

Understanding PollPom

gavinguinn1

June 2020

The goal of this document is to break the latter half of the Pollack/Pomerance argument into digestible pieces with explanation.

- p is reserved for primes
- $P(n)$ denotes the largest prime factor of n

P_y

$$P_y = \frac{\phi(A_y)}{A_y} = \prod_{p \leq y} \left(1 - \frac{1}{p}\right)$$

Because:

$$\phi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

Noting that every $p \leq y$ must divide A_y because that p will be included in $\text{lcm}[1, 2, \dots, y]$

S_y **and** S'_y

$$S_y = \sum_{\substack{P(a) \leq y \\ 2|a}} \frac{1}{a} e^{-a/s(a)}$$

$$S'_y = \sum_{\substack{a|A_y \\ 2|a}} \frac{1}{a} e^{-a/s(a)}$$

The authors note:

$$S'_y \leq S_y$$

Prove: If $a|A_y$ then $P(a) \leq y$

Proof: Assume $a|A_y$ then by definition we know that $a|\text{lcm}[1, 2, \dots, y]$

Case 1: $a \in \{1, 2, \dots, y\}$

Then necessarily $P(a) \leq y$ as a divisor must be less than the number being

divided

Case 2: a is some multiple of the elements in $\{1, 2, \dots, y\}$

$a = d_1 \cdot d_2 \cdot \dots \cdot d_n$ such that $d \in \{1, 2, \dots, y\}$. To find the prime factorization of a it is sufficient to find the prime factors of each value d . Each value of d is either prime or composite, if d is prime then its prime factorization is itself and is less than y . Otherwise if d is composite then all its prime factors are less than d and thus less than y .

So the difference between S_y and S'_y comes down to those a st $P(a) \leq y$ and $a \nmid A_y$. The authors observe that such values of a are divisible by p^k with $p^k > y$ and $p < y$. **I will have to leave the reasoning behind this open**

(3.2)

Where $a = p^k a'$

$$\sum_{\substack{P(a) \leq y \\ a \nmid A_y \\ 2|a}} \frac{1}{a} e^{-a/s(a)} \leq \sum_{p \leq y} \left(\sum_{k > \frac{\log y}{\log p}} \frac{1}{p^k} \sum_{P(a') \leq y} \frac{1}{a'} \right)$$

Is the second term just a way of generating any possible value that fit conditions $P(a) \leq y$ and $a \nmid A_y$ necessitating that it must be larger.

$$\sum_{p \leq y} \left(\sum_{k > \frac{\log y}{\log p}} \frac{1}{p^k} \sum_{P(a') \leq y} \frac{1}{a'} \right) \ll \log y \sum_{p \leq y} \frac{1}{y}$$

Where $f(s) \ll g(s)$ is equivalent to $f(s) \in O(g(s))$. This statement is correct as the second statement sums the same number of terms but with constantly higher values. I'm not sure where the log comes from.

$$\log y \sum_{p \leq y} \frac{1}{y} \ll 1$$

Remember that $f(n) \in O(g(n))$ means $\exists c, k$ st $0 \leq f(n) \leq cg(n)$, $\forall n \geq k$. All the above statement is saying is that the left term goes to a constant when y goes to the limit.

Mertens' Theorem

$$\lim_{n \rightarrow \infty} \ln n \prod_{p \leq n} \left(1 - \frac{1}{p} \right) = e^{-\gamma}$$

So

$$P_y = \frac{\phi(A_y)}{A_y} = \prod_{p \leq y} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\log y}$$

Then we have:

$$P_y S_y - P_y S'_y \sim \left(\frac{e^{-\gamma}}{\log y} \sum_{\substack{P(a) \leq y \\ 2|a}} \frac{1}{a} e^{-a/s(a)} \right) - \left(\frac{e^{-\gamma}}{\log y} \sum_{\substack{a|A_y \\ 2|a}} \frac{1}{a} e^{-a/s(a)} \right)$$

The authors observe that:

$$P_y S_y - P_y S'_y \ll \frac{1}{\log y}$$

Which I suppose makes sense as $S'_y \leq S_y$ so the subtraction will yield a smaller number times $\frac{e^{-\gamma}}{\log y}$

I understand the authors to mean that since

$$P_y S_y - P_y S'_y \ll \frac{1}{\log y}$$

then the limit $P_y S'_y$ exists if and only if $\lim_{y \rightarrow \infty} P_y S_y$ exists. If this is the case then

$$\lim_{y \rightarrow \infty} P_y S'_y = \lim_{y \rightarrow \infty} P_y S_y$$

I believe this comes from the fact that the difference between the 2 terms is going down to zero.

