Distribution of primes

Thursday, April 7, 2016 9:40 AM

Understanding how primes are distributed in integers is an important problem and various results in number theory are dealing with such questions.

Next we would like to understand the asymptotic behavior of

$$\sum_{p\le x} \frac{1}{p}.$$

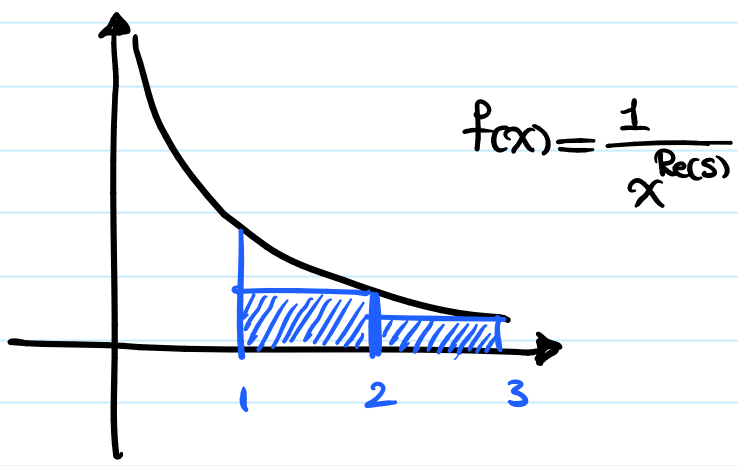
To warm up and mention some properties of absolutely convergent series, let's start with the following:

Proposition. For $\mathrm{Re}(s)>1$, $\sum_{n=1}^{\infty} \frac{1}{n^s}$ is absolutely convergence and $\prod_{p\in\mathcal{P}} \left(1-\frac{1}{p^s}\right)^{-1}$ is convergence and

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p\in\mathcal{P}} \left(1-\frac{1}{p^s}\right)^{-1}.$$

Proof. $\left|\frac{1}{n^s}\right| = \frac{1}{n^{\mathrm{Re}(s)}}$.

$$\sum_{n=1}^{\infty} \frac{1}{n^{\mathrm{Re}(s)}} \le 1+\int_1^{\infty} \frac{dx}{x^{\mathrm{Re}(s)}}$$



$$= 1+\lim_{T\to\infty} \mathrm{Re}(s)\left(1-\frac{1}{T^{\mathrm{Re}(s)-1}}\right) < \infty.$$

Let $M_x := \{ n\in\mathbb{Z}^+ \mid v_p(n)\neq 0 \Rightarrow p\le x\}$. Since $\sum_{n\in\mathbb{Z}^+} \frac{1}{n^s}$ is absolutely convergent, so is $\sum_{n\in M_x} \frac{1}{n^s}$.

Claim. Let p_k be the k^{th} prime number; then

$$\sum_{n\in M_{p_k}} \frac{1}{n^s} = \prod_{i=1}^{k} \left(1-\frac{1}{p_i^s}\right)^{-1}.$$

Pf of claim. We proceed by induction on k.

Base. $k=1$. $p_1=2$ and $M_2=\{1,2,2^2,\dots\} \Rightarrow \sum_{n\in M} \frac{1}{n^s} = 1+\frac{1}{2^s}+\frac{1}{2^{2s}}+\cdots$

Base. $k=1$. $p_1 = 2$ and $M_2 = \{1, 2, 2^2, \dots\} \Rightarrow \sum_{n \in M_2} \frac{1}{n^s} = 1 + \frac{1}{2^s} + \frac{1}{2^{2s}} \dots$

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

Inductive Step. Let's consider the following partition of $M_{p_{k+1}}$:

$$A_i := \{ m \in M_{p_{k+1}} \mid v_{p_{k+1}}(m) = i \}.$$

Since $\sum_{n \in M_{p_{k+1}}} \frac{1}{n^s}$ is absolutely convergent, we have

$$\sum_{n \in M_{p_{k+1}}} \frac{1}{n^s} = \sum_{i=0}^{\infty} \sum_{n \in A_i} \frac{1}{n^s} \underbrace{=}_{\boxed{\text{unique factorization}}} \sum_{i=0}^{\infty} \sum_{m \in M_{p_k}} \frac{1}{(m \cdot p_{k+1}^{i})^s}$$

$$= \sum_{i=0}^{\infty} \frac{1}{p_{k+1}^{is}} \left(\sum_{m \in M_{p_k}} \frac{1}{m^s} \right) = \left(\sum_{m \in M_{p_k}} \frac{1}{m^s} \right) \left(\sum_{i=0}^{\infty} \frac{1}{p_{k+1}^{is}} \right)$$

$$\underbrace{=}_{\boxed{\text{induction hypoth.}}} \left[\prod_{i=1}^{k} \left(1 - \frac{1}{p_i^s}\right)^{-1} \right] \left(1 - \frac{1}{p_{k+1}^s}\right)^{-1}. \qquad \blacksquare$$

$$\left| \zeta(s) - \sum_{n \in M_x} \frac{1}{n^s} \right| = \left| \sum_{n \notin M_x} \frac{1}{n^s} \right| \leq \sum_{n \notin M_x} \frac{1}{n^{\mathrm{Re}(s)}} \underset{\boxed{\text{It is clear that } \mathbb{Z} \cap [1, x] \subseteq M_x}}{\leq} \sum_{n \geq x} \frac{1}{n^{\mathrm{Re}(s)}}.$$

Since $\sum_{n \in \mathbb{Z}^+} \frac{1}{n^s}$ is absolutely convergent, $\forall \varepsilon > 0 \ \exists \ x = x(\varepsilon)$ s.t. $\sum_{n \geq x} 1/|n^s| < \varepsilon$. Hence $\left| \zeta(s) - \prod_{p \leq x} (1 - 1/p^s)^{-1} \right| \leq \varepsilon$.

