}{dx} g(x)$, i.e., $\frac{xf'(x)}{f(x)} = xg'(x)$, i.e., $xf'(x) = f(x)(xg'(x))$. Then we have

$$\sum_{n\geqslant 1} n a_n x^n = \left(\sum_{n\geqslant 0} a_n x^n\right)\left(\sum_{n\geqslant 1} n b_n x^n\right).$$

So $na_n = \sum_{k=0}^{n-1} a_k (n-k) b_{n-k}$ for any $n \in \mathbb{N}$. Thus, just as $e^{e^x - 1}$, we can compute the coefficients of e^g recursively.

(b) Consider $f(x) = g(x)^k$, here $g(0)$ is not assumed to be 0 and if $k < 0$, it will be assumed to be 1. Similarly, $xf'(x) = kxg'(x)g(x)^{k-1}$, i.e., $xf'(x)g(x) = kxg'(x)g(x)^k = kxg'(x)f(x)$. So $\sum_{n\geqslant 1} n a_n x^n \sum_{n\geqslant 0} b_n x^n = \sum_{n\geqslant 1} k n b_n x^n \sum_{n\geqslant 0} a_n x^n$. Hence $\sum_{j=1}^{n} a_j b_{n-j} = \sum_{j=1}^{n} k j b_j a_{n-j}$. So $a_n b_0 = \sum_{j=1}^{n} k j b_j a_{n-j} - \sum_{j=1}^{n-1} a_j b_{n-j}$ for any $n \in \mathbb{N}$. If $g(0) = b_0 = 0$, then we can't do this. But if $k \in \mathbb{N}$ and $b_r \neq 0$ is the first nonzero coefficient of g, then we can write $g(x) = b_r x^r h(x)$ with $h(0) = 1$. Then $f(x) = g(x)^k = b_r^k x^{rk} h(x)^k$. So we can compute, for example, for $k = -1$, $f(x) = \frac{1}{g(x)}$ provided $g(0) \neq 0$, i.e., $f(x)g(x) = 1$, so $[x^n]f(x)g(x) = \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{if } n > 0 \end{cases}$. So $a_0 = \frac{1}{b_0}$ and $\sum_{k=0}^{n} a_k b_{n-k} = 0$, i.e., $a_n = -\sum_{k=0}^{n-1} a_k b_{n-k}$ for any $n \in \mathbb{N}$. When g is a polynomial of degree d with constant term 1, $a_n = -\sum_{k=\max\{n-d,0\}}^{n-1} a_k b_{n-k}$ for any $n \in \mathbb{N}$, this shows that a_n satisfies a d-term recurrence.

(c) Consider $f(x) = \log g(x)$. Assume $g(0) = 1$. Then $xf'(x) = x\frac{d}{dx}\log g(x) = x\frac{g'(x)}{g(x)}$, i.e., $xf'(x)g(x) = xg'(x)$, etc. Alternately $f(x) = \log(1 + xh(x)) = \sum_{k=1}^{\infty}(-1)^{k-1}\frac{x^k h(x)^k}{k}$.

(d) When can we compute $f(g(x))$ as a generating function?

Definition 6.18. An *integer partition* $\lambda \vdash n$ is a nonincreasing sequence $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_k \geqslant 1$ with $\lambda_1, \ldots, \lambda_k \in \mathbb{N}$ such that $n = \lambda_1 + \cdots + \lambda_k$. Define $\underline{i}(\lambda) = (i_1, i_2, \cdots)$, where i_j is the number of copies of j in λ.

Remark. (a) Note # parts of λ is $\sum_{i\geqslant 0} i_j$.

(b) $\underline{i}(\lambda)$ has only finitely many nonzero terms. For example, the partition $5 = 3 + 1 + 1$ can be written as $(2, 0, 1, 0, 0, \cdots)$.

(c) There is a bijection between integer partition and nonnegative integer sequences with finitely many positive entries.

(d) Partitions were studied extensively by Euler.

Definition 6.19. Define $p(n)$ be the number of integer partitions of n.

Remark. $n=0$, the sum of no parts, $p(0)=1$;
$n=1$, 1, $p(1)=1$;
$n=2$, 2, $1+1$, $p(2)=2$;
$n=3$, $3, 2+1, 1+1+1$, $p(x)=3$;
$n=4$, $4, 3+1, 2+2, 2+1+1, 1+1+1+1$, $p(4)=5$;
$n=5$, $5, 4+1, 3+2, 3+1+1, 2+2+1, 2+1+1+1, 1+1+1+1+1$, $p(5)=7$;
Can we get an expression for $p(x)=\sum_{n=0}^{\infty} p(n)x^n$?

Remark. Choose a partition of n by picking i_1, picking i_2, $\cdots$.

(a) How many ways can we write n as a sum of 1's? The generating function for this is $\sum_{n\geqslant 0} 1x^n = \frac{1}{1-x} =: \sum_{k\geqslant 0} a_k x^k$.

(b) How many ways can we write n as a sum of 2's? 0 if n is oddr; and 1 if n is even. $gf = \sum_{n \text{ even}} x^n = \sum_{k\geqslant 0} x^{2k} = \frac{1}{1-x^2} =: \sum_{k\geqslant 0} b_k x^k$.

(c) How many ways can we write n as a sum of 1's and 2's?

$$\sum_k (\text{\# ways of writing } k \text{ as sum of 1's})(\text{\# ways of writing } n-k \text{ as a sum of 2's})$$

$$= \sum_{k\geqslant 0} a_k b_{n-k} = [x^n]\left(\sum_{k\geqslant 0} a_k x^k\right)\left(\sum_{k\geqslant 0} b_k x^k\right).$$

Remark. Suppose we have two sets S, T of objects with weights.
egf for g is $\sum_{k\geqslant 0} x^k(\text{\# objects of weights } k \text{ in } S) = \sum_{k\geqslant 0} a_k x^k$.
egf for T is $\sum_{k\geqslant 0} x^k(\text{\# objects of weights } k \text{ in } T) = \sum_{k\geqslant 0} b_k x^k$.
Then the number of pairs $(s,t), s\in S, t\in T$ of total weight n is

$$\sum_{k=0}^{n} a_k b_{n-k} = [x^n]\left(\sum_{k\geqslant 0} a_k x^k\right)\left(\sum_{k\geqslant 0} b_k x^k\right).$$

The # ways to represent n as a sum of 1's and 2's in non-decreasing order is the # ways of choosing i_1, i_2 so that $i_1 + 2i_2 = n$, which is $[x^n]\frac{1}{x-1}\frac{1}{1-x^2}$. # ways of writing $\sum_{n\geqslant 0} p(n)x^n = \prod_{j=1} \frac{1}{1-x^j}$. $[x^n]\prod_{j=1}^{\infty}\frac{1}{1-x^j} = [x^n]\prod_{j=1}^{n}\frac{1}{1-x^j} = [x^n]\prod_{j=1}^{n}(1+x^j+2^{2j}+\cdots+x^{\lfloor n/j\rfloor j})$, which is the product of n polynomials each of degree $leqn$. So for any n, this gives a finite algorithm for computing $p(n)$ and indeed for generating all $p(n)$ partitions, since if we consider $\prod_{j=1}^{\infty}\frac{1}{1-x^j y_j}$ and truncate products and series. $[x^n]\frac{1}{1-x^j y_j}$ will be a polynomial in $y_1, \dots, y_n$. each monomial will be $y_1^{i_1}\cdots y_n^{i_n}$, where $(i_1, \dots, i_n, 0, 0, \cdots)$ is the sequence of multiplicities for λ. $f(x) = \prod_{j=1}^{\infty}\frac{1}{1-x^j}$, $\log f(x) = \sum_{j=1}^{\infty} -\log(1-x^j) =$. Note $-\log(1-x^j) = \sum_{k=1}^{\infty}\frac{{x^j}^k}{k}$,

$$[x^n]\sum_{j=1}^{\infty} -\log(1-x^j) = \sum_{d|n}\frac{1}{d} = \frac{1}{n}\left(\sum_{d|n}\frac{n}{d}\right) = \frac{1}{n}\sum_{d|n} d = \frac{1}{n}\sigma(n).$$

So $\log f(x) = \sum_{n\geqslant 1} x^n \frac{\sigma(n)}{n} = L(x)$, i.e., $f = e^L$.

Remark. Last time we saw how to compute the exponential of a power series with constant coefficients 0.
Exercise: find a recurrence for $p(n)$ in terms of $p(0), p(1), \ldots, p(n-1)$.
Faa DiBrun considered $f(g(x))$.

(a) $g(x)$ is a power series, $f(x)$ is a polynomial,

(b) $f(x)$ is a power series, $g(x)$ is a power series with $g(0) = 0$.

Coefficient of x^n is $[x^n]f(g(x)) = \frac{1}{n!}\left(\frac{d}{dx}\right)^n f(g(x))\Big|_{x=0}$. So it is enough to be able to compute $\left(\frac{d}{dx}\right)^n f(g(x))$. Note $\frac{d}{dx}f(g(x)) = f'(g(x))g'(x)$, $\frac{d}{dx}\left(\frac{d}{dx}f(g(x))\right) = f''(g(x))(g'(x))^2 + f'(g(x))g''(x)$,

$$\begin{aligned}\left(\frac{d}{dx}\right)^3 f(g(x)) &= f'''(g')^3 + 2f''g''g' + f''g''g' + f'g''' = f^{(3)}{g^{(1)}}^3 + 3f^{(2)}g^{(2)}g^{(1)} + f^{(1)}g^{(3)} \\ &= f^{(3)}g_{1,1,1} + 3f^{(2)}g_{2,1} + f^{(1)}g_3 = \sum_{\lambda\vdash 3} c_\lambda f^{(\pi(\lambda))}g_\lambda,\end{aligned}$$

where λ is a partiton of n, $\pi(\lambda) = \#$ partitiones in λ and if $\lambda = \lambda_1 + \cdots + \lambda_k$, $g_\lambda = g^{(\lambda_1)}\cdots g^{(\lambda_k)}$.
$\left(\frac{d}{dx}\right)^n f(g(x)) = \sum_{\lambda\vdash n} c_\lambda f^{(\pi(\lambda))} g_\lambda$.
Exercise: What are c_λ's? Put $f(y) = e^y$, compute c_λ for $\lambda \mapsto n$ for $n \leqslant 5$. Conjecture a form for c_λ in terms of $\lambda_1, \ldots, \lambda_k$ and $i_1, \ldots, i_n$. Prove you formula works.

