

# Jeffrey Stopple's Primer of Analytic Number Theory Book

Stefanus Koesno<sup>1</sup>

<sup>1</sup> Somewhere in California

San Jose, CA 95134 USA

(Dated: March 4, 2017)

## Abstract

Very good book, had a lot of fun

Saw this somewhere, show that

$$\sum_{n=1}^{\infty} \frac{2}{n(n+1)} = 1 + \frac{1}{3} + \frac{1}{6} + \frac{1}{10} + \dots = 2$$

historically this was first proved by Pietro Mengoli in 1650, the real gist of the story is that Mengoli couldn't deduce the value of the sum  $\sum_n 1/n^2$  which is similar to the one above (this quadratic harmonic series was finally solved by Euler, neither Leibnitz nor Bernoulli could do it). So back to the original problem, my approach was to split  $\frac{2}{n(n+1)}$  into a sum of two fractions

$$\begin{aligned} \frac{2}{n(n+1)} &= \frac{A}{n} + \frac{B}{n+1} \\ &= \frac{An + A + Bn}{n(n+1)} \end{aligned}$$

this means that  $A = 2$  and  $B = -2$  and thus the series can be written as

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{2}{n(n+1)} &= \sum_{n=1}^{\infty} \left( \frac{2}{n} - \frac{2}{n+1} \right) \\ &= \left( \frac{2}{1} - \frac{2}{2} \right) + \left( \frac{2}{2} - \frac{2}{3} \right) + \left( \frac{2}{3} - \frac{2}{4} \right) + \left( \frac{2}{4} - \dots \right) \\ &= \frac{2}{1} + \left( -\frac{2}{2} + \frac{2}{2} \right) + \left( -\frac{2}{3} + \frac{2}{3} \right) + \left( -\frac{2}{4} + \frac{2}{4} \right) + \dots \\ &= 2 \end{aligned}$$

so the series more or less “telescopes” and we are left with just the first term although I believe the method above is not legit as we turned an (absolutely?) convergent series into a conditionally convergent series?  $\backslash(^o\_o)/^-$  but to then again it might still be ok because if we denote

$$\begin{aligned} H_n &= \frac{2}{n} - \frac{2}{n+1} \\ \rightarrow S(k) &= \sum_{n=1}^k H_n \\ &= \frac{2}{1} - \frac{2}{k+1} \end{aligned}$$

and if we take the limit  $k \rightarrow \infty$

$$\begin{aligned} \lim_{k \rightarrow \infty} S(k) &= \frac{2}{1} - \frac{2}{\infty} \\ &= 2 \end{aligned}$$

the only trick remaining is showing that

$$\sum_{n=1}^{\infty} \frac{2}{n(n+1)} = \lim_{k \rightarrow \infty} S(k)$$

which can be shown by noting that if the sum is not infinite the two agree, *i.e.*

$$\sum_{n=1}^k \frac{2}{n(n+1)} = S(k)$$

so if we take both sides to infinity we get

$$\begin{aligned} \rightarrow \lim_{k \rightarrow \infty} \sum_{n=1}^k \frac{2}{n(n+1)} &= \lim_{k \rightarrow \infty} S(k) \\ \sum_{n=1}^{\infty} \frac{2}{n(n+1)} &= 2 \end{aligned}$$

sounds a lot like a roundabout way of saying the same thing LOL  $\neg \setminus (^{\circ} \_ o) / ^{-}$

-----  
 =====

**Exercise 1.2.1**, Page 12, Imitate this argument to get a formula for the hexagonal numbers  $h(n)$ .

A triangular number,  $t_n$  is a number where  $t_1 = 1$  and  $t_2 = 3$  and because it's 3 it's a triangular number, the explicit formula is

$$t_n = \frac{n(n+1)}{2}$$

which is just  $1 + 2 + \dots + n$ .

For the hexagonal number we have  $\Delta^2 h(n) = 6 - 2 = 4$  and so  $\Delta h(n) = 4n + C$  such that

$$\Delta^2 h(n) = \Delta(\Delta h(n)) = \Delta(4n + C) = (4(n+1) + C) - (4n + C) = 4$$