Therefore $\prod_{p \in \mathcal{P}} (1 - 1/p^s)^{-1}$ is convergent and $\prod_{p \in \mathcal{P}} (1 - 1/p^s)^{-1} = \zeta(s)$. $\blacksquare$

Euler used the above equality to prove that there are infinitely many primes.

$$\lim_{s \to 1^+} \zeta(s) = \infty \Rightarrow \lim_{s \to 1^+} \prod_{p \in \mathcal{P}} (1 - 1/p^s)^{-1} = \infty.$$

$\Rightarrow |\mathcal{P}| = \infty$ as otherwise we would have had

$$\lim_{s \to 1^+} \prod_{p \in \mathcal{P}} (1 - 1/p^s)^{-1} = \prod_{p \in \mathcal{P}} (1 - 1/p)^{-1}.$$

$\wr$

This spirit of this argument can give us more. Let's start with

how fast $\sum_{n\le x} \frac{1}{n}$ is going to infinity.

Proposition. For some number $0<\gamma<1$, we have

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

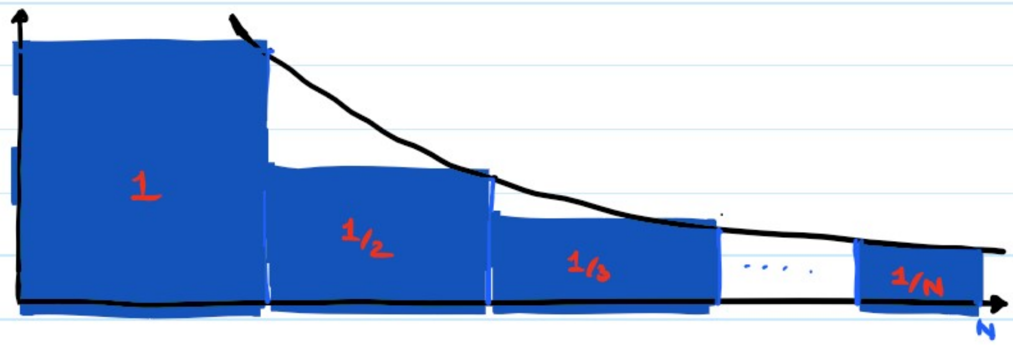
I.e. $\lim_{x\to\infty} \sum_{n\le x} \frac{1}{n} - \ln x = \gamma$ and

moreover $\left|\sum_{n\le x} \frac{1}{n} - \ln x - \gamma\right| \ll \frac{1}{x}$.

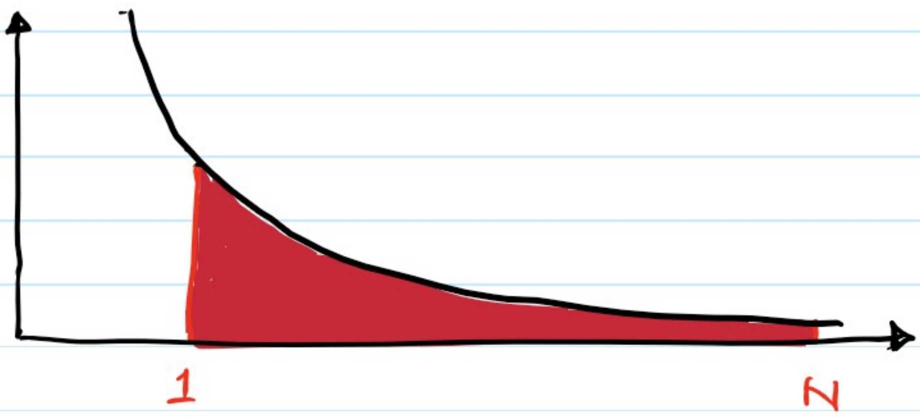
(This constant γ is called Euler's Constant.).

Proof. Let's visualize a few quantities.

