# Notes on Pollack's "Not Always Buried Deep"

Stefanus Koesno[1]

[1] Somewhere in California
San Jose, CA 95134 USA

(Dated: June 29, 2017)

## Abstract

This is one of the best books on analytic number theory I've found to date. The comments here will be based on his notes with the same title instead of the book because the notes is free while the book is not :)

**Page 3**. On *Euclid's second proof.* The trick here is to show that $\varphi(P) > 1$, one thing to note is that $\varphi(n)$ measures the number of numbers that are co-prime to $n$. Since $\varphi(P) > 1$, there is a number that is co-prime to $P$ but in that case that number must contain a new prime. The other details were to make sure there is such a co-prime number.

**Page 4**. This is really clever, first let's do induction to show that $n_i = 2^{2^{i-1}} + 1$. For $i = 1$ this is obvious as $3 = 2^{2^0} + 1$ we just need to show the induction hypothesis is true, say $n_i = 2^{2^{i-1}} + 1$ we need to show the same holds for $n_{i+1}$

$$\begin{aligned} n_{i+1} &= 2 + \prod_{1 \le j < i+1} n_j \\ &= 2 + n_i \prod_{1 \le j < i} n_j \\ &= 2 + n_i(n_i - 2) \\ &= 2 - 2n_i + n_i^2 \\ &= (n_i - 1)^2 + 1 \\ &= 2^{2^i} + 1 \end{aligned}$$

Next we tackle the claim that this upper bound on $p_n$ gives us a lower bound on $\pi(x)$. Start with

$$\begin{aligned} p_i &< n_i \\ &< 2^{2^{i-1}} + 1 \\ &< 2^{2^{\pi(p_i)-1}} + 1 \\ \log\log(p_i) &< \pi(p_i) \end{aligned}$$

where there will be some very small correction terms due to the constant in the exponent and the one outside and of course due to the fact that we are using natural log instead of $\log_2$.

Page 5. Top of page, "the order of 2 (mod $p$) is precisely $2^i$", why is this so? Why can't the order be smaller than $2^i$, this is because

$$2^{2^i} = \left(\left((2)^2\right)^2\right)^2 \dots$$

so we are just squaring over and over again, and since $2^{2^{i-1}} \equiv -1$ this means that there's no lower exponent that produces $+1$ otherwise $2^{2^{i-1}}$ wouldn't have been $-1$.

Next we tackle the comment that "$a_n = 2^n - 1$ has the desired properties (note that $a_{11} = 23 \cdot 89$)." The problem is when $n = 13$ which is also a prime $2^{13} - 1 = 8191$ which is a prime, it even fails for $2^5 - 1 = 31$, in hindsight, this is obvious since there are things called the Mersenne primes of the form $2^n - 1$ :)

I then tried changing it to $a_n = 2^n + 1$ but it fails for $n = 3$ as $a_3 = 9$ and it doesn't have two distinct prime divisors but it's more promising, what we need is that $2^{2j+1} + 1$ to have more than 1 prime divisor, we know that $2^{2j+1} + 1$ is always a multiple of 3, it's obvious when you do (mod 3) as $2 \equiv -1 \pmod 3$.

The question now is if there will be a $k$ such that $3^k = 2^{2j+1} + 1$, well for $j = 1$ we have $3^2 = 2^3 + 1$ so how about for $j > 1$?

Suppose we find a $k$ such that $3^k = 2^{2j+1} + 1$, since $j > 1$, $4|2^{2j+1}$ and so by modulo 4 we know that $k$ is even since $3 \equiv -1 \pmod 4$, so we can denote $k = 2m$

$$\begin{aligned} 3^{2m} &= 2^{2j+1} + 1 \\ 3^{2m} - 1 &= 2^{2j+1} \end{aligned}$$

Focusing on the LHS

$$\begin{aligned} 3^{2m} - 1 &= (3-1)(3^{2m-1} + 3^{2m-2} + 3^{2m-3} + 3^{2m-4} + \cdots + 3 + 1) \\ &= 2(3^{2m-2}(3+1) + 3^{2m-4}(3+1) + \cdots + (3+1)) \\ &= 2^3(3^{2m-2} + 3^{2m-4} + \cdots + 1) \end{aligned}$$

