Analytic Number Theory

January 29, 2019

Contents

0 Introduction

—Lecture 1—

Lecturer: Thomas Bloom (*tb634@cam.ac.uk*, *www.thomasbloom.org/ant.html*)

Printed notes will be updated, but 1-2 weeks behind.

Example classes: weeks 3,5,7, tuesdays 330-5pm; prop-in sessions weeks 2,4,6,8. Rooms to be confirmed later.

What is analytic number theory? It's the study of numbers (regular integers, discrete) using analysis (real/complex, continuous) and some other quantitative questions.

For example, for the famous function $\pi(x)$, the number of primes no greater than x, we know $\pi(x) \sim \frac{x}{\log x}$.

Throughout this course, by *numbers* we'll mean natural numbers excluding 0.

We can also ask how many twin primes there are, i.e. how many p such that $p, p+2$ are both prime. This is not known yet (not even the finiteness); but from 2014, Zhang, Maynard, Polymath showed that there are infinitely many primes at most 246 apart, which is not that far from 2. The current guess is that the number is around $\frac{x}{(\log x)^2}$.

Another question we may ask: how many primes are there $\equiv a \pmod q$, $(a,q) = 1$. We know by Dirichlet's theorem that there are infinitely many.
A natural guess of the count is $\frac{1}{\phi(q)} \frac{x}{\log x}$, where $\phi(x)$ is the Euler Totient function. This is known to hold for small q.

In this course we'll talk about:
(1) Elementary techniques (real analysis);
(2) Sieve methods;
(3) Riemann zeta function/prime number theory (complex analysis);
(4) Primes in arithmetic progressions.

1 Elementary techniques

Review of asymptotic notation:
• $f(x) = O(g(x))$ if there is $c > 0$ s.t. $|f(x)| \leq c|g(x)|$ for all large enough x;
• $f \ll g$ is the same thing as $f = O(g)$. This also defines what $f \gg g$ means in the natural way;
• $f \sim g$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$ (i.e. $f = (1 + o(1))g$);
• $f = o(g)$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$.

1.1 Arithmetic functions

Arithmetic functions are just functions $f : \mathbb{N} \to \mathbb{C}$; in other words, relabelling natural numbers with some complex numbers.
An important operation for multiplicative number theory ($fg = f(n)g(n)$) is multiplicative convolution,

$$f * g(n) = \sum_{ab=n} f(a)g(b)$$

Examples: $1(n) \equiv 1 \forall n$ (caution: 1 is not the identity function, and $1 * f \neq f$).
Möbius function:

$$\mu(n) = \begin{cases} (-1)^k & \text{if } n = p_1...p_k \\ 0 & \text{if } n \text{ is divisible by a square} \end{cases}$$

Liouville function: $\lambda(n) = (-1)^k$ if $n = p_1...p_k$ (primes not necessarily distinct),
Divisor function: $\tau(n) =$ number of d s.t. $d|n = \sum_{ab=n} 1 = 1 * 1$. This is sometimes also known as $d(n)$.

An arithmetic function is multiplicative if $f(nm) = f(n)f(m)$ when $(n, m) = 1$. In particular, a multiplicative function is determined by its values on prime powers.

Fact. If f, g are multiplicative, then so is $f * g$.
All the function we've seen so far $(\mu, \lambda, \tau, 1)$ are multiplicative.

Non-example: $\log n$ is definitely not multiplicative.

Fact. (Möbius inversion)
$1 * f = g \iff \mu * g = f$. That is,

$$\sum_{a|n} f(d) = g(n) \forall n \iff \sum_{d|n} g(d)\mu(n/d) = f(n) \forall n$$

e.g.

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & n = 1 \\ 0 & \text{else} \end{cases} = 1 * \mu$$

is multiplicative: it's enough to check identity for primes powers.
If $n = p^k$ then $\{d|n\} = \{1, p, ..., p^k\}$. So LHS=$1 - 1 + 0 + 0 + ... = 0$, unless $k = 0$ when LHS $= \mu(1) = 1$.

Our goal is to study primes. The first guess might be to work with

$$1_p(n) = \begin{cases} 1 & n \text{ prime} \\ 0 & \text{else} \end{cases}$$

(e.g. $\pi(x) = \sum_{1 \le n \le x} 1_p(n)$). Instead, we work with von Mangoldt funcion

$$\wedge(n) = \begin{cases} \log p & n \text{ is a prime power} \\ 0 & \text{else} \end{cases}$$

(e.g. in a few lectures we'll look at $\psi(x) = \sum_{1 \le n \le x} \wedge(n)$).

Lemma. (1)
$1 * \wedge = \log$, and by Möbius inversion, $\mu * \log = \wedge$.
Note that it's easy to realize that $\wedge$ is not multiplicative, else log will be.