6.2 The Magic of Power Series

Remark. The number of partitions of n into odd parts = the number of partitions of n into distinct parts.
Partitions into distinct parts. $i_1 = 0$ or 1, $i_2 = 0$ or 1, $i_3 = 0$ or 1.
Let $f(x) = (1+x)(1+x^2)(1+x^3)\cdots$. Then

$$\begin{aligned}\frac{1-x}{1-x}f(x) &= \frac{(1-x^2)(1+x^2)(1+x^3)(1+x^4)}{1-x} = \frac{(1-x^4)(1+x^3)(1+x^4)(1+x^5)(1+x^6)\cdots}{1-x} \\ &= \frac{(1-x^4)(1-x^3)(1+x^3)(1+x^4)(1+x^5)(1+x^6)\cdots}{(1-x)(1-x^3)} \\ &= \frac{(1+x^5)(1-x^6)(1-x^8)(1+x^6)(1+x^7)\cdots}{(1-x)(1-x^3)} \\ &= \cdots = \frac{1}{(1-x)(1-x^3)(1-x^5)(1-x^7)\cdots} \\ &= \frac{(1+x)}{(1-x^2)(1-x^3)(1-x^5)\cdots} = \frac{(1+x)(1+x^2)}{(1-x^3)(1-x^4)(1-x^5)\cdots} \\ &= \frac{(1-x^2)(1-x^4)(1-x^6)(1-x^8)\cdots}{(1-x)(1-x^2)(1-x^3)(1-x^4)(1-x^5)(1-x^6)\cdots} \\ &= (1+x)(1+x^2)(1+x^3)(1+x^4)\cdots.\end{aligned}$$

Remark (Euler Maclaurin)**.** Suppose we wish to estimate $\sum_{k=a}^{b} f(k)$, where f is a nice smooth function. We might recall from calculus II that is similar to $\int f(x)dx$ integrating over an interval near $[a,b]$ of width $b-a+1$. This leads to the equation of how big is

$$f(k) - \int_{k-1}^{k} f(x)dx \approx \int_{k-1}^{k} (f(k)-f(x))d(x-c_k) = (x-c_k)(f(k)-f(x))|_{k-1}^{k} + \int_{k-1}^{k} (x-c_k)f'(x)dx$$
$$= -(x-c_k)f(x)\,|_{k-1}^{k} + \int_{k-1}^{k} (x-c_k)f'(x)dx.$$

We're going to want to repeat the process of using integration by parts. It turns out that it is a very good thing if we choose c_k so that $\int_{k-1}^{k}(x-c_k)dx = 0$, i.e., $c_k = k-1/2$ for each k. Then $x-c_k = x-(k-1/2) = x-(k-1)-1/2 = \{x\}-1/2$ for $x \in [k-1,k]$, where $\{x\}$ is the fractional part of x. Then

$$\int_{k-1}^{k} (f(k)-f(x))dx = -(\{x\}-1/2)f(x)|_{k-1}^{k} + \int_{k-1}^{k} (\{x\}-1/2)f'(x)dx$$
$$= f(k-1)/2 - f(k)/2 + \int_{k-1}^{k} (\{x\}-1/2)f'(x)dx.$$

So

$$\sum_{k=a}^{b} f(k) = \int_{a-1}^{b} f(x)dx + \frac{1}{2}\sum_{k=a}^{b}(f(k-1)) - f(k) + \int_{a}^{b} f'(x)r_1(x)dx.$$

$\cdots$. Claim $\int_1^n \frac{1}{x}(\{x\}-1/2)dx$ converges as $n \to \infty$.

$$\int_{k-1}^{k} \frac{1}{x}(\{x\}-1/2)dx = \int_{k-1}^{k-1/2} \frac{1}{x}(\{x\}-1/2)dx + \int_{k-1/2}^{k} \frac{1}{x}(\{x\}-1/2)dx$$
$$= -\int_{k-1}^{k-1/2} \frac{1}{x}(1/2-\{x\})dx + \int_{k-1/2}^{k} \frac{1}{x}(\{x\}-1/2)dx = O(1/k^2)$$

Exercise. Find an explicit d so that $\int_{k-1}^{k} \frac{1}{x}(\{x\}-1/2)dx < \frac{d}{k^2}$.
Consequently, $\log n! - n\log n + n - \frac{1}{2}\log n \to \log c$ for $\log c = 1 + \int_1^\infty \frac{1}{x}(\{x\}-1/2)dx$. So $\frac{n!}{(\frac{n}{c})^n\sqrt{n}} \to c$.
Hence $n! \sim \left(\frac{n}{e}\right)^n \sqrt{n}c$. What is c? How can we show $c = \sqrt{2\pi}$.
Idea $\sum_{k=-n}^{n} \binom{2n}{n+k} = 2^{2n} = 4^n$. $\sum_{k=-n}^{n} \binom{2n}{n+k} = \binom{2n}{n}\sum_{k=-n}^{n} \frac{\binom{2n}{n+k}}{\binom{2n}{n}}$.
How big is $\sum_{k=-n}^{n} \frac{\binom{2n}{n+k}}{\binom{2n}{n}}$? Note that

$$\frac{\binom{2n}{n+k}}{\binom{2n}{n}} = \frac{\frac{(2n!)}{(n+k)!(n-k)!}}{\frac{(2n)!}{n!n!}} = \frac{n(n-1)\cdots(n-k+1)}{(n+1)(n+1)\cdots(n+k)} = \frac{1(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{k-1}{n})}{(1+\frac{1}{n})(1+\frac{2}{k})\cdots(1+\frac{k}{n})}.$$

We know $\log(1-x) \approx -x$ and $\log(1+x) \approx x$, etc when x is small

Remark. Another look at $n!$ $e^z = \sum_{n\geqslant 0} \frac{z^n}{n!}$. Some elementary complex analysis. We know $i^2 = -1$.

$$e^{i\theta} = \sum_{n\geqslant 0} \frac{(i\theta)^n}{n!} = \sum_{n \text{ even}} \frac{\theta^n}{n!}(-1)^{\frac{n}{2}} + \sum_{n \text{ odd}} i\frac{\theta^n}{n!}(-1)^{\frac{n-1}{2}} = \cos\theta + i\sin\theta.$$

Any complex number $z = a + ib$ with $a, b \in \mathbb{R}$. Then $z = re^{i\theta}$, where $r = |z| = \sqrt{a^2 + b^2}$ and $|z|^2 = (a + ib)(a - ib) = z\bar{z}$.
Complex analysis. Functions of a complex variable z which are well-behaved in a region $\mathcal{D}$ are really well-behaved in $\mathcal{D}$. In particular, if $f : \mathcal{D} \to \mathbb{C}$,

(a) first derivative exists everywhere in a disc $\mathcal{D}$ means second derivative does too, hence infinitely differentiable.

(b) $f(z)$ is well approximated by polynomials.

Integration for complex functions. Let γ be a path in $\mathbb{C}$. $\int_\gamma f(z)dz \approx \sum f(z_i)\Delta z_i$, where Δz_i has a direction as well as magnitude. Now let γ be a line $z_0 \to z_1$, we can parametrize γ by $\gamma = \{z_0 + t(z_1 - z_0) \mid 0 \leqslant t \leqslant 1\}$ and $z(t) = z_0 + t(z_1 - z_0)$ with $\frac{dz}{dz} = z_1 - z_0$, then

$$\int_\gamma z dz = \int_0^1 z(t)\frac{dz}{dt}dt = \int_0^1 (z_0 + t(z_1 - z_0))(z_1 - z_0)dt$$
$$= \left(z_0 t + \frac{t^2}{2}(z_1 - z_0)\right)(z_1 - z_0)\Big|_0^1 = \frac{z_1^2}{2} - \frac{z_0^2}{2}.$$

By induction, we have $\int_\gamma z^k dz = \frac{z_1^{k+1}}{k+1} - \frac{z_0^{k+1}}{k+1}$ for any $k \in \mathbb{Z}^+$. Let $p \in \mathbb{C}[z]$ with antiderivative P. Then $\int_\gamma p(z)dz = P(z_1) - P(z_0)$, where $P'(z) = p(z)$. If $\gamma = \gamma_1 + \gamma_2$, where $\gamma_1 : z_0 \to z_1, \gamma_2 : z_1 \to z_2$ are two lines, then

$$\int_\gamma = \int_{\gamma_1} p(z)dz + \int_{\gamma_2} p(z) = P(z_1) - P(z_0) + P(z_2) - P(z_1) = P(z_2) - P(z_1).$$

So if γ is a nice curve from z_0 to z_1, then approximate γ by a piecewise line, then we have $\int_\gamma p(z) = P(z_1) - P(z_0)$ and is 0 when γ is closed.
Consider z^k when $k \in \mathbb{Z}^{<0}$ and $\gamma = \{z \in \mathbb{C} \mid |z| = 1\}$. Then $\int_{|z|=1} z^k dz = i\int_0^{2\pi} e^{ik\theta}e^{i\theta}d\theta = i\int_0^{2\pi} e^{i(k+1)\theta}d\theta = i\frac{1}{i(k+1)}e^{i(k+1)\theta}\Big|_0^{2\pi} = \begin{cases} 0 & \text{if } k \neq -1 \\ 2\pi i & \text{if } k = -1 \end{cases}$, i.e., $\frac{1}{2\pi i}\int_{|z|=1} z^k dz = \begin{cases} 0 & \text{if } k \neq -1 \\ 1 & \text{if } k = -1 \end{cases}$.
Same for $|z| = r$ for other $r > 0$.

Remark. Let $f(z) = \sum_{n\geqslant 0} a_n z^n$ be a generating function with radius of convergence R. Then inside any open disc $\mathcal{D}_r$ of radius $r < R$ centered at 0, $f(z)$ will be as well behaved as we like. In this case, $a_n = \frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z^{n+1}}dz$, where γ is any closed curve enclosing 0 exactly onece in counterclockwise direction inside $\mathcal{D}_r$.
$e^z = \sum_{n\geqslant 0} \frac{z^n}{n!}$ has infinite radius of convergence. So if $\gamma_n = \{z \mid |z| = n\} = \{ne^{i\theta} \mid 0 \leqslant \theta \leqslant 2\pi\}$, then

$$\frac{1}{n!} = \frac{1}{2\pi i}\int_{\gamma_n} \frac{e^z}{z^{n+1}}dz = \frac{1}{2\pi i}\int_0^{2\pi} \frac{e^{ne^{i\theta}}}{e^{i(n+1)\theta}}ie^{i\theta}d\theta$$
$$= \frac{1}{2\pi}\int_0^{2\pi} e^{n\cos\theta + in\sin\theta}e^{-in\theta}d\theta = \frac{1}{2\pi}\int_0^{2\pi} e^{n\cos\theta}e^{in(\sin\theta - \theta)}d\theta.$$

First suppose θ is small, say $\theta << n^{-\frac{1}{4}}$, then $n\cos\theta \approx n - n\frac{\theta^2}{2!} + n\frac{\theta^4}{4} \approx n - n\frac{\theta^2}{2}$ when $\theta = o(n^{-\frac{1}{4}})$. So in this range, $e^{n\cos\theta} = e^{n - n\frac{\theta^2}{2}}(1 + o(1))$. When $\theta = o(n^{-\frac{1}{3}})$, then $in(\sin\theta - \theta) = in(-\frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \cdots) \to$

0, so $e^{in(\sin\theta-\theta)} \to 1$. Note $n^{-\frac{1}{2}} = o(n^{-\frac{2}{5}})$ and $n^{-\frac{2}{5}} = o(n^{-\frac{1}{3}})$. If we let θ lie between $-n^{-\frac{2}{5}}$ and $n^{-\frac{2}{5}}$, we'll get

(a) outside of this range, contribution to the integral is abot $\frac{e^n}{n^n} \cdot e^{-\frac{n^{-\frac{1}{2}}}{2}}$, which is $\frac{e^n}{n^n} \cdot o(1)$. So the contribution to the integral is negligible.

(b) $n^{-\frac{2}{5}}$, $I = \cdots$. So as $n \to \infty$, $I = \frac{n^n}{e^n}\sqrt{n}2\pi \to \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx = \sqrt{2\pi}$. So $\frac{1}{2\pi} \sim \left(\frac{e}{n}\right)^n \frac{1}{\sqrt{2\pi n}}$. Hence $n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$.

Euler's method: for nonnegative a_j's, $a_n \leqslant \frac{f(x)}{x^n}$ for any $x \in (0, R)$. In this case, $f(x) = e^x$, this was minimized.

6.3 Sequences with restrictions

The set of all finite binary strings.

Definition 6.20 (Binary strings)**.** Binary string:

- Could have length 0, the empty string ϵ.
- Starts with some number of 0's (possibly none).
- k blocks of positive number of 1's followed by positive number of 0's., but k may be 0.
- Ends with some number of 1's (possibly none).