Now there are $m$ terms inside the brackets in the RHS, if $m$ is odd then

$$3^{2m-2} + 3^{2m-4} + 3^{2m-6} + 3^{2m-8} + \cdots + 1 = (3^{2m-2} + 3^{2m-4}) + (3^{2m-6} + 3^{2m-8}) + \cdots + 1$$

Each bracket in the RHS is even and so the total is odd and thus

$$3^{2m} - 1 = 2^3(2M + 1)$$

and it can't be $2^{2j+1}$ since $(2M + 1)$ is odd. Now if $m$ is even we get

$$\begin{aligned} 3^{2m-2} + 3^{2m-4} + 3^{2m-6} + 3^{2m-8} + \cdots + 3^2 + 1 &= 3^{2m-4}(3^2 + 1) + 3^{2m-8}(3^2 + 1) + \cdots + (3^2 + 1) \\ &= 2 \cdot 5(3^{2m-4} + 3^{2m-8} + \cdots + 1) \end{aligned}$$

but 5 is prime and so $3^k - 1$ can't be $2^{2j+1}$. Therefore $2^{2j+1} + 1$ contains at least two prime divisors for $j > 1$ only for $j = 1$ does it contain only one prime divisor.

But we cannot use $a_n = 2^n + 1$ for the proof in the book, because then each $a_n$ is a multiple of 3 and they are not co-prime.

The proof in the book still works even is we use $a_n = 2^n - 1$ as long as we include $n = 11$ because then $a_2 a_3 a_5 a_7 a_{11}$ still have $5 + 1$ prime factors, note that $a_2, a_3, a_5, a_7$ are all prime numbers.

**Page 9**. Mid page-ish, "Proceeding as above, we deduce that $\sum_{p \le x}(p-1)^{-1} \ge \log\log x$", it's very tempting to just do log on both sides of (1.4) since the RHS is already $\log x$ but this will lead to some messy situation, salvageable but messy

$$\log\left(\prod_{p \le x} \frac{1}{1 - \frac{1}{p}}\right) \ge \log\log x$$
$$\sum_{p \le x} \log\left(\frac{p}{p-1}\right) \ge \log\log x$$

We can still proceed by showing that

$$\frac{1}{p-1} > \log\left(\frac{p}{p-1}\right)$$

One way to show this is to show $\frac{1}{x-1} > \log\left(\frac{x}{x-1}\right)$ instead with all real number $x \ge 2$ ( we choose 2 as the starting point for simplicity).

Well we know that at $x = 2$, $\frac{1}{2-1} > \log\left(\frac{2}{2-1}\right)$, we now show that the slope for $\frac{1}{x-1}$ is always smaller to that of $\log\left(\frac{x}{x-1}\right)$ for all $x$ and so $\frac{1}{x-1} > \log\left(\frac{x}{x-1}\right)$ always holds (we need a smaller slope because the function is a decreasing one).

The slope for $\frac{1}{x-1}$ is

$$\frac{d}{dx}\left(\frac{1}{x-1}\right) = -\frac{1}{(x-1)(x-1)}$$

while for $\log\left(\frac{x}{x-1}\right)$

$$\frac{d}{dx}\left(\log\left(\frac{x}{x-1}\right)\right) = -\frac{1}{x(x-1)}$$

and so the slope for $\frac{1}{x-1}$ is indeed smaller than that of $\log\left(\frac{x}{x-1}\right)$ and so $\frac{1}{x-1} > \log\left(\frac{x}{x-1}\right)$, this is one way.

Another way to show that $\sum_{p\le x}(p-1)^{-1} \ge \log\log x$ is to go back to the top of Page 9

$$\begin{aligned}\prod_{p\le x}\frac{1}{1-\frac{1}{p}} &= \prod_{p\le x}\left(1+\frac{1}{p-1}\right)\\ &\le \prod_{p\le x} e^{1/(p-1)} = e^{\sum_{p\le x}(p-1)^{-1}}\end{aligned}$$

Combining all the inequalities (the one involving $\log x$ is from (1.4))

$$\log x \le \prod_{p\le x}\frac{1}{1-\frac{1}{p}} \le e^{\sum_{p\le x}(p-1)^{-1}}$$

taking the logarithm of both the far LHS and far RHS we have

$$\log\log x \le \sum_{p\le x}(p-1)^{-1}$$

which is what we want to show in the first place.