Proof. $1 * \wedge(n) = \sum_{d|n} \wedge(d)$. So if $n = p_1^{k_1}...p_r^{k_r}$, then above

$$\begin{aligned} &= \sum_{i=1}^{r} \sum_{j=1}^{k_i} \wedge(p_i^j) \\ &= \sum_{i=1}^{r} \sum_{j=1}^{k_i} \log(p_i) \\ &= \sum_{i=1}^{r} k_i \log(p_i) \\ &= \log n \end{aligned}$$

□

Note that the above tells us

$$\begin{aligned} \wedge(n) &= \sum_{d|n} \mu(d) \log(n/d) \\ &= \log n \sum_{d|n} \mu(d) - \sum_{d|n} \mu(d) \log d \\ &= -\sum_{d|n} \mu(d) \log d \end{aligned}$$

by the famous fact that $\sum_{d|n} \mu(d) = 0$ unless $n = 1$; but when $n = 1$, $\log n = 0$.
Now we can try to evaluate

$$\begin{aligned} -\sum_{1 \le n \le x} \wedge(n) &= \sum_{1 \le n \le x} \sum_{d|n} \mu(d) \log d \\ &= -\sum_{d \le x} \mu(d) \log d \Big(\sum_{1 \le n \le x, d|n} 1 \Big) \quad \text{(reverse order of summation)} \end{aligned}$$

But

$$\sum_{1\leq n\leq x, d|n} 1 = \lfloor x/d \rfloor = x/d + O(1)$$

So we know the original sum is equal to

$$-x\sum_{d\leq x}\mu(d)\frac{\log d}{d} + O(\sum_{d\leq x}\mu(d)\log d)$$

—Lecutre 2—

Lecturer's favourite book: *Multiplicative Number Theory.*

Room for example classes: MR14 (Tues 330-5pm, week 357).

1.2 Summation

Given an arithmetic function f, we can ask for estimates of $\sum_{1\leq n\leq x} f(n)$.
We say that f has *average order* g if $\sum_{1\leq n\leq x} f(n) \sim xg(x)$ (in some sense, the average size of f is g).

For example, if $f \equiv 1$, then $\sum_{1\leq n\leq x} f(n) = \lfloor x \rfloor = x + O(1) \sim x$. So the average order of 1 is 1 (makes a lot of sense).
A slightly less trivial example is the identity function $f(n) = n$: we have $\sum_{1\leq n\leq x} n \sim \frac{x^2}{2}$, so the average order of n is $n/2$.

Lemma. (1, Partial summation)
If (a_n) is a sequence of complex numbers, and f is s.t. f' is continuous. Then $\sum_{1\leq n\leq x} a_n f(n) = A(x)f(x) - \int_1^x A(t)f'(t)dt$, where $A(x) = \sum_{1\leq n\leq x} a_n$.
We can see that this is a discrete version of integration by parts.

Proof. Suppose $x = N$ is an integer. Note that $a_n = A(n) - A(n-1)$. So

$$\begin{aligned}\sum_{1\leq n\leq N} a_n f(n) &= \sum_{1\leq n\leq N} f(n)(A(n) - A(n-1)) \\ &= A(N)f(N) - \sum_{n=1}^{N-1} A(n)(f(n+1) - f(n)) \text{ using } A(0) = 0\end{aligned}$$

Now $f(n+1) - f(n) = \int_n^{n+1} f'(t)dt$. So

$$\begin{aligned}\sum_{1\leq n\leq N} a_n f(n) &= A(N)f(N) - \sum_{n=1}^{N-1} A(n)\int_n^{n+1} f'(t)dt \\ &= A(N)f(N) - \int_1^N A(t)f'(t)dt\end{aligned}$$

To be complete, we should also consider the case where x is not an integer. But if $N = \lfloor x \rfloor$,

$$\begin{aligned} A(x)f(x) &= A(N)f(x) \\ &= A(N)\left(f(N) + \int_N^x f'(t)dt\right) \end{aligned}$$

□

Lemma. (2)

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

where γ is some constant.

Proof. Apply partial summation with $f(x) = \frac{1}{x}$ and $a_n \equiv 1$, so $A(x) = \lfloor x \rfloor$. Then, writing $\lfloor t \rfloor = t - \{t\}$,

$$\begin{aligned} \sum_{1 \le n \le x} \frac{1}{n} &= \frac{\lfloor x \rfloor}{x} + \int_1^x \frac{\lfloor t \rfloor}{t^2} dt \\ &= 1 + O\left(\frac{1}{x}\right) + \int_1^x \frac{1}{t} dt - \int_1^x \frac{\{t\}}{t^2} dt \\ &= 1 + O\left(\frac{1}{x}\right) + \log x - \int_1^\infty \frac{\{t\}}{t^2} dt + \int_x^\infty \frac{\{t\}}{t^2} dt \\ &= \gamma + O\left(\frac{1}{x}\right) + \log x + O\left(\frac{1}{x}\right) \\ &= \log x + \gamma + O\left(\frac{1}{x}\right) \end{aligned}$$

where at the penultimate step we bound the error term by

$$\begin{aligned} \int_x^\infty \frac{\{t\}}{t^2} dt &\le \int_x^\infty \frac{1}{t^2} dt \\ &\le \frac{1}{x} \end{aligned}$$

and we actually know $\gamma = 1 - \int_1^\infty \frac{\{t\}}{t^2} dt$.
This γ is called Euler's constant (Euler-Mascheroni).
We know very little about this constant: we only know $\gamma = 0.577...$, and we don't even know if γ is irrational. □

Lemma. (3)

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

Proof. Use partial summation again, with $f(x) = \log x$ and $a_n = 1$, so $A(x) = \lfloor x \rfloor$:

$$\begin{aligned}\sum_{1\le n\le x} \log n &= \lfloor x \rfloor \log x - \int_1^x \frac{\lfloor t \rfloor}{t} dt \\ &= x \log x + O(\log x) - \int_1^x 1 dt + O(\int_1^x \frac{1}{t} dt) \\ &= x \log x - x + O(\log x)\end{aligned}$$

□

1.3 Dinsar function

Recall that $\tau(n) = 1 * 1(n) = \sum_{d|n} 1$.