Example 6.21. 0001001110101. Break each binary string into maximal blocks, it is given uniquely as an element of $0^*(11^*00^*)^*1^*$, $0^* = \{\epsilon, 0, 00, 000, 0000, \cdots\}$. $1^* = \{\epsilon, 1, 11, 111, 11111, \cdots\}$. $11^*00^* = \{1^i0^j, i, j \geqslant 1\}$. $(11^*00^*)^*$ = set of concatenations of strings in 11^*00^* (possibly none). S^* = set of concatenations of objects in S (possibly none) provided that no object is attainable in more than one way. So we'll not talk about $\{1, 11\}^*$ (Related to regular languages). Let's use generating functions in conjunction with this....
then clearly, $f_{A\cup B}(x) = f_A(x) + f_B(x)$. Second if A, B are ...... (no long necessarily disjoint) and if we consider $A \times B$, where if $a \in A$ and $b \in B$, weight$((a, b))$ = weight(a) + weight(b), then $f_{A\times B}(x) = f_A(x)f_B(x)$. Similarly for multiple sets $A_1, \ldots, A_k$. Suppose we weight strings by length. String with k places has weight k. e.g., for $0^* = \{\epsilon, 0, 00, 000, \cdots\}$ is $\frac{1}{1-x} = \frac{1}{1-(\text{ogf for }\{0\})}$. gf for $1^* = \frac{1}{1-x}$, for 11^* is $\frac{x}{1-x}$, for $00^* = \frac{x}{1-x}$, for $11^*00^* = \frac{x^2}{1-x}$. $(11^*00^*)^*$ has ogf $1+\frac{x^2}{(1-x)^2}+\left(\frac{x^2}{(1-x)^2}\right)^2+\left(\frac{x^2}{(1-x)^2}\right)^3+\cdots = \frac{1}{1-\frac{x^2}{(1-x)^2}}$. So ogf for all binary strings is $\frac{1}{1-x}\frac{1}{1-\frac{x^2}{(1-x)^2}}\frac{1}{1-x} = \frac{1}{(1-x)^2-x^2} = \frac{1}{1-2x} =$.
How many binary strings of length n have k 1's. Bivariate ogf: x marks length, y marks ... the number of 0's, $0^* \to \frac{1}{1-x}$, $1^* \to \frac{1}{1-xy}$. $00^* \to \frac{x}{1-x}$, $11^* \to \frac{xy}{1-xy}$....
Alt. $[y^k][x^n]\frac{1}{1-x-yx} = [y^k][x^n]\frac{1}{1-(1-y)x} = [y^k](1-y)^n = \binom{n}{k}$. Binary strings with/without $||$. $0^*(100^*)^*\{\epsilon, 1\}$. $\frac{1}{1-x}\frac{1}{1-\frac{x^2}{1-x}}(1+x) = \frac{1+x}{1-x-x^2} = \frac{A}{1-\alpha x} + \frac{B}{1-\beta x}$, where $\alpha = \frac{1+\sqrt{5}}{2}$ and $\beta = \frac{1-\sqrt{5}}{2}$.

Example 6.22. Fibonacci strings. F_n .
Fibonacci strings of length n with k 1's? $\frac{1}{1-x}\frac{1}{1-\frac{x^2y}{1-x}}(1+xy) = \frac{1+xy}{1-x-x^2y}$.

Last time: $0^*(100^*)^* = (\epsilon, 1)$. $\frac{1}{1-x} \cdot \frac{1}{1-\frac{x^2y}{1-x}} \cdot (1+xy)$. Computing $[x^n]$ or $[y^k]$ in $\frac{1}{1-x-x^2y}$. $[x^n]$. we wish to write $\frac{1}{1-x-x^2y}$ and $\frac{A(y)}{1-\alpha(y)x} + \frac{B(y)}{1-\beta(y)x}$. $\alpha(y)+\beta(y) = 1$ and $\alpha\beta = -y$, i.e., $\beta = -\frac{y}{\alpha(y)}$ and $\alpha(y) - \frac{y}{\alpha(y)} = 1$, i.e., $\alpha = \frac{1\pm\sqrt{1+4y}}{2}$. So we are goint to compute $[x^n]$ in an expression involving $\left(\frac{1+\sqrt{1+4y}}{2}\right)^n$. Unfortunately, this will have to wait until we've see Lagrange inversion. Insteed, let's try $[y^k]$ first.

$$[y^k] = \frac{1}{1-x-x^2y} = [y^k]\frac{1}{(1-x)(1-\frac{x^2}{1-x}y)} = \frac{1}{1-x}[y^k]\frac{1}{1-\frac{x^2}{1-x}y} = \frac{1}{1-x}\left(\frac{x^2}{1-x}\right)^k.$$

So

$$[x^ny^k]\frac{1}{1-x-x^2y} = [x^n]\frac{x^{2k}}{(1-x)^{k+1}} = [x^{n-2k}]\frac{1}{(1-x)^{k+1}} = \binom{-(k+1)}{n-2k}(-1)^{n-2k} = (-1)...$$

$$[x^ny^k]\frac{xy}{1-x-x^2y} = [x^{n-1}][y^{(k-1)}]\frac{1}{1-x-x^2y} = \binom{(n-1)-(k-1)}{k-1} = \binom{n-k}{k-1}.$$

So $f_{n,k} = \binom{n-k}{k} + \binom{n-k}{k-1} = \binom{n-k+1}{k}$.

Corollary 6.23. $F_n = \sum_k \binom{n-k+1}{k}$. $F_0 = 1$, $F_1 = 2$, $F_2 = 3$, $F_3 = 5$, $F_4 = 8$, $F_5 = 13$. $\binom{6}{0}+\binom{5}{1}+\binom{4}{2}+\binom{3}{3}$, $13 = 1+5+6+1$. Revisiting partition function. Write $\lambda \mapsto n$ in non-decreasing order instead of non-increasing order. Partition are generated by $1^*2^*3^*4^*5^* = \frac{1}{1-x}\frac{1}{1-x^2}\frac{1}{1-x^3}\cdots = \prod_{i=1}^{\infty}\frac{1}{1-x^i}$. keep track of the number of parts: y marks each part. $\frac{1}{1-xy}\frac{1}{1-x^2y}\frac{1}{1-x^3y}\cdots = \prod_{i=1}^{\infty}\frac{1}{1-x^iy}$.

Remark. Partition into odd number of parts vs Partiton into even number of parts. Partitions into distinct parts.
Empiracally, it appears $\prod_{i=1}^{\infty}(-1)^kx^{\frac{k(2k\pm1)}{2}}$. In fact, $\prod_{i=1}^{\infty} =$.

Remark. How to use bivariate generating functions to count two aspects of partitions. Partitons are generated by $1^*2^*3^*\cdots n^*\cdots$. $\frac{1}{1-x}\frac{1}{1-x^2}\frac{1}{1-x^3}\cdots = (1+x+\underbrace{x^2}_{11}+\cdots)(1+x^2+x^4+\underbrace{x^6}_{222}\cdots)(1+x^3+\underbrace{x^6}_{33}+\cdots)\cdots$. $3+3+2+2+2+1+1 \vdash 14$, a partition of 14 into 7 parts. Consider instead $\frac{1}{1-xy}\frac{1}{1-x^2y}\frac{1}{1-x^ny}\cdots = (1+xy+\underbrace{x^2y^2}_{11}+\cdots)(1+x^2y+x^4y^2+\underbrace{x^6y^3}_{222}+\cdots)(1+x^3y+\underbrace{x^6y^2}_{33}+x^8y^4)\cdots$. $x^{14}y^7$.

We could modify this to count something different for example the number of different part sizes used.

$$(1+xy+x^2y+x^3y)(1+x^2y+x^4y+x^6y+\cdots) = (1+\frac{xy}{1-x})(1+\frac{x^2y}{1-x^2})(1+\frac{x^3y}{1-x^3})\cdots.$$

Theorem 6.24 (Euler's Theorem)**.** *Let $E(n) = \#$ partitions into even number of distinct parts. $O(n) = \#$ partitions into odd number of distinct parts. Then*

$$E(n) - O(n) = \begin{cases} 0 & \text{if } n \text{ is not a pentagonal number} \\ (-1)^k & \text{if } n = \frac{k(3k+1)}{2} \text{ for some } k \in \mathbb{Z} \end{cases}$$

$\prod_{i=1}^{\infty}(1-x^i) = \sum_{k=-\infty}^{\infty}(-1)^kx^{k(3j)}$.

Proof. Give a map which either takes a partition with odd number of parts (all distinct) to one with even number, verse vice.

$$\begin{array}{ccccccc} o & o & o & o & o & o & o \\ o & o & o & o & o & o & \\ o & o & o & o & o & & \\ o & o & o & & & & \\ o & o & & & & & \end{array}$$

Trasform it to

$$\begin{array}{cccccccc} o & o & o & o & o & o & o & o \\ o & o & o & o & o & o & o & \\ o & o & o & o & o & & & \\ o & o & o & & & & & \end{array}$$

If smallest row $\leqslant$ size of diagonal, move it up. If diagonal $<$ size of smallest row, move it down.

Problem can rise

(a) smallest row = size of diagonal, but they intersect.

$$\begin{array}{cccccccc} o & o & \cdots & o & o & o & o & o \\ o & o & \cdots & o & o & o & o & \\ \vdots & & & & & & & \\ o & o & \cdots & o & o & o & & \\ o & o & \cdots & o & o & & & \end{array}$$

On the left, we have a $k \times k$ terms and on the right, we have $\frac{k(k-1)}{2}$ terms. In total, it is $k^2 + \frac{k(k-1)}{2} = \frac{k(3k-1)}{2}$.

(b) smallest row = 1 more than diagonal and they intersect.

$$\begin{array}{cccccc} o & o & o & o & o & o \\ o & o & o & o & o & \\ o & o & o & o & & \end{array}$$

Since number of rows in these exceptional Ferrers graph is k contribution to $E(n) - O(n) = (-1)^k$ if $n = \frac{k(3k+1)}{2}$. □

Remark (8550 Conjection)**.** $\frac{P_n(y)}{1+y}$ is a polynomial with positive integer coefficients.

6.4 Other generating functions for partitions

Remark (Durfee Square)**.** For every $\lambda \vdash n$, there is a unique k so that $\lambda_k \geqslant k$ and $\lambda_{k+1} < k+1$. Then since the generating function for partitions into at most k parts = generating functions for partitions into parts of size $\leqslant k = \prod_{i=1}^{k}(1-x^i)^{-1}$. So $\prod_{i=1}^{\infty}(1-x^i)^{-1} = \sum_{k=0}^{\infty} x^{k^2} \prod_{i=1}^{k}(1-x^i)^{-2}$.

Remark (Self conjugate partitions)**.** Partial row and partial columns plus diagonal element on both. Unfold the books: get a partition into distinct odd parts. ogf. $\prod_{i=0}^{\infty}(1+x^{2i+1})$.

Remark. Euler: $\prod_{i=1}^{\infty} x = \sum_{k=-\infty}^{\infty}(-1)^k x^{k(3k+1)}/2$. $P(x) = \prod_{i=1}^{\infty}(1-x^i)$.

So $P(x)\sum_{k=-\infty}^{\infty}(-1)^k x^{k(3k+1)/2} = 1$. Hence $\sum_{\frac{k(3k+1)}{2}\leqslant n} P_{n-\frac{k(3k+1)}{2}}(x)(-1)^k = 0$ unless $n=0$.

How far does P_n grow? $P_n = [x^n]\prod_{i=1}^{\infty}(1-x^i)^{-1} \leqslant \frac{P(x)}{x^n}$, where $x \in (0,1)$.

Euler Criterion: If $\sum|a_k|$ converges, so does $\prod(1+a_k)$, so $\sum x^i$ converges if $|x|<1$, hence $\prod(1-x^i)^{-1}$ converges.

How big is $P(x)$?