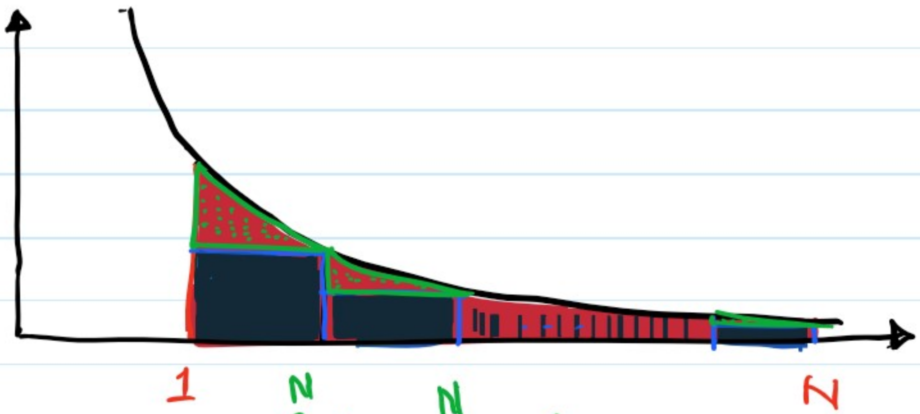
$$\sum_{k=1}^{N} \frac{1}{k}$$



$$\int_1^N \frac{dt}{t}$$

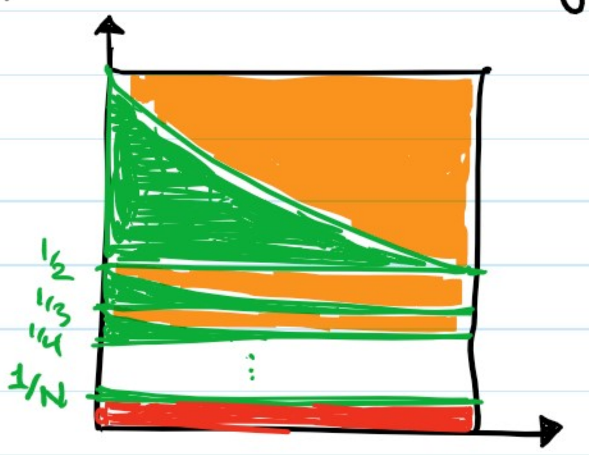


$$\int_1^N \frac{dt}{t} - \sum_{k=2}^{N} \frac{1}{k}$$



To visualize $\displaystyle\sum_{k=1}^{N} \frac{1}{k} - \int_1^N \frac{dt}{t} = 1 - \left(\int_1^N \frac{dt}{t} - \sum_{k=2}^{N} \frac{1}{k}\right)$,

we draw a unit square and shift the green regions into it:



So we get the "orange" region + the red rectangle. We notice that the orange region gets larger as $N \to \infty$, but its area is always ≤ 1. So its area converges to a number γ. Moreover the difference of the limiting area and the area of the orange region at the N^{th} step is at most the area of the red rectangle $= 1/N$. We are done.

More precisely: area of the orange region at the N^{th} step is $\displaystyle\sum_{n=1}^{N} \frac{1}{n} - \ln N - \frac{1}{N}$ which is <u>increasing</u> and converging to γ. So

$$0 < \sum_{n=1}^{N} \frac{1}{n} - \ln N - \gamma < \frac{1}{N}. \qquad \blacksquare$$

(from picture)

<u>Proposition</u> $\displaystyle\prod_{p \leq x} \left(1 - \frac{1}{p}\right)^{-1} > \log x.$

<u>Proof.</u>

$$\prod_{p\leq x}\left(1-\frac{1}{p}\right)^{-1} = \prod_{p\leq x}\left(1+\frac{1}{p}+\frac{1}{p^2}+\cdots\right)$$

$$> \prod_{p\leq x}\left(1+\frac{1}{p}+\frac{1}{p^2}+\cdots+\frac{1}{p^{k_p}}\right) \quad (p^{k_p} \leq x < p^{k_p+1}).$$

$$> \sum_{n \leq x} \frac{1}{n} \quad \text{(by unique factorization)}$$

$$\geq \log x. \qquad \blacksquare$$

<u>Corollary.</u> $\displaystyle\sum_{p\leq x} \frac{1}{p-1} > \log\log x$ and $\displaystyle\sum_{p \leq x} \frac{1}{p} > \log\log x - 1.$

<u>Proof.</u> $\displaystyle\log\left(\prod_{p\leq x}\left(1-\frac{1}{p}\right)^{-1}\right) > \log\log x$

$$\Longrightarrow \sum_{p\leq x} -\log\left(1-\frac{1}{p}\right) > \log\log x.$$

<u>Recall</u>. $\displaystyle\frac{f(a)-f(b)}{a-b} = f'(c)$ for some $c \in (a,b)$.

Recall. $\dfrac{\quad\cdots\quad}{a-b} = f'(c)$ for some $c \in (a,b)$.

apply this to $\underline{\log}$, $a = 1$, $b = 1-y$:

$$\frac{\log 1 - \log(1-y)}{y} = \frac{1}{c} \quad \text{for some} \quad 1-y < c < 1$$

$$\Longrightarrow \quad -\log(1-y) = \frac{y}{c} \quad \text{for some} \quad 1-y < c < 1$$

$$\Longrightarrow \quad y < -\log(1-y) < \frac{y}{1-y}$$

$$\Longrightarrow \quad \sum_{p \le x} \frac{1/p}{1-1/p} > \log\log x$$

$$\Longrightarrow \quad \log\log x < \sum_{p \le x} \frac{1}{p-1}.$$

Suppose $p_1 < p_2 < \cdots < p_k \le x$ is the list of primes at most x.

$$\sum_{i=1}^{k} \frac{1}{p_i} \ge \sum_{i=1}^{k} \frac{1}{p_{i+1}-1} \ge \sum_{3 \le p \le x} \frac{1}{p-1} > \log\log x - 1. \qquad \blacksquare$$

Alternative way to get the needed inequality.

$$-\log(1-y) = y + \frac{y^2}{2} + \frac{y^3}{3} + \cdots \quad \text{for any} \quad 0 < y < 1.$$

$$< y + y^2 + y^3 + \cdots = \frac{y}{1-y}.$$

Later we will see much better estimates, but let's see what this estimate tells us about the number of primes $\le x$.