Theorem. (4)

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

So average order of τ is $\log x$.

Proof. Note that we won't apply partial summation here: PS allows to get $\sum a_n f(n)$ from knowledge of $\sum a_n$; but $\tau(n)$ here is not differentiable, so PS is not going to apply.

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

where we applied lemma 2 at the last step. This is all correct, but the error term is larger than what we wanted. However, we have indeed prove that the average order of $\tau(x)$ is $\log x$.

To reduce error term, we use (Dirichlet's) hyperbola trick:

$$\begin{aligned}\sum_{1\le n\le x}\tau(n) &= \sum_{1\le n\le x}\sum_{ab=n}1\\ &= \sum_{ab\le x}1\\ &= \sum_{a\le x}\sum_{b\le \frac{x}{a}}1\end{aligned}$$

Note that now we're just counting number of integer points below the hyperbola $xy = n$ (relabelling variables).
When summing over $ab \le x$, we can sum over $a, b \le x^{1/2}$ separately, then subtract the repetition off. Then

$$\begin{aligned}\sum_{1\le n\le x}\tau(n) &= \sum_{a\le x^{1/2}}\sum_{b\le \frac{x}{a}}1+\sum_{b\le x^{1/2}}\sum_{a\le \frac{x}{b}}1-\sum_{a,b\le x^{1/2}}1\\ &= 2\sum_{a\le x^{1/2}}\lfloor\frac{x}{a}\rfloor-\lfloor x^{1/2}\rfloor^2\\ &= 2\sum_{a\le x^{1/2}}\frac{x}{a}+O(x^{1/2})-x+O(x^{1/2})\end{aligned}$$

by noting that $\lfloor x^{1/2}\rfloor^2 = (x^{1/2}+O(1))^2$. Now the above equals

$$\begin{aligned}&= 2x\log x^{1/2}+2\gamma x-x+O(x^{1/2})\\ &= x\log x+(2\gamma-1)x+O(x^{1/2})\end{aligned}$$

□

Improving this $O(x^{1/2})$ error term is a famous and hard problem. We should probably get $O(x^{1/4+\varepsilon})$, but this is open. The best known result is $O(x^{0.3149\ldots})$.

Note that this does *not* mean that $\tau(n) \ll \log n$. The average order is small doesn't say about the individual values being small.

We'll state the theorem we're proving and prove it in the next lecture:

Theorem. (5)

$$\tau(n) \le n^{O(\frac{1}{\log\log n})}$$

In particular, $\tau(n) \ll_\varepsilon n^\varepsilon\ \forall\varepsilon > 0$.

—Lecture 3—

Proof. τ is multiplicative, so it's enough to calculate at prime powers.
Now, $\tau(p^k) = k+1$. So if $n = p_1^{k_1}\ldots p_r^{k_r}$, thjen $\tau(n) = \prod_{i=1}^r(k_i+1)$.
Let ε be chosen later, and consider $\frac{\tau(n)}{n^\varepsilon} = \prod_{i=1}^r\frac{k_i+1}{p_i^{k_i\varepsilon}}$. Note as $p\to\infty$, $\frac{k+1}{p^{k\varepsilon}}\to 0$.
In particular, if $p \ge 2^{1/\varepsilon}$, then $\frac{k+1}{p^{k\varepsilon}} \le \frac{k+1}{2^k} \le 1$. What about for small p? We

can't do better than $p \geq 2$, but that's enough.
In this case, $\frac{k+1}{p^{k\varepsilon}} \leq \frac{k+1}{2^{k\varepsilon}} \leq \frac{1}{\varepsilon}$ (as $x + \frac{1}{2} \leq 2^x \implies \varepsilon k + \varepsilon \leq 2^{k\varepsilon} \forall x \geq 0$), for $\varepsilon \leq 1/2$.
So

$$\frac{\tau(n)}{n^\varepsilon} \leq \prod_{i=1, i<2^{1/\varepsilon}}^{r} \frac{k_i+1}{p^{k_i\varepsilon}} \leq (1/\varepsilon)2^{1/\varepsilon}$$

Now choose optimal ε:
(trick!) if you want to choose x to minimise $f(x) + g(x)$, choose x s.t. $f(x) = g(x)$.
So here, $\tau(n) \leq n^\varepsilon \varepsilon^{-2^{1/\varepsilon}} = exp(\varepsilon \log n + 2^{1/\varepsilon} \log 1/\varepsilon)$.
Choose ε s.t. $\log n \approx 2^{1/\varepsilon}$, i.e. $\varepsilon = \frac{1}{\log\log n}$. So

$$\begin{aligned}\tau(n) &\leq n^{1/\log\log n}(\log\log n)^{2^{\log\log n}} \\ &= n^{1/\log\log n} e^{(\log n)^{\log 2}\log\log\log n} \\ &\leq n^{O(\frac{1}{\log\log n})}\end{aligned}$$

□

1.4 Estimates for the Primes

Recall $\pi(x)$ is the number of primes $\leq x = \sum_{1\leq n\leq x} 1_p(n)$, and $\psi(x) = \sum_{1\leq n\leq x} \Lambda(n)$.
The prime number theorem states that $\pi(x) \sim \frac{x}{\log x}$, or equivalently $\psi(x) \sim x$ (justified later).
It was 1850 before the correct magnitude of $\pi(x)$ was proved. Chebyshev showed that $\pi(x) \asymp x/\log x$, where $f \asymp g$ means $g \ll f \ll g$.