Remark. Can we bound $\frac{P(x)}{x^n}$, where $P(x) = \prod_{i=1}^{\infty}(1-x^i)^{-1}$ and use this to give an upper bound for p_n. Suppose $x \in (0,1)$. $B < \frac{P(x)}{x^n} = \exp\{\log\prod_{i=1}(1-x^i)^{-1} - n\log x\}$. Note that

$$\exp\left\{\log\prod_{i=1}^{\infty}(1-x^i)^{-1}\right\} = \exp\left\{\sum\log(1-x^i)^{-1}\right\} = \exp\left\{-\sum\log(1-x^i)\right\}$$

$$= \exp\left\{\sum_{i=1}^{\infty}\sum_{k=1}^{\infty}\frac{x^{ik}}{k}\right\} = \exp\left\{\sum_{k=1}^{\infty}\frac{1}{k}\sum_{i=0}^{\infty}x^{ik}\right\} = \exp\left\{\sum_{k=1}^{\infty}\frac{1}{k}\frac{x^k}{1-x^k}\right\}.$$

Aside: $\frac{(1-x)x^k}{1-x^k} = \frac{x^k}{1+x+x^2+\cdots+x^{k-1}} < \frac{1}{k}$. So

$$S = \prod_{i=1}^{\infty}(1-x^i)^{-1} = \exp\left\{\sum_{k=1}^{\infty}\frac{1}{k}\frac{1}{1-x}\frac{(1-x)x^k}{1-x^k}\right\} < \exp\left\{\sum_{k=1}^{\infty}\frac{1}{1-x}\cdot\frac{1}{k^2}\right\}.$$

Note $\sum_{k=1}^{\infty}\frac{1}{k^2} = \zeta(2) = \frac{\pi^2}{6}$. So $S < \exp\{\frac{\pi^2}{6(1-x)}\}$. Since we expect (do we? will this be borned out by this?) that as $n\to\infty$, this minimizing value of x approaches 1, we'll parametrize by $x = 1-\epsilon$ and minimize w.r.t. ϵ (Note: any ϵ will give a valid upper bound). $p_n < \exp\{\frac{\pi^2}{6\epsilon} - n\log(1-\epsilon)\}$. Since $\exp(\cdot)$ is increasing, we can minimize $\frac{\pi^2}{6\epsilon} - n\log(1-\epsilon)$. $\frac{d}{d\epsilon}(\frac{\pi^2}{6\epsilon} - n\log(1-\epsilon)) = -\frac{\pi^2}{6\epsilon^2} + \frac{n}{1-\epsilon}$. So $\frac{\pi^2}{6\epsilon^2} = \frac{n}{1-\epsilon}$. So $\pi^2 - \pi^2\epsilon = 6\epsilon^2 n$. Then

$$\epsilon = \frac{-\pi^2 + \sqrt{\pi^4 + 24\pi^2 n}}{12n} = -\frac{\pi^2}{12n} + \frac{\pi}{12n}\sqrt{\pi^2+24n} \approx \pi\sqrt{\frac{1}{6n}}.$$

If we now playing $\epsilon = \pi\sqrt{\frac{1}{6n}}$ into $p_n < \exp\{\frac{\pi^2}{6\epsilon} - n\log(1-\epsilon)\}$, we get

$$p_n < \exp\{\pi\sqrt{\frac{n}{6}} - n\log(1-\pi\sqrt{\frac{1}{6n}})\} = \exp\{\pi\sqrt{\frac{n}{6}} + \frac{n\pi}{\sqrt{6n}} + o(\frac{1}{\sqrt{n}})\} \approx \exp\{2\pi\sqrt{\frac{n}{6}}\} = \exp(\pi\sqrt{\frac{2n}{3}}).$$

Remark. $(\frac{1}{10^5}\sum_{n=-\infty}^{\infty} e^{-\frac{n^2}{10^{10}}})^2 \neq \pi$, but it is correct to 42 billion decimal digits.
Jacobi theorem and the inversion theorem.

Remark. How big is the number of patitions of n into distinct parts? # partitions of n into distinct parts = # partitions of n into odd parts.

If we take a partition of n into odd parts, and remove a 1 from each part, we get a partition of something close to n into even parts, corresponds to a partition of something close to $\frac{n}{2}$ into any parts. Heuristically, expect $q_n \cong p_{n/2}$, $p_{n/2} \cong e^{\pi\sqrt{\frac{2n}{3\cdot 2}}} = e^{\pi\frac{n}{3}}$. This turns out to be correct as an order of magnitude.

Next approaches

(a) Try $Q(x) = \prod_{i=1}^{\infty}(1+x^i)$ bound $\frac{Q(x)}{x^n}$, etc.

(b) Compute $\prod_{i=0}^{\infty}(1-x^{2i+1})^{-1}$ and $\prod_{i=1}^{\infty}(1-x^{2i+1})^{-1}$ to $p(x^2)\prod_{i=1}^{\infty}(1-x^{2i})^{-1}$.

What proportion of partitions has a part of size 1? How big is $\frac{p_{n-1}}{p}$. Asymtotically

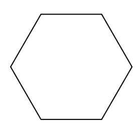
$$\frac{\frac{c}{n-1}\exp\left(\pi\sqrt{\frac{2(n-1)}{3}}\right)}{\frac{c}{n}\exp\left(\pi\sqrt{\frac{2n}{3}}\right)} \approx \exp\left(\pi\sqrt{\frac{2}{3}}(\sqrt{n-1}-\sqrt{n})\right) \approx \exp\left(\pi\sqrt{\frac{2}{3}}\frac{-1}{\sqrt{n-1}+\sqrt{n}}\right)$$
$$\approx \exp\left(-\pi\sqrt{\frac{2}{3}}\frac{2}{\sqrt{n}}\right) \approx 1-\pi\sqrt{\frac{2}{3}}\frac{2}{\sqrt{n}}.$$

Other questions we could ask: can we get this estimate in any other way? What is the distribution of the number of 1's? What is the distribution of the number of 2's, $\ldots$, d's? What is the distribution of the size of the largest part? number of parts?

Remark. Partitions don't have labels. Labels make things fun! We'll see some examples of labelled objects, and ways of combining them.

Our labels will typically be (always be?) $\{1,\ldots,n\}$.

Example 6.25. Directed cycles.



(1 3 4 2 6 7) in cycle form.

How many labelled cycles of length n, labelled with $\{1,\ldots,n\}$ are there? $(n-1)! : c_n$. How many pairs of cycles are there on a total of n vectices, labelled together by $\{1,\cdot,n\}$? There are no cycle on 0 vertices.

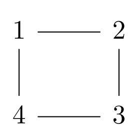
$\frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k}(k-1)!$ Since the c_k's grow quickly, define $c(x) = \sum_{n=0}^{\infty} c_n\frac{x^n}{n!} = \sum_{n=1}^{\infty}\frac{x^n}{n} = -\log(1-x)$. Let $b_n = \#$ pairs of cycles labelled with $\{1,\ldots,n\}$.

Remark. Suppose we have two sets of labelled objects A_1, A_2 so that objects of weights n have exactly n lables $1,\ldots,n$. (e.g., graphs on vertices $\{1,\ldots,n\}$ trees with n edges labelled $\{1,\ldots,n\}$. set partitions of $\{1,\ldots,n\}$).

6.5 Permutation

Definition 6.26. Construct a set $S = A_1\times A_2$ of all parts of objects (α_1,α_2) with $\alpha_1\in A_1$ and $\alpha_2\in A_2$ and label set for (α_1,α_2) is $\{1,2,\ldots,\text{weights}(\alpha_1)+\text{weights}(\alpha_2)\}$ and then labels partition this set.

Example 6.27. A_1 has



A_2 has (1 2 3). (squre,( )) has total weights 7. Pick a 4-set of labels from $\{1,\ldots,7\}$ to label the square and the complement to ( ).

Remark. The number of labelled objects of size n thus created is $\sum_{k=0}^{n}\binom{n}{k}f_k g_{n-k}$, where $f_k =$ # of labelled objects of weight, $g_m =$ # of labelled objects of weight m in A_2. If $f(x)=\sum_{k=0}^{\infty} f_k\frac{x^k}{k!}$ is the edf for A_1........
If $A_1,\ldots,A_m$ are sets of labelled objects with egf $f_1(x),\ldots,f_m(x)$ respectively, then the number of ordered k-tuples of labelled $(\alpha_1,\ldots,\alpha_k)$ with $\alpha_i\in A_i$ has egf $f_1(x)\cdots f_k(x)$. Some sets of combinatorial objects have a natural subset of objects which could be considered "connected". e.g., finite connected graphs as a subset of the set of all finite graphs. finite trees as a subset of all finite forests sets as a subset of set partitions. cycles as a subset of set of partitions.

Suppose our class A is a set of labelled connected objects with egf $f(x)$. Then the egf for $A\times\cdots\times A$ with all labellings is $f(x)^k$. The set of all sets of k objects chosen from A and labelled thus has egf $\frac{1}{k!}f(x)^k$. So as 2 tuples $((1\,2\,4),(3\,5))$ and $((3\,5),(1\,2\,4))$ are distinct, but $\{(1\,2\,4),(3\,5)\}$ is the same as $\{(3,\ 5),(1\ 2\ 4)\}$.

Convention 6.28 (Important!)**.** # of connected objects of weight 0 is 0, that is $f(0)=0$. Then the sets of all labelled objects formed from relabelling sets of connected objects has edf $\sum_{k=0}^{\infty}\frac{f(x)^k}{k!}=\exp\{f(x)\}$.

Example 6.29. Set of all permutations: # of cycles on $\{1,\ldots,n\}$ is $(n-1)!$ if $n\geqslant 1$. Cycles have egf $\sum_{n\geqslant 1}(n-1)!\frac{x^n}{n!}=-\log(1-x)$. So egf for all permutations is $\exp\{-\log(1-x)\}=\frac{1}{1-x}$ as expected.

Remark. The set of non-fixed point cycles has egf $-\log(1-x)-x$. So the egf for all permutations without fixed point is $\exp\{-\log(1-x)-x\}=e^{-x}\frac{1}{1-x}$.

$$\left[\frac{x^n}{n!}\right]\frac{e^{-x}}{1-x}=n![x^n]\frac{1}{1-x}e^{-x}=n![x^n](1+x+\cdots+x^k+\cdots)(1-x+\frac{x^2}{2!}-\cdots+(-1)^k\frac{x^k}{k!}+\cdots)$$
$$=n!\sum_{k=0}^{n}1\frac{(-1)^k}{k!}=n!(e^{-1}-\sum_{k=n+1}^{\infty}\frac{(-1)^k}{k!})=n!e^{-1}-n!\sum_{k=n+1}^{\infty}\frac{(-1)^k}{k!}.$$

$\left|\sum_{k=n+1}^{\infty}\frac{(-1)^k}{k!}\right|.$

Remark. We saw that #permutations of $\{1,\ldots,n\}$ without a fixed point is the closest integer to $\frac{n!}{e}$ (also, if $n\geqslant 2$, $\lfloor\frac{n!}{e}+\frac{1}{3}\rfloor$, check this).

One way to interpret this is the following: random permution has probability $\approx\frac{1}{e}$ of having no fixed points. How about the probability that a permutation has one fixed point? two? seven?

Precisely one fixed point. Pick the fixed point in n ways. Pick a derangement on remaining set in $\approx\frac{(n-1)!}{e}$ ways. So $\approx\frac{n(n-1)!}{e}$ permutations with one fixed point.

Remark (Exercise)**.** (say $n\leqslant 8$), use sage to count # with $0,1\cdots,8$ fixed points.

How far apart can $[\frac{n!}{e}]$ and $n[\frac{(n-1)!}{e}]$ be?. Here [] is a temporary nearest integer-notation.

Clearly, fix k, # with exactly k fixed points is $\binom{n}{k}[\frac{(n-k)!}{e}]$ when $n\neq k$, 1, otherwise. So $n!=\sum_{k=0}^{n}\binom{n}{k}[\frac{(n-k)!}{e}]+1$. Is this reasonable? Divide by $n!$ and approximate $1=\sum\frac{1}{k!}\times^{\text{closest to}}\frac{1}{e}$, reasonable.

How many cycles does a ramdonn permutation have? Return to "all = exp{connected}".Mixed generating function, exponential for # of labels, ordinary for # cycles. egf for connected permuations is $-y\log(1-x)$. Then # permutations of $\{1,\ldots,n\}$ with k-cycles

$$[y^k\frac{x^n}{n!}]\exp(-y\log(1-x)) = [y^k\frac{x^n}{n!}](1-x)^{-y} = [y^k]n!\binom{-y}{n}(-1)^n = [y^k]y(y+1)\cdots(y+n-1).$$

Unsigned Stirling number of the first kind.