Definition. $\pi(x) := |\mathcal{P} \cap [1, x]|$.

Prime number theorem. $\displaystyle\lim_{x \to \infty} \frac{\pi(x)}{x / \ln x} = 1.$

(Hadamard, de la Valleé Poussin) independently '1896

Legendre made a conjecture on how $\pi(x)$ is growing.

$\mapsto \dfrac{x}{\ln \qquad}$ and $A(x) \to 1.08\ldots$

$\longmapsto \dfrac{x}{\ln x - A(x)}$ and $A(x) \longrightarrow 1.08\ldots$

which is not quite true.

· Gauss said $li(x) = \displaystyle\int_2^x \frac{dt}{\ln t}$ is a better estimate. (It is called logarithmic integral.)

Exercise. $li(x) = \dfrac{x}{\ln x} + 1!\,\dfrac{x}{(\ln x)^2} + 2!\,\dfrac{x}{(\ln x)^3} + \cdots + q!\,\dfrac{x}{(\ln x)^q}\,(1+\varepsilon(x))$

where $\varepsilon(x) \longrightarrow 0$ as $x \longrightarrow \infty$.

Chebyshev 1851 proved the first result about growth of $\pi(x)$.

$$\underline{\lim}\, \frac{\pi(x)}{li\, x} \le 1 \le \overline{\lim}\, \frac{\pi(x)}{li\, x}$$

(In particular, if limit exists, then it is 1.)

Let's start with a rather easy upper bound.

Let $p_1 < p_2 < \cdots < p_r$ be the list of first r primes. And let

$A_d(x) = \{ m \in [1,x] \text{ s.t. } d \mid m \}$. Then

$$\mathcal{P} \cap [1,x] \subseteq \{p_1, \ldots, p_r\} \cup \Big(A_1(x) \setminus \bigcup_{i=1}^{r} A_{p_i}(x)\Big).$$

$\Longrightarrow$ By inclusion-exclusion

$$\pi(x) \le r + |A_1(x)| - \sum |A_{p_i}(x)| + \sum |A_{p_{i_1}p_{i_2}}(x)| - \cdots$$

Definition. The floor of y or the integer part of y is

$$\lfloor y \rfloor = \max\{ k \in \mathbb{Z} \mid k \le y \}.$$

Ex. $\lfloor 1.7 \rfloor = 1$ and $\lfloor -1.7 \rfloor = -2$

Basic Property. $\lfloor y \rfloor$ is the unique integer s.t.

$$\lfloor y \rfloor \le y < \lfloor y \rfloor + 1.$$

· $|A_d(x)| = \lfloor \frac{x}{d} \rfloor$.

$$\Rightarrow \pi(x) \le r + \lfloor x \rfloor - \sum \Big\lfloor \frac{x}{p_i} \Big\rfloor + \sum \Big\lfloor \frac{x}{p_{i_1}p_{i_2}} \Big\rfloor - \cdots$$

$$\Rightarrow \pi(x) \le r + \lfloor x \rfloor - \sum \left\lfloor \frac{x}{p_i} \right\rfloor + \sum \left\lfloor \frac{x}{p_{i_1} p_{i_2}} \right\rfloor$$

$$\begin{aligned}
&\le r + x - \sum \frac{x}{p_i} + \sum \frac{x}{p_{i_1} p_{i_2}} - \cdots \\
&\qquad + 1 + r + \binom{r}{2} + \cdots + \binom{r}{r} \\
&= r + 2^r + x \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right) \\
&\le 2^{r+1} + \frac{x}{\ln p_r} \\
&< 2^{r+1} + \frac{x}{\ln r}
\end{aligned}$$

This technique is called sieve. A fundamental combinatorial sieve was developed by Brun.

Let $r = \lfloor \ln x \rfloor + 1 \implies \pi(x) < 4 \cdot x^{\ln 2} + \dfrac{x}{\ln \ln x}$.

$$\Rightarrow \pi(x) \ll \frac{x}{\ln \ln x} \quad \text{as} \quad \ln 2 < 1.$$

Let's follow Chebyshev's idea to get a baby prime number theorem.

Suppose we would like to understand $\pi(2n) - \pi(n)$.

