Chebyshev's results on distribution of primes

Notation: $p =$ prime number, $n \geq 1$ integer, $x \in \mathbb{R}$

Def: $\pi(x) = \sum_{p \leq x} 1$, $\vartheta(x) = \sum_{p \leq x} \log p$, $\Lambda(n) = \begin{cases} \log p, & n = p^r \\ 0, & n \neq p^r \end{cases}$; $\psi(x) = \sum_{n \leq x} \Lambda(n)$

Note: $\psi(x) = \sum_{p^k \leq x} \log p = \sum_{p \leq x} \left\lfloor \frac{\log x}{\log p} \right\rfloor \log p = \sum_{m \geq 1} \vartheta(x^{1/m})$

Thm 1. $\exists$ explicit constants $0 < a < 1 < b$ such that, for large enough $x \geq x_0$,

$$ax \leq \vartheta(x) \leq bx, \quad ax \leq \psi(x) \leq bx, \quad a\frac{x}{\log x} \leq \pi(x) \leq b\frac{x}{\log x}.$$

Rmk. It is enough to treat one of the functions ϑ, ψ, π, because of:

Exercise. $\liminf_{x \to +\infty} \vartheta(x)/x = \liminf_{x \to +\infty} \psi(x)/x = \liminf_{x \to +\infty} \pi(x)/\left(\frac{x}{\log x}\right)$ $\left(\text{idem for } \limsup\right)$.

Cor of Thm 1. $\exists$ explicit constants $0 < A < B$ such that, for large enough $n \geq n_0$,

$$An \log n \leq p_n = \text{the } n\text{-th prime number} \leq Bn \log n.$$

Pf of Thm 1. One compares estimates on the size of $N = \binom{2n}{n} = \frac{(2n)!}{n!\,n!}$ with its prime factorisation. Naive estimates suffice:

$$N \leq \sum_{k=0}^{2n} \binom{2n}{k} = (1+1)^{2n} = 2^{2n} \leq (2n+1) \max_k \binom{2n}{k} = (2n+1)N.$$

$\underline{b = 4\log 2}$: if $n < p \leq 2n \implies p \mid N$, hence

$$\vartheta(2n) - \vartheta(n) = \sum_{n < p \leq 2n} \log p \leq \log N \leq 2n \log 2$$

$$\implies \vartheta(2^m) = \sum_{k=0}^{m-1} (\vartheta(2^{k+1}) - \vartheta(2^k)) \leq \left(\sum_{k=0}^{m-1} 2^{k+1}\right) \log 2 < 2^{m+1} \log 2.$$

If $2^{m-1} \leq x < 2^m \implies \underline{\vartheta(x) \leq \vartheta(2^m) < 2^{m+1}\log 2 \leq 4x\log 2}$

$\underline{a = \log 2}$: $v_p(n!) = \sum_{k \geq 1} \left\lfloor \frac{n}{p^k} \right\rfloor \implies v_p(N) = \sum_{1 \leq k \leq \left\lfloor \frac{\log 2n}{\log p} \right\rfloor} \underbrace{\left(\left\lfloor \frac{2n}{p^k} \right\rfloor - 2\left\lfloor \frac{n}{p^k} \right\rfloor\right)}_{0 \text{ or } 1} \leq \left\lfloor \frac{\log 2n}{\log p} \right\rfloor$,

$$\log N = \sum_p v_p(N) \log p \leq \sum_p \left\lfloor \frac{\log 2n}{\log p} \right\rfloor \log p = \psi(2n)$$

$$\implies \psi(2n) \geq \log(2^{2n}/(2n+1)) = 2n\log 2 - \log(2n+1) \implies \liminf_{x \to +\infty} \psi(x)/x \geq \log 2.$$

Alternative argument: write $f = O(g)$ if $\exists C > 0\ \exists x_0\ \forall x \geq x_0\ |f(x)| \leq C|g(x)|$

$L(x) := \sum_{n \leq x} \log n = \log(\lfloor x \rfloor !) = \sum_{p^k \leq x} \left\lfloor \frac{x}{p^k} \right\rfloor \log p = \sum_{n \leq x} \left\lfloor \frac{x}{n} \right\rfloor \Lambda(n) = \sum_{m \geq 1} \psi(x/m)$ satisfies