Let $f(n,k) = [y^k\frac{x^n}{n!}]\exp\{-y\log(1-x)\}$.

$$\begin{aligned}[y]y(y+1)(y+2)\cdots(y+n-1) &= [1](y+1)(y+2)\cdots(y+(n-1))\\ &= [1](y+1)(y+1)(y+2)\cdots(y+n-1)\\ &= 1\cdot 2\cdot 3\cdot(n-1) = (n-1)!.\end{aligned}$$

Probability of a random permutation has 1 cycle is $\frac{(n-1)!}{n!} = \frac{1}{n}$.

$$f(n,2) = [y^2]y(y+1)(y+2)\cdots(y+n-1) = [y](y+1)(y+2)\cdots(y+n-1).$$

Can we find recurrence for $f(n,k)$? $f(n,k) = f(n-1,k-1)+(n-1)f(n-1,k)$, $f(n-1,k) = f(n,k-1)+nf(n,k)$. Expected # cycles in a random permutation is

$$\frac{1}{n!}\sum_k k[y^k]y(y+1)\cdots(y+n-1).$$

$$\frac{1}{n!}\frac{d}{dx}y(y+1)\cdots(y+n-1)|_{y=1} = \frac{1}{n!}y(y+1)\cdots(y+n-1)(\frac{1}{y}+\frac{1}{y}+\cdots+).$$

Euler Mach,

Remark. # k-cycles in a permutation. Let y count k-cycles. The egf for cycles is

$$-\log(1-x)-\frac{x^k}{k}+y\frac{x^k}{k} = -\log(1-x)+(y-1)\frac{x^k}{k}.$$

So the egf for all permutations with y combining k-cycles is $\exp\{-\log(1-x)+(y-1)\frac{x^k}{k}\} = \frac{1}{1-x}\exp\{(y-1)\frac{x^k}{k}\}$. $f(n,k) = [\frac{x^k}{n!}y^k]\frac{1}{1-x}\exp\{(y-1)\frac{x^k}{k}\}$. To compute the expected # of cycles of length k, $f(n,j) =$ # permutations on $\{1,\ldots,n\}$ having exactly j k-cycles. The expected number is

$$\sum_{j=0}\frac{jf(n,j)}{n!} = \frac{1}{n!}\sum_{j=0}[\frac{x^n}{n!}y^j]\frac{d}{dy}\frac{1}{1-x}e^{-(y-1)\frac{x^k}{k}} = \frac{1}{n!}[\frac{x^n}{n}]\frac{x^k}{k}\frac{e^0}{1-x} = \frac{1}{k}$$

provided $1\leqslant k\leqslant n$.

So expected number of k-cycles in a random permutation is exactly $\frac{1}{k}$!

Remark. Revisiting set partitions. What does a connected set partiton of $\{1,2,3,4,5\}$ look like? $\{1,2,3,4,5\}$? The number of connected set partitions of $\{1,\ldots,n\}$ is 1 for $n\geqslant 1$. So egf for connected set partitions is $\sum_{n\geqslant 1}\frac{x^n}{n!} = \exp(x)-1$. So the egf for all set partitions is $e^{e^x}-1$. How about counting parts? If y marks the number of parts, then $f(n,k) =$ # set partitions with k parts.

$f(n,k) = [\frac{x^n}{n!}y^k]\exp\{y(e^x-1)\}$. So to try to calculate the expected number of parts, we'd differentiate w.r.t. y as before. But now we run into difficulties since it is harder to compute $[\frac{x^n}{n!}]e^{e^x-1}$, then in $\frac{1}{1-x}$.

Instead. How many set partitions of $\{1,\dots,n\}$ have exactly k parts?

$$[\frac{x^n}{n!}y^k]e^{y(e^x-1)} = [\frac{x^n}{n!}]\frac{(e^x-1)^k}{k!} = [\frac{x^n}{n!}]\frac{1}{k!}\sum_{j=0}^{k} e^{jx}\binom{k}{j}(-1)^{k-j}$$
$$= \frac{1}{k!}\sum_{j=0}^{k}\binom{k}{j}(-1)^{k-j}[\frac{x^n}{n!}]e^{jx} = \frac{1}{k!}\sum_{j=0}^{k}\binom{k}{j}j^n(-1)^{k-j}.$$

Remark (Exercise)**.** Can you recover the recurrence for $S(n,k)$ from week (2?). From $\exp\{y(e^x-1)\}$?

Corollary 6.30. The number of onto functions from $\{1,\dots,n\}$ to $\{1,\dots,k\}$ is $k!$.

$$\frac{1}{k!}\sum_{j=0}^{k}\binom{k}{j}(-1)^{k-j}j^n = \sum_{j=0}^{k}\binom{k}{j}(-1)^{k-j}j^n = k^n - \binom{k}{1}(k-1)^n + \binom{k}{2}(k-2)^n - \cdots.$$

When k is small, # onto functions is thus $k=1$: $1^n - k\cdot 0^n = 1$, $k=2$: $2^n - \binom{2}{1}1^n = 2^n-2$, $k=3$: $3^n - \binom{3}{1}\cdot 2^n + \binom{3}{2}1^n = 3^n - 3\cdot 2^n + 3$.

It is similar to permutations. We could try to count # of parts of size k instead?

6.6 Labelled trees and forest

6.6.1 Labelled notes

Definition 6.31. Tree is a connected graph without cycles. # of trees on n vertices: $n=1$: $1, 1^{-1}$, 1,

$n=2$: $1, 2^0$, $1-2$,

$n=3$: $3, 3^1$, $\cdot - - - - \cdot - - - \cdot$: 3 choices for lable, other two lables are forces.

$n=4$: $16, 4^2$.

6.7 Pm̈fer sequences

Remark. Since the number of labelled trees on n vertices is n^{n-2}, we look for a bijection from the set of trees on $\{1,\dots,n\}$ to the set of sequences of length $n-2$ having entries $1,\dots,n$ with repetition allowed. (graph).

Observe if a finite graph has minimum degree 2, then it has a cycle. Pick x_1. Given $x_1,\dots,x_k$, where $(x_1,x_2),\cdots,(x_{k-1},x_k)$ are edges of G, x_k has another vertex adjacent to it. If it is in $x_1,\dots,x_{k-2}$, it will create a cycle. Otherwise, get x_{k+1}, k is bounded by # vertices. So eventually we must have created cycle.

Corollary 6.32. If T is a tree, it has a vertex of degree $\leqslant 1$. If T has n vertices, $n \geqslant 2$ degree. In fact, easy to see if $n \geqslant 2$, T has $\geqslant 2$ vertices of degree 1.

To create a Pm̈fer sequence for T, create an empty list, find the leaf of T with smallest label, delete it, and append the lable of its neighbor to the list. (graph) $[2,1,8,5,10,8,8,1,10,2]$. Clearly, this process creates a list on $n-2$ numbers from $\{1,\ldots,n\}$. Distinct trees clearly give distinct lists. Exercise. Write a careful proof.

To remove the tree from the list: note n (in this case 12) is never the lowest label of a leaf since T has $\geqslant 2$ leaves. Given the sequence $[2,1,8,5,10,8,8,1,10,2]$. Any number not in the sequence was a leaf in T. Numbers are not in $[2,1,8,5,10,8,8,1,10,2]$ are $\{3,4,6,7,9,11,12\}$. Least of these is 3. So $(3,2)$ is in T. This leaves out list as $[1,8,5,10,8,8,1,10,2]$ and $\{4,6,7,9,11,12\}$ as leaves to consider. $(4,1)$ is an edge. $[8,5,10,8,8,1,10,2]$ and $\{6,7,9,11,12\}$. $(6,8)$ is an edge $[5,10,8,8,1,10,2]$ and $\{7,9,11,12\}$. $[10,8,8,1,10,2]$ and $\{5,9,11,12\}$. (graph).

2,4,6,7,5,9,11,8,

$[8,8,1,10,2]$ and $\{9,11,12\}$. $[8,1,10,2]$ and $\{11,12\}$. $[1,10,2]$ and $\{8,12\}$. $[10,2]$ and $\{12\}$. $[2]$ and $\{10,12\}$. $[]$ and $\{2,12\}$. This process works in general.

This gives us a way to uniformly generate randomly labelled trees. Furthermore, it gives us extra information about.......

6.8 Connection between labelled trees and forests

. Functional equation for forests. Why this has a unique solution? Next time: Lagrange Inverse Formula.

Aim: show $r(x) = \sum_{n\geqslant 0} r_n \frac{x^n}{n!}$ and $t(x) = \sum_{n\geqslant 0} t_n \frac{x^n}{n!}$. $r(x) = \exp(t(x))$, but this is not particularly helpful as the sage computation shows. Switch to $t_n = \#$ rootted labelled trees distinguish one of the n vertices as a root $= n^{n-2}\cdot n = n^{n-1}$. $r = \#$ rooted labelled forests: each subtree get a root.

Rooted lablled trees on $n+1$ vertices give rise to rooted labelled forests on n vertices, erase root of tree: make adjoint vertices roots of the forest. Relabel with $\{1,\ldots,n\}$, $n+1$ rooted labelled trees on $(n+1)$ vertices give rise to same rooted labelled forest. So $t_{n+1} = (n+1)r_n$. Thus, $r_n = \frac{(n+1)^n}{n+1} = (n+1)^{n-1}$. $r(x) = \sum_{n=1}^{\infty} (n+1)^{n-1} \frac{x^n}{n!}$ and $t(x) = \sum_{n\geqslant 1} n^{n+1} \frac{x^n}{n!}$. So

$$xr(x) = \sum_{n\geqslant 0} (n+1)^{n-1} \frac{x^{n+1}}{n!} = \sum_{n\geqslant 0} (n+1)^n \frac{x^{n+1}}{(n+1)!} = \sum_{n\geqslant 1} n^{n-1} \frac{x^n}{n!} = t(x),$$

i.e., $t(x) = xr(x) = x\exp(t(x))$.

So we obtain a functional equation for $t(x)$, $t(x) = x\exp(t(x))$. We can hence write $x(t)$ via $x(t) = t\exp(-t)$. We can hence write $x(t)$ via $x(t) = t\exp(-t)$ enabling us to give a power series $x(t)$ for x in terms of t: $[t^n]x(t) = [t^{n-1}]e^{-t} = \frac{(-1)^{n-1}}{(n-1)!}, n \geqslant 1$. This has positive (indeed ∞ radius of convergence). Why does this ensure that we can find coefficients $t_1, t_2, \ldots, t_n, \cdots$ so that $t(x) = \sum_{n\geqslant 1} t_n \frac{x^n}{n!}$?

Remark (Lagrange Inversion)**.** Given a relationship like $x(t) = \frac{t}{p(t)}$ find coefficients t_n so that $t(x) = \sum t_n \frac{x^n}{n!}$ satifies $t(x) = x\varphi(t(x))$.... We'll need certain conditions and furthermore, ...

Recall 6.33. Labelled rooted trees, then $t(x) = xe^{t(x)}$, which is easy to solve for $x : x = te^{-t} = \sum_{n=1}^{\infty} t^{n+1} \frac{(-1)^n}{n!}$.

What we want to do is to invest this power series $x(t)$ to find a compositional inverse $t(x)$.

Little results (but useful). We've seen making liberal use of $[x^n]$. However, this is not the fundamental operator to use. The best operator to use is $[x^{-1}]$. We can then recover $[x^n]f(x)$

by computing $[x^{-1}]x^{-(n+1)}f(x)$. This require us to work in the realm of Laurent series instead of power series, series of the form $\sum_{k=-m}^{\infty} a_k x^k$. Why is $[x^{-1}]$ different from all other $[x^n]$? Because $\frac{1}{x}$ is different in nature from x^n for any $n \neq -1$. $\frac{1}{x}$ is not the derivative of x^{-n} or any Lauranet series. So if $f \in \mathbb{R}(x)$ or $\mathbb{C}[x]$, $[x^{-1}]f'(x) = 0$. If $f(x)g(x) \in \mathbb{R}[[x]]$, then $f(x)g(x) \in \mathbb{R}[x]$, $[x^{-1}](fg)' = 0$ and $[x^{-1}]fg' = [x^{-1}]fg$.