Theorem. (6, Chebyshev)
$\psi(x) \asymp x$.
We'll show below that $(\log 2)x \leq \psi(x) \leq (\log 4)x$ (remember that the default base for log is e, so $\log 2 < 1$ while $\log 4 > 1$).

Proof. First we'll prove the lower bound. Recall $1 * \Lambda = \log$, i.e. $\sum_{ab=n} \Lambda(a) = \log n$. The (genuine) trick is to find a sum Σ s.t. $\varepsilon \leq 1$(?). We'll use the identity $\lfloor x \rfloor \leq 2\lfloor \frac{x}{2} \rfloor + 1 \forall x \geq 0$. Why? Say $\frac{x}{2} = n + \theta$, $\theta \in [0, 1)$ Then $\lfloor \frac{x}{2} \rfloor = n$, and $x = 2n + 2\theta$, and so $\lfloor x \rfloor = 2n$, or at most $2n + 1$.

So

$$\begin{aligned}
\psi(x) &\geq \sum_{n\leq x} \Lambda(n)(\lfloor \frac{x}{n} \rfloor - 2\lfloor \frac{x}{2n} \rfloor) \\
&= \sum_{n\leq x} \Lambda(n) \sum_{m\leq x/n} 1 - 2\sum_{n\leq x} \Lambda(n) \sum_{m\leq \frac{x}{2n}} 1 \\
&= \sum_{nm\leq x} \Lambda(n) - 2 \sum_{nm\leq x/2} \Lambda(n), \text{ write } d = nm, \\
&= \sum_{d\leq x} 1 * \Lambda(d) - 2 \sum_{d\leq x/2} 1 * \Lambda(d) \\
&= \sum_{d\leq x} \log d - 2 \sum_{d\leq x/2} \log d \\
&= x\log x - x + O(\log x) - 2(\frac{x}{2}\log\frac{x}{2} - \frac{x}{2} + O(\log x)) \\
&= (\log 2)x + O(\log x) \gg x
\end{aligned}$$

For the upper bound, note that $\lfloor x \rfloor = 2\lfloor x/2 \rfloor + 1$ for $x \in (1,2)$, so

$$\sum_{x/2<n<x} \Lambda(n) = \sum_{x/2<n<x} \Lambda(n)(\lfloor x/n \rfloor - 2\lfloor x/2n \rfloor) \leq \sum_{1\leq n\leq x} \Lambda(n)(\lfloor x/n \rfloor - 2\lfloor x/2n \rfloor)$$

so $\psi(x) - \psi(x/2) \leq (\log 2)x + O(\log x)$.
So $\psi(x) = (\psi(x) - \psi(x/2)) + (\psi(x/2) - \psi(x/4)) + ... \leq \log 2(x + x/2 + x/4 + ...) = (2\log 2)x$ (note only $\log x$ error terms at most). □

Lemma. (7)

$$\sum_{p\leq x, p \text{ primes}} \frac{\log p}{p} = \log x + O(1)$$

Proof. Recall that $\log = 1 * \Lambda$. So

$$\begin{aligned}
\sum_{n\leq x} \log n &= \sum_{ab\leq x} \Lambda(a) \\
&= \sum_{a\leq x} \Lambda(a) \sum_{b\leq x/a} 1 \\
&= \sum_{a\leq x} \Lambda(a)\lfloor x/a \rfloor \\
&= x\sum_{a\leq x} \frac{\Lambda(a)}{a} + O(\psi(x)) \\
&= x\sum_{a\leq x} \frac{\Lambda(a)}{a} + O(x)
\end{aligned}$$

But $\sum_{n\leq x} \log n = x\log x - x + O(\log x)$. So

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

It remains to note that

$$\begin{aligned}\sum_{p\le x}\sum_{n=2}^{\infty}\frac{\log p}{p^n} &= \sum_{p\le x}\log p\sum_{k=2}^{\infty}\frac{1}{p^k}\\ &= \sum_{p\le x}\frac{\log p}{p^2-p}\\ &\le \sum_{p=2}^{\infty}\frac{1}{p^{3/2}} = O(1)\end{aligned}$$

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

—Lecture 4—

Drop-in: Tuesday 4pm-5pm.

Lemma. (8)
$\pi(x) = \frac{\psi(x)}{\log x} + O(\frac{x}{(\log x)^2})$.
In particular, $\pi(x) \asymp \frac{x}{\log x}$ and prime number theorem: $\pi(x) \sim \frac{x}{\log x}$ is equivalent to $\psi(x) \sim x$.
So from now on, we'll call this the prime number theorem instead.