Let's list all the primes in this range:

$$n < p_1 < p_2 < \cdots < p_r \le 2n.$$

$$\Rightarrow \quad p_1 \cdot p_2 \cdots p_r \;\Big|\; \frac{(2n)!}{n!\,n!} = \binom{2n}{n}.$$

What can we say about power of p in $\binom{2n}{n}$?

$$v_p\binom{2n}{n} = v_p((2n)!) - 2\, v_p(n!).$$

Proposition. $v_p(m!) = \left\lfloor \frac{m}{p} \right\rfloor + \left\lfloor \frac{m}{p^2} \right\rfloor + \cdots$

Proof.

$$\begin{aligned}
v_p(m!) &= \sum_{i=1}^{m} v_p(m) = \sum_{j=1}^{\infty} j \,\big|\{k \in [1..m] \mid v_p(k) = j\}\big| \\
&= \sum_{j=1}^{\infty} \Big(\sum_{i \le j} 1 \Big) \big|\{k \in [1..m] \mid v_p(k) = j\}\big|
\end{aligned}$$

$$= \sum_{1\le i\le j} |\{k \in [1..m] \mid v_p(k) = j\}|$$

$$= \sum_{i=1}^{\infty} |\{k \in [1..m] \mid v_p(k) \ge i\}|$$

$$= \sum_{i=1}^{\infty} |A_{p^i}(m)| = \sum_{i=1}^{\infty} \left\lfloor \frac{m}{p^i} \right\rfloor . \qquad \blacksquare$$

Hence $v_p\binom{2n}{n} = \sum_{i=1}^{\infty} \left(\left\lfloor \frac{2n}{p^i} \right\rfloor - 2\left\lfloor \frac{n}{p^i} \right\rfloor \right)$

(Ex. $\lfloor 2x \rfloor = \lfloor x \rfloor + \lfloor x + 1/2 \rfloor$)

$$= \sum_{i=1}^{\infty} \left(\left\lfloor \frac{n}{p^i} + \frac{1}{2} \right\rfloor - \left\lfloor \frac{n}{p^i} \right\rfloor \right)$$

$$\le \sum_{1\le i\le \lfloor \log_p(2n) \rfloor} 1 = \lfloor \log_p(2n) \rfloor .$$

$$\Rightarrow \binom{2n}{n} \,\Big|\, \prod_{p\le 2n} p^{\lfloor \log_p(2n) \rfloor} .$$

$$\Rightarrow n^{\pi(2n)-\pi(n)} < \prod_{n<p\le 2n} p \le \binom{2n}{n} \le \prod_{p\le 2n} p^{\lfloor \log_p(2n) \rfloor}$$

$$\le (2n)^{\pi(2n)} .$$

$$\Rightarrow (\pi(2n)-\pi(n)) \ln(n) \le \ln\binom{2n}{n} \le \pi(2n) \ln(2n)$$

- $\binom{2n}{n} \le 2^{2n} \Rightarrow \quad \pi(2n) - \pi(n) \le 2\ln(2) \cdot \dfrac{n}{\ln n} \quad (*)$

- $\binom{2n}{n} = \dfrac{(2n)(2n-1)\cdots(n+1)}{(n)(n-1)\cdots(1)} \ge 2^n \Rightarrow$

$$\ln(2) \cdot \frac{n}{\ln(2n)} \le \pi(2n) .$$

$$\Rightarrow \boxed{\pi(x) \ge \pi\left(2\left\lfloor \frac{x}{2} \right\rfloor\right) \gg \frac{x}{\ln x} .}$$

$$\ln x$$

- $\pi(x) - \pi\left(\frac{x}{2}\right) = \pi(x) - \pi\left(\lfloor \frac{x}{2} \rfloor\right)$

$$\leq 1 + \pi\left(2\lfloor \tfrac{x}{2} \rfloor\right) - \pi\left(\lfloor \tfrac{x}{2} \rfloor\right)$$

$$\ll \frac{x}{\ln x}. \qquad \text{(because of } \circledast\text{)}$$

$$\Rightarrow \ln x \cdot \pi(x) - \ln x \cdot \pi\left(\frac{x}{2}\right) \ll x$$

$$\Rightarrow \ln x \cdot \pi(x) - \ln \frac{x}{2} \cdot \pi\left(\frac{x}{2}\right) \ll x + \ln(2) \cdot \pi\left(\frac{x}{2}\right)$$

$$\ll x.$$

$$\Rightarrow \sum_{k=0}^{\lfloor \log_2 x \rfloor} \left[\ln\left(\frac{x}{2^k}\right)\pi\left(\frac{x}{2^k}\right) - \ln\left(\frac{x}{2^{k+1}}\right)\pi\left(\frac{x}{2^{k+1}}\right)\right] \ll \sum_{k=0}^{\lfloor \log_2 x \rfloor} \frac{x}{2^k}$$

$$\Rightarrow \ln(x)\,\pi(x) \ll x \Rightarrow \boxed{\pi(x) \ll \frac{x}{\ln x}}.$$

Let's summarize what we have proved:

$\boxed{\text{Theorem}}$ (Chebyshev) $\quad \dfrac{x}{\ln x} \ll \pi(x) \ll \dfrac{x}{\ln x}.$

Let's use $n!$ further to get an estimate on certain weighted sums over primes.