Following Wilf's notation, (completely incompatible with last class's notation). We'll have a function equation of the form $u = t\varphi(u)$. (in our earlier discussion, we had $t = e^{t(x)}$.) So t plays the role of x and u plays the role of t, φ plays the role of exp().

Theorem 6.34. *Let $f(u)$ and $\varphi(u)$ be formal power series in u with $\varphi(0) = 1$. Then there is a unique formal power series $u(t)$ satisfying $u(t) = t\varphi(u(t))$, i.e., $\varphi = \frac{u}{t}$ and furthermore if we compute $f(u(t))$ as a power series in t, about $t = 0$, satisfies $[t^n]f(u(t)) = \frac{1}{n}[u^{n-1}](f'(u)\varphi(u)^n)$.*

Example 6.35. Before we prove this, let's apply it. If $\varphi() = \exp()$, $f(u) = u$, then $[t^n]f(u) = \frac{1}{n}[u^{n-1}](1\exp(un)) = \frac{1}{n}\frac{n^{n-1}}{(n-1)!} = \frac{n^{n-1}}{n!}$. So # of labelled rooted trees is n^{n-1}. It is sufficient to prove for f a polynomial in u and for φ a in u....

$$= \frac{1}{2\pi i}\int t^{-n} f'(ut)u'(t)dt = \frac{1}{2\pi i}\int t^{-n}\frac{d}{dt}f(u(t))dt = [t^{-1}]t^{-n}\frac{d}{dt}f(u(t)) = [t^{-1}]t^{-(n+1)}t\frac{d}{dt}f(u(t)).$$

Since $[t^{-1}]g(t)h'(t) = -[t^{-1}]g'(t)h(t)$, we have $[t^{-1}]t^{-n}f' = [t^n]$.

Remark (Lagrange Version)**.** Let $f(u), \varphi(u)$ be power series in u with $\varphi(u) = 1$. Then we can find a unique power series $u(t)$ so that $u(t) = t\varphi(u(t))$ and furthermore, $[t^n]f(u(t)) = \frac{1}{n}[u^{n-1}]f'(u)\varphi(u)^n$.

Theorem 6.36 (Galois, Abeletc)**.** *"You can't solve some quintics".*

(a) Consider the degree 5: $x^5 - x + z = 0$, where z is a parameter we can vary. Clearly for some z, we can solve this, e.g., $z = 0$. It is in fact known that this has solutions in radicals if and only if it has an integer solution, or if $z = \pm 15, \pm 22440$ or ± 2759640. We'll see how to solve $x^p - x + z = 0$ (not in radical). Rewrite this as $z = x - x^p$. Want to get a formula for $x = x(z)$ so that this is satisfied. Note $x = \frac{z}{1-x^{p-1}} = z\varphi(x)$, where $\varphi(x) = \frac{1}{1-x^{p-1}}$. So z plays the role of t in LIF and x plays the role of u. Hence $[z]x(z) = \frac{1}{n}[x^{n-1}]\varphi(x)^n$. So we need to consider $\frac{1}{n}[x^{n-1}](\frac{1}{1-x^{p-1}})^n$. Let $y = x^{p-1}$. Then

$$(1-x^{p-1})^{-n} = (1-y)^{-n} = 1+ny+\frac{n(n+1)}{2}y^2+\cdots = \sum_{k=0}^{\infty}\binom{n+k-1}{k}y^k = \sum_{k=0}^{\infty}\binom{n+k-1}{k}x^{k(p-1)}.$$

So want $k(p-1) = n-1$. If $n-1 = k(p-1)$, then $n-1+k = kp$, so we get $\frac{1}{k(p-1)+1}\binom{kp}{k}$ as coefficient. So we get $x(z) = \sum_{k=0}^{\infty}\frac{1}{(p-1)k+1}\binom{kp}{k}z^{k(p-1)+1}$, which converges inside. $|z| \leqslant (p-1)p^{-p(p-1)}$. If $p = 5$, we get $4, 5^{-\frac{4}{5}} = \frac{4}{5^{14}}$. In particular, if you are interested in solving problems involving interest rates. e.g., converges these technique can often be helpful.

Remark. Story: Minutes before the drawing for a lottery, you buy a ticket. You know current total # tickest T (include you) for this drawing, the (constant for all tickets) probability p of a particular ticket winning, and the current prize value V. When a ticket is bought, it is assigned uniformly and independent from all tickets. Want to know the expected value of the ticket you but. There are $n = \frac{1}{p}$ possible tickets, each equally likely to win. Only one of these can win (though multiply copies may have been sold). If w copies of the winning ticket are sold, then the payoff to each winning ticket is $\frac{V}{w}$. Let X be how much you win.

$$\begin{aligned} E[X] &= V \cdot P(\text{you win, } w=1) + \frac{V}{2}P(\text{you win, } w=2) + \cdots + \frac{V}{T}P(\text{you win, } w=T) \\ &= Vp\sum_{k=0}^{T-1}\frac{1}{k+1}\binom{T-1}{k}p^k(1-p)^{T-1-k}. \end{aligned}$$

3 approaches to get the final answer.

(a) "correct approach". introduce an x^{k+1} in the k^{th} term, differentiate w.r.t. x, sum and integrate.

(b) Be smart, get lucky: $\frac{1}{k+1}\binom{T-1}{k} = \frac{(T-1)!}{(k+1)k!(T-1-k)!} = \frac{(T-1)!}{(k+1)!(T-(k+1)!} = \frac{1}{T}\binom{T}{k+1}$.

$$E[X] = Vp\sum_{k=0}^{T-1}\frac{1}{T}\binom{T}{k+1}p^k(1-p)^{T-1-k} = \frac{V}{T}\sum_{k=0}^{T-1}\binom{T}{k+1}p^{k+1}(1-p)^{T-(k+1)} = \frac{V}{T}(1-(1-p)^T).$$

(c) Your ticket has the same expected value as everyone elses. $\sum_{\text{all ticket sold}} E[X_t] = V(1-(1-p)^T) = V \times$ probability tiket at least one wins.

As p is very small, $(1-p)^T \approx e^{-p^T}$.

Theorem 6.37 (Ramsey Theorem)**.** $R(3,3) = 6$ *meaning that if we color the edges of* K_6 *with two colors, yellow and brown, either there will be a yellow triangle or a brown triangle, but not true for* K_5*.*

Take a coloring of K_6*. Pick a vertex of* K_6*. Five edges implies* $\geqslant 3$ *yellow edges or* $\geqslant 3$ *brown edges. Suppose yellow. (graph). Similar proof shows that* $R(m,k)$ *is finite, there is at least* n *so that 2 coloring of* K_n *has either a yellow* K_m *or a brown* K_k*.*

The best lower bounds are much smaller than the best upper bounds.

Remark (The probabilistic methodin combinatorics)**.** Idea: generate combinatorial objects at random from some distribution and investigate their properties.

How to obtain information about $R(m,m)$? Take K_n with equal probability and independently color each edge y or b. For a given sd $S \subseteq V$ of m vertices, $P(S \text{ is monochromatic } K_m) = 2 \cdot 2^{-\binom{m}{2}}$. Let $X_S = 1$ of S is monochromatic and 0 otherwise. Let $X = \sum_{|S|=m, S\subseteq V} X_S$. Then $E[X] = \binom{n}{m}2^{1-\binom{m}{2}}$. If $E[X] \leqslant 1$, then since X is non-negative integer valued, there must exist a coloring of K_n with 0 monochromatic K_m's. $E[X] = \sum_k kP_r(X=k)$. How big is m if $\binom{n}{m} = 2^{\binom{m}{2}-1}$. Can $m = cn$? (c $fixed$, $n \to \infty$). LHS $< n^m = n^{cn}$. Exponent $cn\log n$. RHS $= 2^{\frac{c^2n^2}{2}-1}$. Exponent Cn^2. How about n^c some fixed $c \in (0,1)$. $\binom{n}{n^c} < n^{n^c} = e^{n^c \log n}$. $2^{\binom{m}{2}} \approx 2^{\frac{n^{2c}}{2}}$. So $m = o(n^c)$ for all $c > 0$. The nice thing about this is that if $m = o(n^{1/2})$, then $\binom{n}{m} = \frac{n(n-1)\cdots(n-m)+1}{m!} \approx \frac{n^m}{m!}$.

$$\begin{aligned} n(n-1)\cdots(n-m+1) &= n^m(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{m-1}{m}) \\ &\approx n^m\exp\{-\frac{1}{n}-\frac{2}{n}-\cdots-\frac{m-1}{n}\} = n^m\exp\{-\frac{\binom{m}{2}}{n}\}. \end{aligned}$$

Now want $\frac{n^m}{m!} \approx 2^{\binom{m}{2}-1}$. Forget (-1), $\frac{n^m}{(\frac{m}{e})^m\sqrt{2\pi m}} = 2^{m\frac{m-1}{2}}$. $\frac{ne}{m} = 2^{\frac{m-1}{2}}(\sqrt{2\pi m}^{\frac{1}{m}}) \approx 2^{\frac{m-1}{2}}$, i.e., $n = \frac{m2^{\frac{m-1}{2}}}{e}$. $\log_2 n = \frac{m}{2} + \log_2 m + \log c$. First approximation $m = 2\log_2 n$. Better $m = 2\log_2 n - \log_2 m - \log c$. $\log_2 m < \log_2(\log_2 n) + \log 2$. So $m > 2\log_2 n - \log_2\log_2 n - \log c$.

Definition 6.38. A set $S \subset \mathbb{N}$ is sum free if there is no $x, y, z \in S$ with $x + y = z$.

Theorem 6.39. *If $|T| < \infty$, is a non-empty subset of $\mathbb{N}$, then there exists $S \subseteq T$ with $|S| > \frac{|T|}{3}$ so that S is a sum free.*

Proof. □

Remark. If x is small, say $0 < x < \frac{1}{2}$, then

$$\begin{aligned}-\log(1-x) &= x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots < x + \frac{x^2}{2} + \frac{1}{3}(x^3 + x^4 + x^5 + \cdots) = x + \frac{x^2}{2} + \frac{x^3}{3(1-x)} \\ &< x + \ldots.\end{aligned}$$

Remark. How big is $\binom{n}{k}$ when n is large? How big is $\sum_{k \leqslant l}\binom{n}{k}$.

(a) Case 1: k is not big, say $k = o(n^{1/2})$. $n^k \exp(*) < n(n-1)\cdots(n-k+1) < \exp()$. So provided $\frac{k^3}{n^2} \to 0$, i.e., $k = o(n^{2/3})$. $n(n-1)\cdots(n-k+1) = n^k \exp\left(-\frac{\binom{k}{2}}{n}\right)(1+o(1))$. This of course explains the Birthday Paradov. How about when k is on the same order of magnitude as n? Fix $\alpha n(0,1)$, $k = \alpha n$ (more precisely, $k = k_n$, $\frac{k_n}{n} \to \infty$ as $n \to \infty$). $\binom{n}{k} = \frac{n!}{(k!)((n-k)!)} = \frac{n!}{(\alpha n)!((1-\alpha)n)!}$. By Stirling again, $\binom{n}{k} \sim (\frac{n}{e})^n\sqrt{2\pi n}$. ...

Special case $\alpha = \frac{1}{2}$. $\frac{2^{\frac{1}{2}n}2^{\frac{1}{2}n}}{\sqrt{\pi\frac{1}{2}n}} = \sqrt{\frac{2}{\pi n}}2^n$. Recall...