$L(x) = \int_1^x \log t \, dt + O(\log x) = x\log x - x + O(\log x)$, $\underbrace{L(x) - 2L\left(\frac{x}{2}\right)}_{x\log 2 + O(\log x)} = \sum_{m \geq 1} (-1)^{m-1} \psi\left(\frac{x}{m}\right)$

$\implies x\log 2 + O(\log x) \leq \psi(x)$ (ψ is non-decreasing) $\implies \underline{a = \log 2}$

$\psi(x) - \psi\left(\frac{x}{2}\right) \leq$

$\psi(x) = \sum_{m \geq 0} \left(\psi\left(\frac{x}{2^m}\right) - \psi\left(\frac{x}{2^{m+1}}\right)\right) \leq 2x\log 2 + O(\log x)$

$\implies \underline{b = 2\log 2}$.

Thm 2. $\forall n \geq 1 \ \exists$ prime p, $n < p \leq 2n$ ("Bertrand's postulate").

Pf (after S.S. Pillay). Better estimates for $N = \binom{2n}{n} = 2^{2n} P$, $P = \frac{(2n-1)!!}{(2n)!!}$:

$$\left.\begin{aligned} 1 &> \prod_{k=1}^{n}\left(1 - \frac{1}{(2k)^2}\right) = \frac{(2n+1)(2n-1)!!^2}{(2n)!!^2} = (2n+1)P^2 > 2nP^2 = \frac{2n}{2^{4n}} N^2 \\ 1 &> \prod_{k=2}^{n}\left(1 - \frac{1}{(2k-1)^2}\right) = \frac{(2n)!!^2}{4n\ (2n-1)!!^2} = \frac{1}{4nP^2} = \frac{2^{4n}}{4nN^2} \end{aligned}\right\} \Longrightarrow \frac{2^{2n}}{2\sqrt{n}} < N < \frac{2^{2n}}{\sqrt{2n}}$$

Inequality $\forall n \geq 1 \quad \vartheta(n) < 2n \log 2$: assume this holds for some $m \geq 2$.

$$N/2 = \binom{2n-1}{n-1} \geq \prod_{n < p \leq 2n-1} p \implies \log(N/2) \geq \vartheta(2n-1) - \vartheta(n) > \vartheta(2n-1) - 2n\log 2$$

$$\implies \vartheta(2n) = \vartheta(2n-1) < 2n\log 2 + \log(N/2) < (4n-1)\log 2 - \tfrac{1}{2}\log(2n) < 2(2n-1)\log 2 \implies \text{result by induction}$$

Proof of $\vartheta(2n) - \vartheta(n) > 0$ for $n \geq 2^6$:

$$\log N = \sum_{p \leq 2n} \nu_p(N) \log p, \quad \nu_p(N) = \sum_{k \geq 1}\left(\left\lfloor \frac{2n}{p^k}\right\rfloor - 2\left\lfloor \frac{n}{p^k}\right\rfloor\right).$$ Write $\log N = \Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_4$,

$$\Sigma_1 = \sum_{n < p \leq 2n}, \quad \underset{(n\geq 3)}{\Sigma_3} = \sum_{\sqrt{2n} < p \leq \frac{2n}{3}}, \quad \underset{(n \geq 3)}{\Sigma_2} = \sum_{\frac{2n}{3} < p \leq n}, \quad \Sigma_4 = \sum_{p \leq \sqrt{2n}}$$

$$\Sigma_1 = \vartheta(2n) - \vartheta(n), \quad \Sigma_2 = 0 \ (\text{if } n \geq 3), \quad \Sigma_3 \leq \sum_{\sqrt{2n} < p \leq \frac{2n}{3}} \log p = \vartheta\left(\frac{2n}{3}\right) - \vartheta(\sqrt{2n}) \underset{(n \geq 5)}{\leq} \vartheta\left(\frac{2n}{3}\right) - \pi(\sqrt{2n})\log 2$$

$$\Sigma_4 \leq \sum_{p \leq \sqrt{2n}} \left\lfloor \frac{\log 2n}{\log p}\right\rfloor \log p \leq (\log 2n)\, \pi(\sqrt{2n})$$