Proof. The idea is to use partial summation:

$$\theta(x) := \sum_{p\le x}\log p = \pi(x)\log x - \int_1^x \frac{\pi(t)}{t}dt$$

but this doesn't work immediately, since $\psi(x) = \sum_{n\le x}\Lambda(n) = \sum_{p^k\le x}\log p$. However, we have

$$\begin{aligned}\psi(x) - \theta(x) &= \sum_{k=2}^{\infty}\sum_{p^k\le x}\log p\\ &= \sum_{k=2}^{\infty}\theta(x^{1/k})\\ &\le \sum_{k=2}^{\infty}\psi(x^{1/k})\\ &= \sum_{k=2}^{\log x}\psi(x^{1/k})\end{aligned}$$

as the larger terms are all zero. Then the above

$$\begin{aligned}&\ll \sum_{k=2}^{\log x} x^{1/k}\\ &\ll x^{1/2}\log x\end{aligned}$$

(Obviously we could do better, but that's less important). Now $\psi(x) = \pi(x)\log x + O(x^{1/2}\log x) - \int_1^x \frac{\pi(t)}{t}dt$. Note that we have $\pi(t) \ll \frac{t}{\log t}$, so we can bound the above by

$$\begin{aligned}\psi(x) &= \pi(x)\log x + O(x^{1/2}\log x) + O(\int_1^x \frac{1}{\log t}dt)\\ &= \pi(x)\log x + O(\frac{x}{\log x})\end{aligned}$$

(For $\pi(t) \ll \frac{t}{\log t}$, note that from $\pi(t) \le t$, $\psi(x) = \pi(x)\log x + O(x^{1/2}\log x) + O(x)$. So $\pi(x)\log x = O(x)$). □

Lemma. (9)
$\sum_{p\le x}\frac{1}{p} = \log\log x + b + O(\frac{1}{\log x})$, where b is some constant.

Proof. We use partial summation. Let $A(x) = \sum_{p\le x}\frac{\log p}{p} = \log x + R(x)$ (so $R(x) \ll 1$ (by lemma 7)). Then

$$\begin{aligned}\sum_{2\le p\le x}\frac{1}{p} &= \frac{A(x)}{\log x} - \int_2^x \frac{A(t)}{t(\log t)^2}dt\\ &= 1 + O(\frac{1}{\log x}) + \int_2^x \frac{1}{t\log t}dt + \int_2^x \frac{R(t)}{t(\log t)^2}dt\end{aligned}$$

Note $\int_2^\infty \frac{R(t)}{t(\log t)^2}dt$ exists, say $= c$. Then

$$\begin{aligned}\sum_{2\le p\le x}\frac{1}{p} &= 1 + c + O(\frac{1}{\log x}) + \log\log x - \log\log 2 + O(\int_x^\infty \frac{1}{t(\log t)^2}dt)\\ &= \log\log x + b + O(\frac{1}{\log x})\end{aligned}$$

□

Theorem. (10, Chebyshev)
If $\pi(x) \sim c\frac{x}{\log x}$, then $c = 1$.
Note that this is weaker than PNT itself: this only says that *if* that relation exists, then we must have $c = 1$.
(Also, if $\pi(x) \sim \frac{x}{\log x - A(x)}$, then $A \sim 1$)

Proof. Use partial summation on $\sum_{p\le x}\frac{1}{p}$:

$$\sum_{p\le x}\frac{1}{p} = \frac{\pi(x)}{x} - \int_1^x \frac{\pi(t)}{t^2}dt$$

If $\pi(x) = (c + o(1))\frac{x}{\log x}$, then

$$\begin{aligned}&= \frac{c}{\log x} + o(\frac{1}{\log x}) + (c + o(1))\int_1^x \frac{1}{t\log t}dt\\ &= O(\frac{1}{\log x}) + (c + o(1))\log\log x\end{aligned}$$

But $\sum_{p\le x}\frac{1}{p} = (1 + o(1))\log x$. Hence $c = 1$. □

Lemma. (11)
$\prod_{p\leq x}(1-\frac{1}{p})^{-1} = c\log x + O(1)$, where c is some constant.

Proof.

$$\begin{aligned}\log(\prod_{p\leq x}(1-\frac{1}{p})^{-1}) &= -\sum_{p\leq x}(1-\frac{1}{p}) \\ &= \sum_{p\leq x}\sum_{k}\frac{1}{kp^k} \\ &= \sum_{p\leq x}\frac{1}{p} + \sum_{k\geq 2}\sum_{p\leq x}\frac{1}{kp^k} \\ &= \log\log x + c' + O(\frac{1}{\log x})\end{aligned}$$

where we used the expansion $\log(1-t) = -\sum_k \frac{t^k}{k}$.
Now note that $e^x = 1 + O(x)$ for $|x| \leq 1$. So

$$\begin{aligned}\prod_{p\leq x}(1-\frac{1}{p})^{-1} &= c\log x e^{O(\frac{1}{\log x})} \\ &= c\log x(1+O(\frac{1}{\log x})) \\ &= c\log x + O(1)\end{aligned}$$

□

It turns out that $c = e^{\gamma} \approx 1.78...$, where γ is the Euler constant that we've seen previously.