If we fix $p \in (0,1)$ and look instead at $1 = \sum_{k=0}^n \binom{n}{k}p^k(1-p)^{n-k}$. We could ask where is the summand maximized? At the maximum, we should have something like $\binom{n}{k}p^k(1-p)^{n-k} \approx \binom{n}{k+1}p^{k+1}(1-p)^{n-k-1}$. $\frac{1}{n-k}(1-p) \approx \frac{1}{k}p$. $\frac{1}{n-k} = p(\frac{1}{k} + \frac{1}{n-k}) = \frac{np}{k(n-k)}$ or $p = \frac{k}{n}$ or $k = np$. Then more careful, estimate show quadratic decay away from pn.

Remark (Partial sums of binomial coefficients, n even)**.** Fix $l < \frac{n}{2}$ for $\sum_{k\leqslant l}\binom{n}{k}$. If $l > \frac{n}{2}$, $\sum_{k\leqslant l}\binom{n}{k} = 2^n - \sum_{k\leqslant n-l-1}\binom{n}{k}$. If $l = \frac{n}{2}$, $\sum_{k\leqslant\frac{n}{2}}\binom{n}{k} = 2^{n-1} - \frac{1}{2}\binom{n}{n/2}$.

Since $l < \frac{n}{2}$,

$$\sum_{k\leqslant l}\binom{n}{k} = \binom{n}{l} + \binom{n}{l-1} + \binom{n}{l-2} = \binom{n}{l}\left(\frac{\binom{n}{l-1}}{\binom{n}{l}} + \frac{\binom{n}{l-2}}{\binom{n}{l}}\right).$$

6.9 Open problems

....

Heuristically speaking: What proposition of integer up to N are 3 good? If $N < 3^k$, then 2^k integer are 3-good. So "Pr" n is 3-good $= (\frac{2}{3})^k = (\frac{2}{3})^{\log 3N} = e^{\log(\frac{2}{3})\frac{\log N}{\log 3}}$.

Remark. Representating #'s in binary corresponds $\frac{1}{1-x} = (1+x)(1+x^2)(1+x^4)(1+x^8)(1+x^{16})\cdots$. Likewise, base 10 becomes $\frac{1}{1-x} = (1+x+x^2+\cdots+x^9)(1+x^{10}+x^{20}+\cdots+x^{90})\cdots$. What about non-constant bases? ...

Every integer has a unique representation as $n = \sum_{i=1}^{\infty} a_i i!$ with $a_i \in \{0,\ldots,i\}$.

Let p be a prime. Base p has a nice extensions to number theory. We can define a distance, the "p-adic" metric on $\mathbb{Z}$. Given $\gamma > 1$, usually $\gamma_p = p$. $d_p(m,n) = |m-n|_p = p^{-r}$, where $p^r \mid\mid (m-n)$....

We can define a distance metric $d_x(p(x),q(x))$ on polynomials in X.

$$d_x(p(x),q(x)) = |p(x)-q(x)| = 2^{-n}$$

if the smallest x^k with non-zero coefficients in $p(x)-q(x)$ is x^n. This defines a topology on $k[X]$.

Remark.

6.10 Inclusion Exclusion

Theorem 6.40. *Given finite sets $A_1,\ldots,A_n$, we have* $|\bigcup_{i=1}^n A_i| = \sum_{k=1}^n \sum_{|S|=k} (-1)^{k+1} \left|\bigcap_{s\in S} A_s\right|$.

Proof. Let $x \in \bigcup_{i=1}^n A_i$. Fix the subset $S \subseteq \{1,\ldots,n\}$ so that $x \in A_i$ if $i \in S$ and $x \notin A_j$ if $j \notin S$. Then $S \neq \emptyset$ since $x \in \bigcap_{i=1}^n A_i$....

Let $\mathcal{O}$ = sets of all objects. Given $S \subseteq \mathcal{P}$, $\mathcal{O}_S$ = objects with all of the properties in S (plus possibly some more). $N_{\geqslant S} = |\mathcal{O}_S|$. $N_{=S}$ = #objects having precisely the properties in S. So typically, $N_{\geqslant S}$ is easy to compute, and we want to be able to compute $N_{=S}$.

Easy $N_{\geqslant S} = \sum_{T\supseteq S} N_{=T}$. We'll assume $S = \emptyset$. So $N_{\geqslant\emptyset} = \sum_{T\subseteq\mathcal{P}} N_{=T}$. $\mathcal{P} = \emptyset$, $N_{\geqslant\emptyset} = N_{=\emptyset}$. $\mathcal{P} = \{1\}$. $N_{\geqslant\emptyset} = N_{=\emptyset} + N_{\geqslant 1}$. $N_{=\emptyset} = N_{\geqslant\emptyset} - N_{\geqslant 1}$. $N_{=\{2\}} = N_{\geqslant\{2\}} - N_{\geqslant\{1,2\}}$. $N_{=\{1\}} = N_{\geqslant\{1\}} - N_{\geqslant\{1,2\}}$. $\mathcal{P} = \{1,2\}$. $N_{\geqslant\emptyset} = N_{=\emptyset} + N_{=\{1\}} + N_{=\{2\}} + N_{=\{1,2\}}$. So $N_{\emptyset} = N_{\geqslant\emptyset} - N_{=\{1\}} - N_{=\{2\}} - N_{=\{1,2\}} = \ldots = \sum_{T\subseteq\mathcal{P}} N_{\geqslant T}(-1)^{|T|}$.

Exercise. Prove $N_{=\emptyset} = \sum_{T\subseteq\mathcal{P}} (-1)^{|T|} N_{\geqslant T}$.

Examples $\mathcal{P} = \{P_1,\ldots,P_n\}$. $\mathcal{O}$ = set of all permutations of $\{1,\ldots,n\}$. $P_i = \pi$ fixes i. $N_{=0}$ = # of dearrangement of $\{1,\ldots,n\}$ = # of fixed point-free permutations of $\{1,\ldots,n\}$. $N_{\geqslant T} = (n-|T|)!$. $N_{=0} = \sum_{T\subseteq\mathcal{P}} (-1)^{|T|}(n-|T|)! = \sum_{k=0}^n (-1)^k \binom{n}{k}(n-k)!$. ☐

We've seen how to compute $N_{\emptyset}$ from $N_{\geqslant T}$ over all subsets $T \subseteq \mathcal{P}$. This allows us to compute $N_{=0}$ = #objects with exactly 0 properties from P. What if we want $N_{=k}$ = # of objects with exactly k properties from P. Equivalently, what if we want to compute $\sum_{|T|=k} N_{=T}$? (for fixed k). Sometime not too bad.

Example 6.41. Permutation with exactly k fixed points could be written as

$$\binom{n}{k} \cdot D_{n-k} = \binom{n}{k} \sum_{j=0}^{n-1} (-1)^j \frac{(n-k)!}{j!}.$$

Can we improve on this (either in general or in this particular example)?

Then $N_{=T} = \sum_{T\subseteq S}(-1)^{|S|-|T|}N_{\geqslant S} = \sum_{j=k}^{n}\sum_{|S|=j,S\supseteq T}(-1)^{j-k}N_{\geqslant S}$. So

$$\sum_{N_{|T|=k}} N_{=T} = \sum_{N_{|T|=k}}\sum_{j=k}^{n}\sum_{|S|=j,T\subseteq S}(-1)^{j-k}N_{\geqslant S} = \sum_{j=k}^{n}\sum_{|S|=j}\sum_{|T|=k,T\subseteq S}N_{\geqslant S} = \sum_{j=k}^{n}\sum_{|S|=j}N_{\geqslant S}\binom{j}{k}.$$

In the case of derangement if $|S| = j$, $N_{\geqslant S} = (n-j)!$, so we get $\sum_j \binom{n}{j}(n-j)!(-1)^j\binom{j}{k}$.

6.11 Combinatorics and Probability

Suppose we have a non-negative integer valued random variable X with $E[X] < \infty$. So we can think of interpreting X as follows. We have a set Ω of objects, e.g., graphs. For each $G \in \Omega$, we count some non-negative integer parameter $X(G)$. Then $P(X > 0) = \sum_{k=1}^{\infty} P(X = k) \leqslant \sum_{k=1}^{\infty} kP(X = k) = \sum_{k=0}^{\infty} kP(X = k)$. So if $E[X] < 1$. then $P(X = 0) = 1 - P(X > 0) > 0$. Hence there exists a $G \in \Omega$ so that $X(G) = 0$. Similarly, if $E[X]$ is large, then there must exist $G \in \Omega$ with $X(G) \geqslant E[X]$; and if there exists $G \in \Omega$ with $X(G) < E[X]$, then there exists $G \in \Omega$ with $X(G) > E[X]$.

Example 6.42. A set $T \subseteq \{1, 2, 3, \cdots\}$ is sum-free if for $x, y, z \in T$, $x + y \neq z$.

Theorem 6.43 (AlonErdos Kleitmen)**.** *If S is a non-empty set of positive integer, then there exists $T \subseteq S$ with $|T| > |S|$ and T is sum-free.*

Q(Erdos) Can $\frac{1}{3}$ be replaced by some $\alpha > \frac{1}{3}$?

Proof. For $\alpha \in (0,1)$, defines $S_\alpha = \{n \in S : \{n\alpha\} \in (\frac{1}{3}, \frac{2}{3})\}$, where $\{n\alpha\}$ is the fractional part of $n\alpha$. Then, if $n_1, n_2 \in S_\alpha$, we have $\{(n_1 + n2)\alpha\} = \{n_1\alpha\} + \{n_2\alpha\} \pmod 1 \in [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. So $n_1 + n_2 \notin S_\alpha$. Hence S_α is sum-free.

Now let $X_\alpha = |S_\alpha|$. $X_\alpha = \sum_{n\in S} I_{n\in S_\alpha}$.

Exercise 6.44. Let $n > 0$ be an integer. Pick α uniformly in $[0, 1)$. Show $P(\{n\alpha\}) \in (\frac{1}{3}, \frac{2}{3}) = \frac{1}{3}$.

Hence $E[X_\alpha] = \sum_{n\in S}\frac{1}{3} = \frac{|S|}{3}$. Let $m = \max_{n\in S} n$. Since $|S_\alpha| = 0$ if $\alpha \in (0, \frac{1}{3m})$, $P(X_\alpha = 0) > \frac{1}{3m}$. Hence $P(X_\alpha > \frac{|S|}{3}) > 0$. Hence there exists α such that $T = S_\alpha$ satisfying $|T| > \frac{|S|}{3}$. How hard to improve upper bounds? Find a good set S of cardinality $|S| = s$. Consider all $\binom{S}{S/3+t}$ subsets of size $S/3 + t$. Huge search space.

Question. Can the probabilistic proof inform good choices for S with small max $|T|$? Can you show if S has "too many small element", then S has a larger sum-free subset?

Another combinatorial number theory question. Reference: B. Lindstrom. An Inequality for B_2 Sequences 1969.

Theorem 6.45. *There exists $C > 0$ such that if $S \subseteq \{1, \ldots, n\}$ is Sidon, $|S| < n^{\frac{1}{2}} + Cn^{\frac{1}{2}} + 1$, $\frac{7}{6}$.*

What is max$|S|, S$. □

Remark. Suppose you play a game, in which there is a non-negative integer payoff X, a random variable, $P(X = k) = p_k$, $0 \leqslant k \leqslant N$. How much should you pay to play the game? It is $E[X] = \sum_{k=0}^{N} kp_k$. Indeed, if you play the game for a sufficiently long period of time, then if you pay $< X$, you win over the long term, if you pay $> X$, you lose over the long term? How long is sufficiently long? If we play one game, at the rate of a Cesuim atom vibrating, about 10^{-10} period until the heal death of the universe, say 10^{110} seconds, we "only" get to play $T = 10^{120}$ rounds of the game. Can we use up with reasonable games whose $E[X]$ is too much to pay if we only get to pay T times?