$$\left.\begin{aligned} &\implies \log N \leq \vartheta(2n) - \vartheta(n) + \vartheta\left(\tfrac{2n}{3}\right) - \pi(\sqrt{2n})(\log 2 - \log 2n) \\ &n \geq 2 \quad \vartheta\left(\tfrac{2n}{3}\right) = \vartheta\left(\left\lfloor\tfrac{2n}{3}\right\rfloor\right) < 2\left\lfloor \tfrac{2n}{3}\right\rfloor \log 2 \\ &\qquad \log N \geq 2n\log 2 - \log(2\sqrt{n}), \quad \pi(n) \leq \tfrac{n}{2} \quad (n \geq 8) \end{aligned}\right\} \Longrightarrow \begin{aligned} &\forall n \geq 2^5 \\ &\vartheta(2n) - \vartheta(n) > \\ &\underbrace{\left(\tfrac{2n}{3} - 1\right)\log 2 - \left(\tfrac{\sqrt{2n}+1}{2}\right)\log n}_{>0 \text{ for } n \geq 2^6} \end{aligned}$$

For $n < 2^6$: primes 2, 3, 5, 7, 13, 23, 43, 67

Chebychev's original proof used $M = \frac{n!(30n)!}{(6n)!(10n)!(15n)!}$ $\quad (1 + 30 = 6 + 10 + 15)$

$$\nu_p(M) = \sum_{p^k \leq 30n} f\left(\frac{30n}{p^k}\right), \quad f(a) = [a] - \left[\frac{a}{2}\right] - \left[\frac{a}{3}\right] - \left[\frac{a}{5}\right] + \left[\frac{a}{30}\right]$$

$f: \mathbb{R} \longrightarrow \{0, 1\}$, $f(a) = 1$ if $1 \leq a < 6$. As above,

$$L_f(x) := L(x) - L\left(\frac{x}{2}\right) - L\left(\frac{x}{3}\right) - L\left(\frac{x}{5}\right) + L\left(\frac{x}{30}\right) = \sum_{p^k \leq x} f\left(\frac{x}{p^k}\right)\log p = \sum_{n \leq x} f\left(\frac{x}{n}\right)\Lambda(n) =: \psi_f(x)$$

$$\left.\begin{aligned} &L_f(x) = \alpha x + O(\log x), \quad \alpha = \tfrac{1}{2}\log 2 + \tfrac{1}{3}\log 3 + \tfrac{1}{5}\log 5 - \tfrac{1}{30}\log(30) = 0.9212\ldots \\ &f(a) \in \{0,1\} \implies \psi_f(x) \leq \psi(x) \\ &f(a) = 1 \text{ if } 1 \leq a < 6 \implies \psi(x) - \psi\left(\tfrac{x}{6}\right) = \sum_{x/6 < n \leq x} \Lambda(n) = \sum_{x/6 < n \leq x} f\left(\tfrac{x}{n}\right)\Lambda(n) \leq \psi_f(x) \end{aligned}\right\} \Longrightarrow$$

$$\alpha x + O(\log x) \leq \psi(x) \leq \left(\sum_{m \geq 0} 6^{-m}\right)\alpha x + O(\log x) = \frac{6}{5}\alpha x + O(\log x)$$

So, $\forall \varepsilon > 0 \ \exists x_\varepsilon \ \forall x \geq x_\varepsilon \quad (\alpha - \varepsilon)x \leq \vartheta(x) \leq \left(\frac{6}{5}\alpha + \varepsilon\right)x \implies \vartheta(2x) - \vartheta(x) \geq$
$\geq \left(\frac{4}{5}\alpha - 3\varepsilon\right)x > 0.$

Explicit estimates of error terms $\implies$ result for $x < x_\varepsilon$

Thm 3. If $c = \lim\limits_{x \to +\infty} \psi(x)/x$ exists $\left(\implies c = \lim\limits_{x \to +\infty} \vartheta(x)/x = \lim\limits_{x \to +\infty} \pi(x)\Big/\frac{x}{\log x}\right) \implies c = 1$.

Pf: $x \log 2 + O(\log x) = \sum_{m \geq 1} (-1)^{m-1}\psi(x/m) \implies \log 2 = c\sum_{m \geq 1} (-1)^{m-1}/m = c \log 2$.