$$n! = \prod_{p \leq n} p^{v_p(n!)} \Rightarrow \ln(n!) = \sum_{p \leq n} v_p(n!) \ln p.$$

$$\Rightarrow \ln(n!) = \sum_{p \leq n} \left(\sum_{k=1}^{\lfloor \log_p n \rfloor} \left\lfloor \frac{n}{p^k} \right\rfloor\right) \ln p$$

$$\Rightarrow \left| \ln(n!) - \sum_{\substack{p \leq n \\ k < \log n}} \frac{n}{p^k} \ln p \right| \leq \sum_{p \leq n} \ln p \leq \ln(n) \cdot \pi(n)$$

$k \le \log_p n$

$$\Rightarrow \left| \frac{1}{n}\ln(n!) - \sum_{p\le n} \frac{\ln p}{p} \right| \le \frac{\ln(n)}{n}\pi(n) + \sum_{p\le n} \ln(p)\cdot\left(\frac{1}{p^2}+\frac{1}{p^3}+\cdots\right)$$

$$\overset{\boxed{\text{By Chebyshev}}}{\le} C + \sum_{p\le n} \frac{\ln(p)}{p^2}\cdot\left(\frac{1}{1-1/p}\right)$$

Ex. $\displaystyle\sum_{k=2}^{\infty} \frac{\ln(k)}{k(k-1)}$ is convergent.

$$\Rightarrow \left| \frac{1}{n}\ln(n!) - \sum_{p\le n} \frac{\ln p}{p} \right| \ll 1.$$

$$\Rightarrow \boxed{\sum_{p\le n} \frac{\ln p}{p} = \frac{1}{n}\ln(n!) + O(1)}$$

We need to understand the rate of growth of $\frac{1}{n}\ln(n!)$.

Lemma. $\frac{1}{n}\ln(n!) = \ln(n) + O(1)$.

Proof. $$\frac{1}{n}\ln(n!) - \ln(n) = \frac{1}{n}\sum_{k=1}^{n}(\ln k - \ln n)$$

$$= \underbrace{\frac{1}{n}\sum_{k=1}^{n}\ln\left(\frac{k}{n}\right)}_{\text{Riemann partial sum}} \xrightarrow[n\to\infty]{} \int_0^1 \ln(t)\,dt$$

$$\int_0^1 \ln(t)\,dt = t\cdot\ln t\Big|_0^1 - \int_0^1 dt$$

$$\boxed{\begin{aligned} du &= dt \Rightarrow u = t \\ v &= \ln t \Rightarrow dv = \frac{dt}{t} \end{aligned}}$$

$$\lim_{t\to 0^+} t\ln t \underset{\boxed{t=1/x}}{=} \lim_{x\to\infty} \frac{-\ln x}{x} \underset{\boxed{\text{l'Hopital}}}{=} \lim_{x\to\infty}\frac{-1/x}{1} = 0.$$

$$\Rightarrow \int_0^1 \ln(t)\,dt = -1. \quad \blacksquare$$

Overall we get the following:

Overall we get the following:

Proposition. $\displaystyle\sum_{p\le n} \frac{\ln p}{p} = \ln(n) + O(1).$

Next we will see how the above Proposition can help us to understand the true rate of growth of $\displaystyle\sum_{p\le x} \frac{1}{p}$.

Lemma (Integration-by-part)

- $\lambda_1 < \lambda_2 < \cdots$; $\lambda_n \to +\infty$ as $n \to +\infty$;
- $c_1, c_2, \cdots \in \mathbb{C}$ and $C(x) = \displaystyle\sum_{\lambda_n \le x} c_n$.
- $f(t)$ has continuous derivative for $t \ge \lambda_1$.

$$\Rightarrow \quad \sum_{\lambda_n \le x} c_n f(\lambda_n) = C(x)f(x) - \int_{\lambda_1}^{x} C(t)\, f'(t)\, dt.$$

Proof. Since $\lambda_n > \lambda_{n-1}$, $C(\lambda_n) - C(\lambda_{n-1}) = c_n$.

$$\Rightarrow \sum_{\lambda_n \le x} c_n f(\lambda_n) = C(\lambda_1)f(\lambda_1) + (C(\lambda_2) - C(\lambda_1))f(\lambda_2) + (C(\lambda_3) - C(\lambda_2))f(\lambda_3) + \cdots + (C(\lambda_{n_0}) - C(\lambda_{n_0-1}))f(\lambda_{n_0})$$

(where n_0 is the largest integer s.t. $\lambda_{n_0} \le x$.)