Next

For a prime q , we have $P_q = P_{q-1} \left(1 - \frac{1}{q}\right)$ This makes sense as

$$P_y = \prod_{p \leq y} \left(1 - \frac{1}{p}\right)$$

Kinda neat that this captures going directly to the next lesser prime.

For $q > 2$:

$$\begin{aligned} S_q - S_{q-1} &= \left(\sum_{\substack{P(a) \leq q \\ 2|a}} \frac{1}{a} e^{-a/s(a)} \right) - \left(\sum_{\substack{P(a) \leq (q-1) \\ 2|a}} \frac{1}{a} e^{-a/s(a)} \right) \\ &= \sum_{\substack{P(a) < q \\ 2|a}} \left(\sum_{i \geq 1} \frac{1}{a q^i} e^{-a q^i / s(a q^i)} \right) \end{aligned}$$

I have no fucking clue why that holds

They continue with:

$$\sum_{\substack{P(a) < q \\ 2|a}} \left(\sum_{i \geq 1} \frac{1}{aq^i} e^{-aq^i/s(aq^i)} \right) > \sum_{\substack{P(a) < q \\ 2|a}} \left(\sum_{i \geq 1} \frac{1}{aq^i} e^{-a/s(a)} \right)$$

Which again I don't know how to justify. My first thought is that the extra terms in the exponent would make it larger. However since it is over a fraction it seems ambiguous.

And finishing it up with:

$$\sum_{\substack{P(a) < q \\ 2|a}} \left(\sum_{i \geq 1} \frac{1}{aq^i} e^{-a/s(a)} \right) = \frac{1}{q-1} S_{q-1}$$

Beats me

Then:

$$P_q S_q > P_{q-1} \left(1 - \frac{1}{q} \right) S_{q-1} \left(1 - \frac{1}{q-1} \right)$$

We know that $P_q = P_{q-1} \left(1 - \frac{1}{q} \right)$ so we are simply using a different expression for P_q . Im not quite sure about the S_q term.

Continuing:

$$P_{q-1} \left(1 - \frac{1}{q} \right) S_{q-1} \left(1 - \frac{1}{q-1} \right)$$

I have no idea why this holds as well. However the result is interesting. This proves that the sequence $P_q S_q$ increases monotonically and the authors establish that the sequence is bounded above by 1/2. This means the limit exists!

Attempt to prove my limit exists

$$\begin{aligned} \Delta_k &= \lim_{y \rightarrow \infty} \frac{\phi(A_y)}{A_y} \sum_{a|A_y} \frac{a^{k-1}}{k! \cdot s(a)^k} \cdot e^{-a/s(a)} \\ &= \lim_{y \rightarrow \infty} \frac{\phi(A_y)}{A_y \cdot k!} \sum_{a|A_y} \left(\frac{a}{s(a)} \right)^k \frac{1}{a} e^{-a/s(a)} \end{aligned}$$

k I wanted to know what was happening with $\left(\frac{a}{s(a)}\right)^k$. It is not clear if the term is really having an impact as on average $\frac{a}{s(a)}$ over even values is quite close to one. For instance running the computations if you calculate the sum with the term removed the values are quite close to the computations with the term included. This trend did seem to break at Δ_4 however with the term-less computation being exactly double than with the term in. lets call:

$$\delta_k = \lim_{y \rightarrow \infty} \frac{\phi(A_y)}{A_y \cdot k!} \sum_{a|A_y} \frac{1}{a} e^{-a/s(a)}$$

and where the average of $a/s(a)$ over even value of a is ≈ 1.0561 let

$$\mathbb{A}_k = \lim_{y \rightarrow \infty} \frac{\phi(A_y)}{A_y \cdot k!} \sum_{a|A_y} \frac{1.0561^k}{a} e^{-a/s(a)}$$

k	0	1	2	3	4	5
Δ_k	0.1646	0.1659	0.0971	0.0435	0.0164	0.0054
δ_k	0.1646	0.1646	0.0823	0.0274	0.0069	0.0014
\mathbb{A}_k	0.1646	0.1738	0.0918	0.0323	0.0085	0.0018
$\Delta_k - \delta_k$	0	0.0013	0.0148	0.0161	0.0095	0.004

So δ_k is constantly smaller. The difference is relatively small but the magnitude between both is large. It appears that the difference between \mathbb{A}_k and Δ_k maybe some slow growing function.