**Page 11**. There's a simple mistake on *J. Perott's proof*, "by removing those divisible by $1^2\ldots$", we cannot remove numbers divisible by $1^2$ because then we will remove every number :)

**Page 12**. On J. Perott's proof, top of page, "It follows that $A(N)/N$ is bounded below $\ldots$ so the squarefree numbers have positive lower density", what it all means is that since $A(N)/N$ has a lower bound, if there are only a finite number of squarefree numbers then after a while (once we've passed the last one) $A(N)/N$ will get smaller and smaller and eventually gets smaller than the lower bound which is not allowed :)

**Page 13**. First paragraph, the conclusion of Erdos's super smart proof, "But $C\sqrt{N} < N/2$ whenever $N$ is large", the question is so what? we know that $C\sqrt{N}$ is the number of integers $\le N$ whose prime factors $\le M$. The problem here is that we know that from page 12, the number of integers $\le N$ whose prime factors $> M$ is $< N/2$, so the total number of integers (those with prime factors $\le M$ + those with prime factors $> M$) must be $N$ but here we have $< N/2 + < N/2$ which can never reach $N$

*Exercise* 1.2.7. **Page 9**. It's interesting that removing $\sum_p 1/p$ from $\sum_n 1/n$ makes the latter a convergent series. The question now is what is the minimum amount we need to subtract from $\sum_n 1/n$ to make it convergent?

Anyway, what we want to show is that $\sum_n 1/n$ with only *squarefull* $n$ converges. We start with

$$\prod_p \left(1 + \frac{1}{p^2} + \frac{1}{p^3} + \ldots\right) = \prod_p \left(\frac{1}{1-\frac{1}{p}}\right) - \frac{1}{p}$$

*i.e.* we remove the $1/p$ term from each series. Note that 1 is a squarefull number because the statement is that if $p|n$ then $p^2|n$ is a must, but for 1 no $p$ divides 1 so it is considered squarefull a well. We then follow the same method as shown on Page 9

$$\prod_p \left[\left(\frac{1}{1-\frac{1}{p}}\right) - \frac{1}{p}\right] = \prod_p \left(1 + \frac{1}{p-1} - \frac{1}{p}\right) = \prod_p \left(1 + \frac{1}{p(p-1)}\right)$$

Now note that $1/(p-1) < 2/p$ so

$$\prod_p \left(1 + \frac{1}{p(p-1)}\right) < \prod_p \left(1 + \frac{2}{p^2}\right) < \prod_p e^{2/p^2} = e^{\sum_p 2/p^2} < e^{2C}$$

We can safely deduce that $\sum_p 1/p^2 < C$ because it is a convergent series, since the above is convergent by Theorem 1.2.2, $\sum_n 1/n$ with $n$ only squarefull numbers is also convergent.

As for $\alpha$ so that $\sum_n 1/n^\alpha$ is also convergent, my first guess will be $\alpha > 1/2$, this is because with $\alpha = 1/2$ we will have $1/p$ term in the series and it will cause things to blow up. Another clue is in the fact that involving an exponent $\alpha$ means that the upper bound becomes

$$e^{\sum_p 2/p^{2\alpha}}$$

and if $\alpha$ goes below $\leq 1/2$ we will have an divergent series while $\sum_n 1/n^\beta$ is convergent if $\beta > 1$ by the integral test.

*Exercise* 1.3.1. **Page** 22. We want to show that

$$\int_2^x \frac{dt}{\log t} \sim \frac{x}{\log x}$$

using L'Hospital's rule. From what I know L'Hospital's rule only works if the limit is indefinite like $0/0$ or $\infty/\infty$ but looks like I was wrong, you can use is even if it's not indefinite.