$$= -C(\lambda_1)(f(\lambda_2) - f(\lambda_1)) - C(\lambda_2)(f(\lambda_3) - f(\lambda_2)) - \cdots - C(\lambda_{n_0-1})(f(\lambda_{n_0}) - f(\lambda_{n_0-1})) + C(\lambda_{n_0})f(\lambda_{n_0})$$

$$= C(x)f(x) - C(\lambda_{n_0})(f(x) - f(\lambda_{n_0})) - \sum_{i=1}^{n_0-1} C(\lambda_i)(f(\lambda_{i+1}) - f(\lambda_i))$$

$$= C(x)f(x) - C(\lambda_{n_0}) \int_{\lambda_{n_0}}^{x} f'(t)\, dt - \sum_{i=1}^{n_0-1} C(\lambda_i) \int_{\lambda_i}^{\lambda_{i+1}} f'(t)\, dt$$

$$\left\{ \cdot\ C(t) = C(\lambda_i) \quad \text{for } \lambda_i \le t < \lambda_{i+1} \cdot \right\}$$

$$= C(x) f(x) - \int_{\lambda_1}^{x} C(t) f'(t)\, dt. \qquad \blacksquare$$

$$\sum_{p \le x} \frac{1}{p} = \sum_{p \le x} \frac{\ln p}{p} \cdot \frac{1}{\ln p} = C(x) \frac{1}{\ln x} - \int_2^x C(t) \cdot \frac{-1}{t (\ln t)^2}\, dt$$

$$\boxed{\text{Let } C(x) = \sum_{p \le x} \frac{\ln p}{p} \text{ and } f(t) = \frac{1}{\ln t}.}$$

$$= 1 + O\left(\frac{1}{\ln x}\right) + \int_2^x \ln(t) \cdot \frac{1}{t (\ln t)^2}\, dt$$

$$+ O\left(\int_2^x \frac{1}{t (\ln t)^2}\, dt \right)$$

$$= 1 + O\left(\frac{1}{\ln x}\right) + \int_2^x \frac{dt}{(\ln t) \cdot t} + O\left(\int_2^x \frac{1}{(\ln t)^2} \cdot \frac{dt}{t} \right)$$

$$= 1 + O\left(\frac{1}{\ln x}\right) + \int_{\ln 2}^{\ln x} \frac{du}{u} + O\left(\int_{\ln 2}^{\ln x} \frac{du}{u^2} \right)$$

$$= 1 + O\left(\frac{1}{\ln x}\right) + \ln \ln x - \ln \ln 2$$
$$+ O\left(\frac{1}{\ln x} - \frac{1}{\ln 2} \right)$$

$$= \ln \ln x + c + O\left(\frac{1}{\ln x}\right) \quad \text{for some constant } c.$$

So we get

<u>Theorem</u>. $\displaystyle\sum_{p \le x} \frac{1}{p} = \ln \ln x + C + O\left(\frac{1}{\ln x}\right)$

for some real number C.

Let's go back to $\displaystyle\sum_{p \le x} \frac{\ln p}{p} = \ln x + O(1)$.

$\Rightarrow \exists\, c > 0$ s.t. $\displaystyle\ln x - c < \sum_{p \le x} \frac{\ln p}{p} < \ln x + c$ $\Rightarrow$

$$\ln x - c < \sum_{p \le x} \frac{\ln p}{p} < \ln x + c$$

$$\rightarrow \quad \ln x + c = \ln\left(e^{2c} x\right) - c < \sum_{p \le e^{2c} x} \frac{\ln p}{p}$$

$$\sum_{p \le x} \frac{\ln p}{p} < \sum_{p \le e^{2c} x} \frac{\ln p}{p}$$

$$\Rightarrow \quad \boxed{\forall x,\ \exists p : \text{prime s.t. } \quad x < p \le e^{2c} x.}$$

__Theorem__ (Bertrand's hypothesis)

$$\forall x > 1,\ \exists p : \text{prime} \qquad x \le p \le 2x.$$

__Lemma.__ $\displaystyle\prod_{p \le n} p < 4^n.$

__Proof.__ We proceed by strong induction on n. The base case is clear.

__Strong induction step.__ $\displaystyle \forall\, 1 \le k < n,\ \prod_{p \le k} p < 4^k \overset{?}{\Longrightarrow} \prod_{p \le n} p < 4^n.$

If n is even and $n \ne 2$, then

$$\prod_{p \le n} p \underset{\boxed{n \text{ is not prime}}}{=} \prod_{p \le n-1} p \underset{\boxed{\text{Strong induction hypothesis}}}{<} 4^{n-1} < 4^n$$

If $n = 2$, then $\displaystyle\prod_{p \le 2} p = 2 < 4^2.$

If $n = 2m+1$ for some positive integer m, then

$$\prod_{p \le 2m+1} p = \prod_{p \le m+1} p \cdot \prod_{m+1 < p \le 2m+1} p$$

Strong induction hypothesis & $\displaystyle\prod_{m+1 < p \le 2m} p \,\Big|\, \binom{2m+1}{m}$

$$< 4^{m+1} \cdot \binom{2m+1}{m}$$

$$< 4^{m+1} \cdot \underbrace{2^{2m+1} / 2}_{\blacktriangle} = 4^{2m+1}. \qquad \blacksquare$$

$m+1 \le 2m \ldots$

$$\boxed{\begin{aligned} 2^{2m+1} &= \binom{2m+1}{0}+\binom{2m+1}{1}+\cdots+\underbrace{\binom{2m+1}{m}+\binom{2m+1}{m+1}}+\cdots+\binom{2m+1}{2m+1} \\ &\ge 2\binom{2m+1}{m} \end{aligned}}$$

Proof of Bertrand's hypothesis.