Flip a coin n times? Payoff $X = c \cdot 3^k$ when $k = \#$ heads. $E[X] = \sum_{j=0}^{n} 3^j \binom{n}{j} 2^{-n} = c\frac{(1+3)^n}{2^n} = c2^n$. (So if we choose $c = 2^{-n}$, $E[X] = \$1$. What happens if we play the game just a few times? With high probability, we see that $\frac{n}{2}$ heads, say $\frac{n}{2} + i$, where i is not too much bigger than $\sqrt{n}$. So we win $2^{-n} \cdot 3^{\frac{n}{2}+1} = (\frac{4}{3})^{\frac{n}{2}} \cdot 3$ which is very small. How many heads would we need to see for us to get a payoff of \$1. Want $2^n = 3^j$. $j = \frac{\log 2}{\log 3} n = \alpha n$, where $\alpha = 0.6309 > \frac{1}{2}$.

What is P(in n coin tosses we see more than αn heads)? $\frac{1}{2^n} \sum_{k=\alpha n}^{n} \binom{n}{k} \approx \frac{1}{2^n} \binom{n}{\alpha} n d_\alpha$.

$$\frac{n!}{(\alpha n)!(1-\alpha)n!} = \frac{(\frac{n}{e})^n \sqrt{2\pi n}}{(\frac{\alpha n}{e})^{\alpha n} (\frac{(1-\alpha)n}{e})^{(1-\alpha)n} \sqrt{2\pi \alpha n}\sqrt{2\pi(1-\alpha)n}}.$$

$\frac{(\alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)})^n}{\sqrt{2\pi\alpha(1-\alpha)n}}$. When is $\frac{\alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)}}{2} \approx p^n = 0.97^n$.

Theorem 6.46 (Kruth)**.** *Given a rooted tree T, we wish to count the number of leaves in T. Algorithm. Take a random walk from the root to a leaf, choosing uniformly at each vertex among its children. Let X = product of the degrees you see. Then $E[X] = \#$leaves of T. Prove: give a leaf l. $X(l) = \prod d_i$ of degrees on path to l. Probability we end at is $\frac{i}{\prod d_i}$. So $\sum_l X(l)P(l) = \sum_l 1$.*

Extension: Label each edge from v to its children with a probability (not necessarily $\frac{1}{d(v)}$, so probability sum to 1).

Aside. Allowing for favorite children, how do you pick your favarite and how do you assign relative probability.

Remark. Non-attacking configuration of King. In one dimension, written vertically $F(k, 1) = \#k \times 1$ boards without attacking $k = F(k-1, 1) + F(k-2, 1)$ with $k \geqslant 3$. $F(1, 1) = 2$ and $F(0, 1) = 1$. (Check out Pingla, Sanskrit poetry, "Pascal Identity" and Fibonacci numbers. Also, Keith Devlin on why golding ratio is less interesting than you've been told.) To compute $F(k, m)$. Use Kruth. T has a depth n. Each vertex at depth j is labelled by a $k \times 1$ board B. Its children are the $k \times 1$ boards that can be adjacent to B.(graph). Let's check the $1 \times n$ case, 2 possible 1×1 boards. (graph).

Remark (depth 10000)**.** 2×10^7 simulations. X never exceed $E[X]$. Want to contral max $\frac{X(l_1)}{X(l_2)}$ for trees l_1 and l_2.

Theorem 6.47. *If we modifying the weight on the edge of T so that (graph), where $p + p^2 = 1$, i.e., $p = 0.6801$, whenever we are at, then all leaves have one of two X-values, with* max $\frac{X(l_1)}{X(l_2)} = \frac{1}{p}$.

Remark (More on Kruthian counting)**.** Counting 2 dimensional $k \times n$ configuration of non-attacking kings. Set up a matrix A, labelled by permissible $k \times 1$ columns of non-attacking kings. e.g., $k = 3$, possible 3×1 configuration are (graph). $A_3 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$. .. $A_k(i, j) = 1$ if columns i, j can be adjacent w/o attacking kings.

Remark (Constructing the Kruthian tree)**.** Start with the root, labelled with (empty col). Each vertex is labelled with a column. Children are permissible columns like thoese that can be adjacent to the current column. (graph). How should we adjust the probability of choosing each child so that the values of X of the random varible, when sampled, give a good estimation of $E[X]$.

The adjacency matrix A has dominant eigenvalue (i.e., the Peron Frobenius eigenvalue) $\varphi = \frac{1+\sqrt{5}}{2}$ and the corresponding eigenvector is $\begin{bmatrix} 1/\varphi \\ 1/\varphi^2 \end{bmatrix}$. Normalize to have $\frac{1}{\varphi} + \frac{1}{\varphi^2} = 1$. We saw last time that if $p = \frac{1}{\varphi}$, so $p + p^2 = 1$. (graph) then with these probabilities for children, the Kruthian variable X, when sampled, does give good estimates for $E[X]$.

For small k, we've verified that if we label the edges with probabilities derived from the PF eigenvector, (the only eigenvector with strictly positive entries), so for example, the probabilities in (graph) are normalized so that $\alpha + \beta = 1$ and $\alpha\alpha$ (graph) entry of eigenvector $\beta\alpha$ (graph) entry of eigenvector. Then at each level of the tree, there are a fixed number of values taken by X.

Conjecture: # values taken by X is at most F_k, the dimension of the matrix A_k.

Conjecture. Given a (big) tree, we're doing a weighted Kruthian sample from, then

(a) if when sampling X, the ratio of largest value of X seen to smallest is not too big, we have confidence $\overline{X} \approx E[X]$;

(b) if ratio of largest X to smallest X is too big, $\overline{X} << E[X]$.

Remark (Possibly BS in general)**.** In the case where the tree corresponds to long walks on finite graphs G, conjecture is probably trees with weights chosen from eigenvectors, should get convergence.

Let A be a non-negative symmetric (not necessarily) matrix, in which the undirected graph (with loops) G having $(i, j) \in E$ if and only if $A_{ij} > 0$ is connected, and not bipartite. (So high power of A have stricly positve entries). (In our case, A_k^2 has strictly positve entries), then there is a unique eigenvalue λ with $\lambda > |\lambda_j|$ for all other eigenvalues λ_j, $\lambda > 0$ and the corresponding eigenvector has strictly positve entries.

How to take a random Kruthian sample when G is too big to do an eigen analysis?

Continuing with the Kings problems.

In 2 dimensional $F(n, k) = \# \ k \times n$ configs. In 3 dimensional $F(k, m, n) = \#k \times m \times n$ configs. In any finite number of dimensions, we have two inequalities. If we take logs of $F()$, writing $h(n_1, \dots, n_d) = \log(F(n_1, \dots, n_d))$, then $h(n_1 + m_1, n_2, \dots, n_d) \leqslant h(n_1, n_2, \dots, n_d) + h(m_1, n_2, \dots, n_d)$ (with the same being true for the j^{th} component) and $h(n_1+m_1+1, n_2, \dots, n_d) \geqslant h(n_1, n_2, \dots, n_d)+h(m_1, n_2, \dots, n_d)$. So for example, with $d = 1$ and $F(0) = 1$, $F(1) = 2$, $F(2) = 3$. $F(3) = 5$, $F(4) = 8$, $F(5) = 13$, $F(6) = 21$, $F(7) = 34$, we have $F(m + n) \leqslant F(m)F(n)$ and $F(m + n + 1) \geqslant F(m)F(n)$. Example $F(2 + 3) = 13 \leqslant F(2)F(3) = 15$, $F(2 + 3 + 1) = 21 \geqslant F(2)F(3) = 15$. The inequality for $h(n_1, \dots, n_d)$ imply that $h(rn_1, \dots, rn_d) \leqslant r^d h(n_1, \dots, n_d)$. $h(rn_1 + r - 1, \dots, rn_d + r - 1) \geqslant r^d h(n_1, \dots, n_d)$. $(r - 1)^d h(n_1, \dots, n_d)$. This will imply that $\lim_{n_1,\dots,n_d \to \infty} \frac{1}{n_1 \cdots n_d} h(n_1, \dots, n_d)$ exists.

We can defined $\eta_d = \lim_{n_1,\dots,n_d \to \infty} \frac{1}{n_1 \cdots n_d} h(n_1, \dots, n_d)$. We can call this the entropy of the configuration of non-attacking king in d dimension.

What is known $\eta_1 = \log(\varphi) = \log(1.618) = \cdots = \log(\frac{1+\sqrt{5}}{2})$. In 2 dimensions, Probably know to 6 digits. Morally known to 60 digits. In dimensions greater or equl to 3 know first digit of e^{μ_d}, $e^{\mu_d} = 1$?

Restricting ourselves to $d = 2$. Recall $A_k =$. If we denote the largest eigenvalue of A_k by λ_k, then for fixed k, $F(k, n) = \#k \times n$ boards $= k \times (n+2)$ boards starting with (graph: emptyboard) and ending with (graph: emptysboard) $= (A_k^{n+1})_{1\times 1}$, (1,1)-entry in the $F_2 \times F_2$ matrix A_k^{n+1}.

Now $(A_k^{n_1}) = \sum_{\text{eigenvalue } \lambda \text{ of } A_k} c_\lambda \lambda^{n+1}$ for fixed and computational values λ. A_k is real and symmetric, so all λ's are real, by Person Frobenius, λ_k, the largest eigenvalue of A_k, is positive and has magnitude strictly greater than all other eigenvalues of A_k, implying that $F(k, n) = c_{\lambda_k} \lambda_k^{n+1}(1 + \sum_{\lambda \neq \lambda_k} \frac{c_\lambda}{c_{\lambda_k}} (\frac{\lambda}{\lambda_k})^{n+1})$. So $\frac{F(k,n)}{c_{\lambda_k} \lambda_k^{n+1}} - \log(c_{\lambda_k} \lambda_k^{n+1}) \to 0$, $\frac{1}{n} \log F(k, n) - \frac{1}{n} \log(c_{\lambda_k}) - \frac{n+1}{n} \log \lambda_k \to 0$ as $n \to \infty$, $c_{\lambda_k} > 0$.

So $\frac{1}{n} \log(F(k, n)) \to \log(\lambda_k)$ or $F(k, n)^{\frac{1}{n}} \to \lambda_k$. Hence $\lim_{k\to\infty} \frac{1}{k} \log(\lambda_k) = \eta_2$. Experimentally it appears that $\frac{1}{k} \log \lambda_k - \eta$ alternates positive and negative. (Compare to the alternating series test. $\sum_{k=-1}^{n} \frac{(-1)^{k+1}}{k}$ alternating above and below i^{th} limit).

Example 6.48.

Remark. Two views of permutations.

(a) Function from $\{1, \ldots, n\}$ to $\{1, \ldots, n\}$. Represented by for example

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 1 & 2 & 4 & 3 & 7 & 6 \end{pmatrix},$$

which we could abbreviate to

$$(5\ 1\ 2\ 4\ 3\ 7\ 6).$$

(b) Can list a permutation by its cycles

$$(1\ 5\ 3\ 2)(4)(6\ 7),$$

where

$$(1\ 5\ 3\ 2) = (5\ 3\ 2\ 1) = (3\ 2\ 1\ 5) = (2\ 1\ 5\ 3).$$

Replace each cycle by the one with the largest entry first.

$$(5\ 3\ 2\ 1)(4)(7\ 6).$$

Write cycles in increasing order of their largest element

$$(4)(5\ 3\ 2\ 1)(7\ 6).$$

Erase parenthesis,

$$4\ 5\ 3\ 2\ 1\ 7\ 6.$$

This gives us a 1-line representation of

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 5 & 3 & 2 & 1 & 7 & 6 \end{pmatrix}.$$

So we have a map from $S_n \to S_n$. Can we undo it? 4 5 3 2 1 7 6. $4 < 5$, so (4). $5 > 3 > 2 > 1$ and $1 < 7$, so (5 3 2 1). $7 > 6$, so (7 6). This transformation is due to Domimiefiite referred to as the Foate Transformation.

Q: Pick a permutation in S_n uniformly at random. What is the probability that j is in a cycle of length k. Observation: by relabelling j as in If n is in a cycle of length k, then the modified cycle notation will be $(\quad)(\quad)(\underbrace{n\quad}_{k\text{terms}})$.....