So we need to do a derivative w.r.t $x$ on the integral above, how do we do it? for now assume that the integral is indefinite

$$\int \frac{dt}{\log t} = F(t) \qquad \longrightarrow \qquad \int_2^x \frac{dt}{\log t} = F(x) - F(2)$$

Taking the derivative with respect to $x$ means that

$$\frac{d}{dx}\int_2^x \frac{dt}{\log t} = \frac{d}{dx}F(x)$$

But in this case

$$\frac{d}{dx}F(x) = \frac{d}{dt}F(t)\bigg|_{t=x} = \frac{d}{dt}\left(\int \frac{dt}{\log t}\right)\bigg|_{t=x} = \frac{1}{\log x}$$

While (on the other hand)

$$\frac{d}{dx}\left(\frac{x}{\log x}\right) = \frac{\log x - 1}{\log^2 x} = \frac{1}{\log x} - \frac{1}{\log^2}$$

taking the ratio

$$\begin{aligned}\lim_{x\to\infty}\frac{x/\log x}{\int_2^x dt/\log t} &= \lim_{x\to\infty}\frac{1/\log x - 1/\log^2 x}{1/\log x}\\ &= \lim_{x\to\infty} 1 - \frac{1}{\log x}\\ &= 1\end{aligned}$$

**Page 24**. Before Lemma 1.4.2. This is a bit confusing, well more than a bit actually :) "Another way of viewing this argument is that all but finitely many primes fall into the unique reduced residue class (mod 2); from this perspective, the correct generalization is in plain sight. Namely, take any integer $q$; then every prime $p \nmid q$ has to fall into one of the $\varphi(q)/q$ reduced residue classes (mod $q$), and this forces the proportion of primes to be at most $\varphi(q)/q$."

First, $\varphi(q)$ is the Euler totient function, *i.e.* the number of numbers $< q$ that are co-prime to $q$. The original logic was that since half of all integers are even, so the proportion of prime numbers can only be at most 1/2. So at first, this looks a lot like the logic of sieve. Since you know that 2 is prime any multiple of 2 cannot be prime and so half of the numbers are gone. But this is not what we are doing, we are not doing sieve.

Instead what happens is this, if we choose $q = 4$ instead of 2, now if we group the integers as follows

$$\{\boldsymbol{1}, 2, \boldsymbol{3}, 4\}, \{\boldsymbol{5}, 6, \boldsymbol{7}, 8\}, \{\boldsymbol{9}, 10, \boldsymbol{11}, 12\}, \ldots$$

*i.e.* we are grouping them as residue classes of (mod 4), we see that each ***bold*** number is co-prime to 4, now the number itself might bot be prime, *e.g.* $\boldsymbol{9}$, but ***they*** are all co-prime to 4. Recall that $\gcd(a,q) = \gcd(a+qm,q)$ so if $a \equiv b \pmod{q}$ then $\gcd(a,q) = \gcd(b,q)$ and so in each residue class the number of numbers that are co-prime to $a$ stays the same.

A number can only be prime if it is co-prime to $q$, which in this case is 4, this is a necessary condition, not a sufficient one, but we are only interested in the upper bound here so the necessary condition is sufficient :)

So if we group the integers into the residue classes of $q$ there will be $\varphi(q)$ numbers that are co-prime to $q$ and so the upper bound of the ratio of the prime numbers to all of the numbers is obviously $\varphi(q)/q$. But of course there's a caveat, this is correct only for the second residue class and beyond, in the first residue class, $\{\boldsymbol{1},\underline{2},\boldsymbol{3},4\}$, 2 is also prime but it doesn't contribute towards $\varphi(q)$ because $2|4$, but this doesn't matter since the upper bound of the total ratio is given by

$$\frac{(\Delta + \varphi(q)) + \varphi(q) + \varphi(q) + \dots}{q + q + q + \dots} = \frac{\Delta + \infty \cdot \varphi(q)}{\infty \cdot q} = \frac{\varphi(q)}{q}$$

Now, the statement "all but finitely many primes fall into the unique reduced residue class (mod 2);" means this, say for our $q = 4$ example above, each residue class, whether the first, second or third, each of them can only take a finite amount of prime numbers and each prime number must of course fall into only one of them, ergo the word "unique", that's what the statement above means.

**Page 24**. Lemma 1.4.2. The first equality, recall that due to its multiplicity $\varphi(q) = \prod_{p_i} p_i^{\alpha_i - 1}(p_i - 1) = \prod_{p_i} p_i^{\alpha_i}(1 - 1/p_i)$