Suppose to the contrary that $\exists x \ge 1$ s.t. there is no prime between x and $2x$. Let $n = \lfloor x \rfloor$. Hence there is no prime p s.t. $n < p \le 2n$.

$$\Rightarrow \ v_p\binom{2n}{n} = 0 \quad \text{if} \quad n < p \le 2n.$$

$$v_p\binom{2n}{n} = \sum_{k=1}^{\infty}\left(\left\lfloor \frac{2n}{p^k}\right\rfloor - 2\left\lfloor \frac{n}{p^k}\right\rfloor\right).$$

$$\left\lfloor \frac{2n}{p^2}\right\rfloor = 0 \iff 2n < p^2 \iff \sqrt{2n} < p.$$

So, if $\sqrt{2n} < p \le n$, then
$$\begin{aligned} v_p\binom{2n}{n} &= \left\lfloor \frac{2n}{p}\right\rfloor - 2\left\lfloor \frac{n}{p}\right\rfloor \\ &= \left\lfloor \frac{n}{p}+\frac{1}{2}\right\rfloor - \left\lfloor \frac{n}{p}\right\rfloor. \end{aligned}$$

So, for $\sqrt{2n} < p \le n$,
$$v_p\binom{2n}{n} \ne 0 \iff \underbrace{\frac{n}{p} - \left\lfloor \frac{n}{p}\right\rfloor}_{\text{the fractional part of } \frac{n}{p}} \ge \frac{1}{2}$$
$$\Updownarrow$$
$$v_p\binom{2n}{n} = 1$$

In particular, if $1 \le \frac{n}{p} < \frac{3}{2}$, then $\left\lfloor \frac{n}{p}\right\rfloor = 1$ and $\frac{n}{p} - \left\lfloor \frac{n}{p}\right\rfloor < 1/2$.

Hence $\frac{2}{3}n < p \le n \Rightarrow p \nmid \binom{2n}{n}$.

As we have seen before, $v_p\binom{2n}{n} \le \lfloor \log_p(2n) \rfloor$.

Hence
$$\binom{2n}{n} = \prod_{p \le \frac{2}{3}n} p^{v_p\binom{2n}{n}} \le \prod_{p \le \sqrt{2n}} p^{v_p\binom{2n}{n}} \cdot \prod_{p \le \frac{2}{3}n} p$$

$\sqrt{}$

$$\le (2n)^{\sqrt{2n}} \cdot 4^{\frac{2}{3}n}$$

$$\left.\begin{matrix} \\ \binom{2n}{n}(2n+1) \ge \sum_{i=0}^{2n} \binom{2n}{i} = 2^{2n} \end{matrix}\right\} \Rightarrow$$

$$4^n \le (2n+1)\cdot(2n)^{\sqrt{2n}} \cdot 4^{\frac{2}{3}n}$$

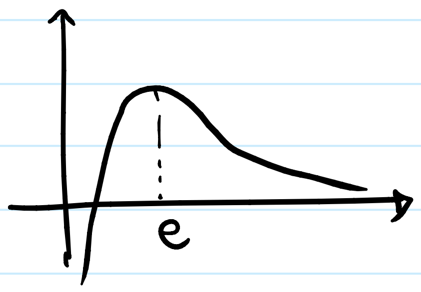
$$\Rightarrow \quad 4^{\frac{n}{3}} \le (2n+1)\cdot(2n)^{\sqrt{2n}} \le 2\cdot(2n)^{\sqrt{2n}+1}$$

$$\Rightarrow \quad \frac{\ln 4}{3}\, n \le \ln 2 + (\sqrt{2n}+1)\ln(2n).$$

$$\Rightarrow \quad \frac{\ln 4}{3} \le \frac{\ln 2}{n} + \left(\sqrt{2}+\frac{1}{\sqrt{n}}\right)\frac{\ln(2n)}{\sqrt{n}}$$

$$\Rightarrow \quad \frac{\ln 4}{3} \le \frac{\ln 2}{n} + \left(\sqrt{2}+\frac{1}{\sqrt{n}}\right)\cdot 2\sqrt{2}\cdot\frac{\ln(\sqrt{2n})}{\sqrt{2n}}$$

__Ex.__



is the graph of $\frac{\ln t}{t}$.

So, for $n \ge 2^9$, the RHS $\le \frac{\ln 2}{2^9} + \left(\sqrt{2} + \frac{1}{2^4\sqrt{2}}\right)\cdot 2\sqrt{2}\cdot\frac{\ln(2^5)}{2^5}$

$$= \frac{\ln 2}{2^9} + \left(4+\frac{1}{8}\right)\cdot\frac{5\ln 2}{2^5}$$

$$\Rightarrow \quad \frac{2}{3} \le \frac{1}{2^9} + \frac{33}{2^8}\cdot 5.$$

$$\Rightarrow \quad 2^{10} \le (3)(1+2\cdot 33\cdot 5) = (3)(331)$$

$\Rightarrow$ $1024 \le 991$ which is a contradiction

$\Rightarrow$ $n < 2^9$; one can check these finitely many integers do satisfy Bertrand's hypothesis. ■