So why is PNT hard, given that we've proved so many results? From probabilistic heuristic, we have the 'probability' that $p|n$ is $\frac{1}{p}$.
What is the probability that n is prime then? n is prime iff n has no prime divisors $\leq n^{1/2}$. Our guess is that the events 'divisible by p' are independent, so the probability that n is prime should be something like $\prod_{p\leq n^{1/2}}(1-\frac{1}{p}) \approx \frac{1}{c\log n^{1/2}} = \frac{2}{c}\frac{1}{\log n}$. So

$$\pi(x) = \sum_{n\leq x} 1_{n\ prime} \approx \frac{2}{c}\sum_{n\leq x}\frac{1}{\log n} \approx \frac{2}{c}\frac{x}{\log x} \approx 2e^{-\gamma}\frac{x}{\log x}$$

if the above guesses are correct; but from theorem 10 we know that the constant should be 1 instead of $2e^{-\gamma} \approx 1.122....$
What have gone wrong? It turns out that the error terms accumulated are too overwhelming that they've actually contributed to the main term. So PNT is not something like we find the main term and prove that the error terms are negligible.

Recall that $1 * \Lambda = \log$, so $\mu * \log = \Lambda$. So

$$\begin{aligned}\psi(x) &= \sum_{n \leq x} \Lambda(n) \\ &= \sum_{ab \leq x} \mu(a) \log b \\ &= \sum_{a \leq x} \mu(a) (\sum_{b \leq \frac{x}{a}} \log b)\end{aligned}$$

Recall that

$$\begin{aligned}\sum_{m \leq x} \log m &= x \log x - x + O(\log x), \\ \sum_{m \leq x} \tau(m) &= x \log x + (2\gamma - 1)x + O(x^{1/2})\end{aligned}$$

so their main terms agree.
So

$$\begin{aligned}\psi(x) &= \sum_{a \leq x} \mu(a) \left(\sum_{b \leq \frac{x}{a}} \tau(b) + 2\gamma \frac{x}{a} + O\left(\frac{x^{1/2}}{a^{1/2}} \right) \right) \\ &= \sum_{ab \leq x} \mu(a) \tau(b) \\ &= \sum_{abc \leq x} \mu(a) \\ &= \sum_{b \leq x} \sum_{ac \leq x/b} \mu(a) \\ &= \sum_{b \leq x} \sum_{d \leq x/b} \mu * 1(d) \\ &= \lfloor x \rfloor = x + O(1)\end{aligned}$$

since we know the last term is 0 unless d is 1.

The error term:

$$-2\gamma \sum_{a \leq x} \mu(a) \frac{x}{a} = O(x \sum_{a \leq x} \frac{\mu(a)}{a})$$

so we need to show that $\sum_{a \leq x} \frac{\mu(a)}{a} = o(1)$. However, this is still the same as PNT, so we haven't gained anything.

—Lecture 5—

1.5 Selberg's identity, and an elementary proof of the PNT

Recall that PNT is

$$\psi(x) = \sum_{n \leq x} \Lambda(n) = x + o(x)$$

Let (*Selberg's function*)

$$\Lambda_2(n) = \mu * (\log^2)(n) = \sum_{ab=n} \mu(a)(\log b)^2$$

(Recall $\Lambda = \mu * \log$).

The idea is to prove a 'PNT for Λ_2' with elementary methods.

Lemma. (12)
(1) $\Lambda_2(n) = \Lambda(n)\log n + \Lambda * \Lambda(n)$;
(2) $0 \leq \Lambda_2(n) \leq (\log n)^2$;
(3) If $\Lambda_2(n) \neq 0$, then n has at most 2 distinct prime factors.

Proof. For (1), we use Möbius inverison, so it is enough to show that

$$\begin{aligned}
\sum_{d|n}(\Lambda(d)\log d + \Lambda * \Lambda(d)) &= (\log n)^2 \\
&= \sum_{d|n} \Lambda(d)\log d + \sum_{ab|n} \Lambda(a)\Lambda(b) \text{ as } 1 * \Lambda = \log \\
&= \sum_{d|n} \Lambda(d)\log d + \sum_{a|n} \Lambda(a) \underbrace{\left(\sum_{b|\frac{n}{a}} \Lambda(b)\right)}_{=\log(\frac{n}{d})} \\
&= \sum_{d|n} \Lambda(d)\log d + \sum_{d|n} \Lambda(d)\log(\frac{n}{d}) \\
&= \log n \sum_{d|n} \Lambda(d) = (\log n)^2
\end{aligned}$$

For (2), $\Lambda_2(n) \geq 0$ since both terms on RHS in (1) are ≥ 0, and since $\sum_{d|n} \Lambda_2(d) = (\log n)^2$, $\Lambda_2(n) \leq (\log n)^2$.
For (3), note that if n is divisible by 3 distinct primes, then $\Lambda(n) = 0$, and $\Lambda * \Lambda(n) = \sum_{ab=n} \Lambda(a)\Lambda(b) = 0$ since at least one of a or b has ≥ 2 distinct prime divisors. □

Theorem. (13, Selberg)

$$\sum_{n \leq x} \Lambda_2(n) = 2x\log x + O(x)$$

Proof.

$$\begin{aligned}\sum_{n\le x}\Lambda_2(n) &= \sum_{n\le x}\mu * (\log)^2(n)\\ &= \sum_{ab\le x}\mu(a)(\log b)^2\\ &= \sum_{a\le x}\mu(a)\left(\sum_{b\le \frac{x}{a}}(\log b)^2\right)\end{aligned}$$

By PS,

$$\sum_{m\le x}(\log m)^2 = x(\log x)^2 - 2x\log x + 2x + O((\log x)^2)$$

By PS, (let $A(t) = \sum_{n\le t}\tau(n) = t\log t + ct + O(t^{1/2})$)

$$\begin{aligned}\sum_{m\le x}\frac{\tau(m)}{m} &= \frac{A(x)}{x} + \int_1^x \frac{A(t)}{t^2}dt\\ &= \log x + c + O(x^{-1/2}) + \int_1^x \frac{\log t}{t}dt + c\int_1^x \frac{1}{t}dt + O(\int_1^x \frac{1}{t^{3/2}}dt)\\ &= \frac{(\log x)^2}{2} + c_1\log x + c_2 + O(x^{-1/2})\end{aligned}$$

So

$$\frac{x(\log x)^2}{2} = \sum_{m\le x}\tau(m)\frac{x}{m} + c_1'\sum_{m\le x}\tau(m) + c_2'x + O(x^{1/2})$$

So

$$\sum_{n\le x}(\log m)^2 = 2\sum_{m\le x}\tau(m)\frac{x}{m} + c_3\sum_{m\le x}\tau(m) + c_4x + O(x^{1/2})$$

So

$$\sum_{n\le x}\Lambda_2(n) = 2\sum_{a\le x}\mu(a)\sum_{b\le\frac{x}{a}}\frac{\tau(b)x}{ab} + c_5\sum_{a\le x}\mu(a)\sum_{b\le\frac{x}{a}}\tau(b) + c_6\sum_{a\le x}\mu(a)\frac{x}{a} + O(\sum_{a\le x}\frac{x^{1/2}}{a^{1/2}})$$

First, note that $x^{1/2}\sum_{a\le x}\frac{1}{a^{1/2}} = O(x)$ (by PS or just comparing with the integral. Secondly,

$$\begin{aligned}x\sum_{a\le x}\frac{\mu(a)}{a} &= \sum_{a\le x}\mu(a)\lfloor\frac{x}{a}\rfloor + O(x)\\ &= \sum_{a\le x}\mu(a)\sum_{b\le\frac{x}{a}}1 + O(x)\\ &= \sum_{d\le x}\mu * 1(d) + O(x)\\ &= O(x)\end{aligned}$$

since the sum is either 1 (when $d = 1$) or 0 (otherwise).
Thirdly,[1]

$$\begin{aligned}\sum_{a\le x}\mu(a)\sum_{b\le x}\tau(b) &= \sum_{a\le x}\mu(a)\sum_{b\le \frac{x}{a}}\sum_{cd=b}1\\ &= \sum_{a\le x}\mu(a)\sum_{cd\le \frac{x}{a}}1\\ &= \sum_{acd\le x}\mu(a)\\ &= \sum_{d\le x}\sum_{ac\le \frac{x}{d}}\mu(a)\\ &= \sum_{d\le x}\sum_{e\le \frac{x}{d}}\mu*1(e)\\ &= \sum_{d\le x}1 = O(x)\end{aligned}$$

So

$$\begin{aligned}\sum_{n\le x}\Lambda_2(n) &= 2\sum_{a\le x}\mu(a)\sum_{b\le \frac{x}{a}}\frac{\tau(b)x}{ab}+O(x)\\ &= 2x\sum_{d\le x}\frac{1}{d}\mu*\tau(d)+O(x)\end{aligned}$$

Recall that $\tau = 1*1$, so $\mu*\tau = \mu*1*1 = 1$. So the above

$$\begin{aligned}&= 2x\sum_{d\le x}\frac{1}{d}+O(x)\\ &= 2x\log x+O(x)\end{aligned}$$

□

(Non-examinable from now, but lecturer still recommends us to think about it)
A 14-point plan to prove PNT from Selberg's identity:
Let $r(x) = \frac{\psi(x)}{x} - 1$, so PNT is equivalent to $\lim_{x\to\infty}|r(x)| = 0$.
1) Selberg's identity $\implies$

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

2) Considering 1) with x replaced $\frac{x}{m}$, summing over m, show

$$|r(x)|(\log x)^2 \le \sum_{n\le x}\frac{\Lambda_2(n)}{n}|r(\frac{x}{n})|+O(\log x)$$

[1] Jaspal noticed that this can be obtained much more easily by $\mu*\tau = 1$ from Möbius inversion.

3)

$$\sum_{n\leq x}\Lambda_2(n) = 2\int_1^{\lfloor x\rfloor}\log t dt + O(x)$$

4-6) (Let's skip some of the steps)

$$\sum_{n\leq x}\frac{\Lambda_2(n)}{n}|r(\frac{x}{n})| = 2\int_1^x \frac{|r(x/t)|}{t\log t}dt + O(\log x)$$

7) Let $V(u) = r(e^u)$. Show that

$$u^2|V(u)| \leq 2\int_0^u\int_0^v |V(t)|dtdv + O(u)$$

8) Show

$$\limsup|r(x)| \leq \limsup\frac{1}{u}\int_0^u |V(t)|dt = \beta$$

9-14) (!) If $\alpha > 0$, then can show from 7) that $\beta < \alpha$, contradiction; so $\alpha = 0$, and PNT.

2 Sieve Methods

—Lecture 6—

Hand in by Monday if you want some questions (q2 and q3) to be marked.

Everyone knows how Sieve of Eratosthenes works (some demonstration by lecturer). Our interest is in using the sieve to *count* things. If we apply Sieve of Eratosthenes to the interval $[1, 20]$, then we get an equality between two ways of counting how many numbers are left:

$$\pi(20) + 1 - \pi(\sqrt{20}) = 20 - \lfloor 20/2 \rfloor - \lfloor 20/3 \rfloor + \lfloor 20/6 \rfloor$$

where both sides evaluate to 7.

2.1 Setup

We'll have the following: • Finite set $A \subset \mathbb{N}$ (the set to be sifted);
• Set of primes p (the set of primes we sift out by), usually all primes;
• Sifting limit z (sift all primes in P less than z)
• sifting function

$$S(A, P; z) = \sum_{n \in A} 1_{(n, \prod_{p \in P, p<z} p)=1}$$

Let $\prod_{p \in P, p<z} p = P(z)$. Our goal is to estimate $S(A, P; z)$.

• For d, let

$$A_d = \{n \in A : d|n\}$$

• We write $|A_d| = \frac{f(d)}{d} X + R_d$ (most textbooks use ω in place of f here, our use here is to avoid confusion), where f is completely multiplicative ($f(mn) = f(m)f(n) \forall m, n$), and $0 \leq f(d) \forall d$.
• Note that $|A| = \frac{f(1)}{1} X + R_1 = X + R_1$;
• R_d is an 'error' term;
• We choose f so that $f(p) = 0$ if $p \notin P$ by convention (so in that case $R_p = |A_p|$).
• Let $W_p(z) = \prod_{p<z, p \in P} (1 - \frac{f(p)}{p})$.

Example. 1) Take $A = (x, x + y] \cap \mathbb{N}$, P the set of all primes. So $|A_d| = \lfloor \frac{x+y}{d} \rfloor - \lfloor \frac{x}{d} \rfloor = \frac{y}{d} + O(1)$.
Here $f(d) \equiv 1$, and $R_d = O(1)$.
So $S(A, P; z) = |\{x < n \leq x + y : p|n \implies p \geq z\}|$.
e.g. if $z \approx (x + y)^{1/2}$, then

$$S(A, P; z) = \pi(x + y) - \pi(x) + O((x + y)^{1/2})$$

2) $A = \{1 \le n \le y : n \equiv a \pmod q\}$, $A_d = \{1 \le m \le \frac{x}{d} : dm \equiv a \pmod q\}$. This congruence only has solutions if $(d,q)|a$. So
$|A_d| = \frac{(d,q)}{dq}y + O((d,q))$ if $(d,q)|a$, and $= O((d,q))$ otherwise.
So here $X = y/q$, and $f(d) = d(d,q)$ if $(d,q)|a$, and 0 otherwise.

3) How about twin primes, i.e. $p, p+2$ both primes? We have
$A = \{n(n+2) : 1 \le n \le x\}$ (so if $p|n(n+2) \iff n \equiv 0, -2 \pmod p$);
P is all primes except 2;
$|A_p| = \frac{2x}{p} + O(1)$ (so $f(p) = 2$). So $f(d) = 2^{\omega(d)}$ for f to be completely multiplicative, where $\omega(d)$ denote the number of primes divisors of d.
$S(A, P; x^{1/2}) = |\{1 \le p \le x : p, p+2 \text{ both prime}\} + O(x^{1/2})$. Denote the main term as $\pi_2(x)$, then as mentioned in the first lecture we expect $\pi_2(x) \approx \frac{x}{(\log x)^2}$. We will prove the upper bound using sieves.

Theorem. (1, Sieve of Eratosthenes Legendre)
$S(A, P; z) = XW_p(z) + O(\sum_{d|p(z)} R_d)$.

Proof.

$$\begin{aligned}
S(A,P;z) &= \sum_{n\in A} 1_{(n,p(z))=1} \\
&= \sum_{n\in A} \sum_{d|(n,p(z))} \mu(d) \\
&= \sum_{n\in A} \sum_{d|n, d|p(z)} \mu(d) \\
&= \sum_{d|p(z)} \mu(d) \sum_{n\in A} 1_{d|n} \\
&= \sum_{d|p(z)} \mu(d)|A_d| \\
&= X \sum_{d|p(z)} \frac{\mu(d)f(d)}{d} + \sum_{d|p(z)} \mu(d)R_d \\
&= X \prod_{p\in P, p<z} \left(1 - \frac{f(p)}{p}\right) + O(\sum_{d|p(z)} |R_d|)
\end{aligned}$$

□

Corollary. $\pi(x+y) - \pi(x) \ll \frac{y}{\log\log y}$.

Proof. In example 1, $f \equiv 1$, and $|R_d| \ll 1$, and $X = y$. So

$$W_p(z) = \prod_{p\le z} \left(1 - \frac{1}{p}\right) \ll (\log z)^{-1}$$

and

$$\sum_{d|p(z)} |R_d| \ll \sum_{d|p(z)} 1 \le 2^z$$

So $\pi(x+y) - \pi(x) \ll \frac{y}{\log z} + 2^z \ll \frac{y}{\log\log y}$ if we choose $z = \log y$.