This means that  $h(n)$  should contain  $Cn + D$  such that  $\Delta h(n) = (C(n+1) + D) - (Cn + D) = C$  but this should not be the only term  $h(n)$  contains, it should also contain a term whose difference is  $4n$  because  $\Delta h(n)$  contains  $4n$  and that is provided by

$$\Delta(4t(n-1)) = 4t(n) - 4t(n-1) = 4 \frac{n(n+1)}{2} - 4 \frac{n(n-1)}{2} = 4n$$

so

$$h(n) = 4t(n-1) + Cn + D = 2n(n-1) + Cn + D$$

applying initial conditions  $h(1) = 1$  and  $h(2) = 6$

$$0 + C + D = 1$$

$$4 + 2C + D = 6$$

giving  $C = 1$  and  $D = 0$ , so

$$h(n) = 2n(n-1) + n = n(2n-1)$$

This seems to be correct since it gives

$$\begin{aligned} h(n) &: 1 \ 6 \ 15 \ 28 \ 45 \ 66 \ 91 \ 120 \ \dots, \\ \Delta h(n) &: 5 \ 9 \ 13 \ 17 \ 21 \ 25 \ 29 \ 33 \ \dots, \\ \Delta^2 h(n) &: 4 \ 4 \ 4 \ 4 \ 4 \ 4 \ 4 \ 4 \ \dots \end{aligned}$$

and in general ...

**Exercise 1.2.4**, Page 14, find a formula for polygonal numbers with  $a$  sides, for any  $a$ , *i.e.* a function  $f(n)$  with

$$\Delta^2 f(n) = a - 2, \quad \text{with } f(1) = 1 \text{ and } f(2) = a.$$

Using the result from Exercise 1.2.1 the generic formula we have is

$$f(n) = (a-2)t(n-1) + Cn + D = (a-2)n(n-1)/2 + Cn + D$$

applying initial conditions

$$0 + C + D = 1$$

$$(a-2) + 2C + D = a$$

which means it's actually independent of  $a$ , interesting, and therefore  $C = 1$  and  $D = 0$  all the time ... and therefore

$$f(n) = n \left[ \left( \frac{a}{2} - 1 \right) n - \frac{a}{2} + 2 \right]$$

the good thing about this is that it works for **any**  $a$ , zero, positive and negative.

**Exercise 1.2.5**, Page 16, verify that

$$n^1 + 3n^2 + n^3 = n^3$$

Now use this fact to find formulas for

$$\sum_{0 \leq k < n+1} k^3.$$

$$n^1 = n$$

$$3n^2 = 3n(n-1)$$

$$= 3n^2 - 3n$$

$$n^3 = n(n-1)(n-2)$$

$$= n^3 - 3n^2 + 2n$$

so it's obvious if we sum them we get  $n^3$ , now we need to “integrate” these things

$$\Sigma(n^1 + 3n^2 + n^3) = \frac{n^2}{2} + \frac{3}{3} \frac{n^3}{3} + \frac{n^4}{4}$$

using this in the big sum we get

$$\begin{aligned} \sum_{0 \leq k < n+1} k^3 &= \Sigma(n^1 + 3n^2 + n^3) \Big|_0^{n+1} \\ &= \frac{n^2}{2} + n^3 + \frac{n^4}{4} \Big|_0^{n+1} \\ &= \frac{(n+1)^2}{2} + (n+1)^3 + \frac{(n+1)^4}{4} \\ &= \frac{(n+1)n}{2} + (n+1)n(n-1) + \frac{(n+1)n(n-1)(n-2)}{4} \\ &= \frac{n(n+1)\{2 + 4(n-1) + (n-1)(n-2)\}}{4} \\ &= \frac{n^2(n+1)^2}{4} \end{aligned}$$

**Exercise 1.2.12**, Page 21, Use the fact that if  $\Delta f(k) = g(k)$  and  $a$  is any constant, then  $\Delta(f(k+a)) = g(k+a)$ , and the fact that  $2(k-1)^{-2} = 1/t_k$  to find the sum of the reciprocals of the first  $n$  triangular numbers

$$\frac{1}{t_1} + \frac{1}{t_2} + \dots + \frac{1}{t_n}.$$

Next compute

$$\frac{1}{T_1} + \frac{1}{T_2} + \dots + \frac{1}{T_n},$$

the sum of the reciprocals of the first  $n$  tetrahedral numbers.

So what we want is

$$\begin{aligned} \sum_{1 \leq k < n+1} \frac{1}{t_k} &= \Sigma \frac{1}{t_k} \Big|_1^{n+1} \\ &= \Sigma 2(k-1)^{-2} \Big|_1^{n+1} \\ &= \Sigma \frac{2(k-1)^{-1}}{-2+1} \Big|_1^{n+1} \\ &= -2 \left[ (n)^{-1} - 0^{-1} \right] \\ &= -2 \left[ \frac{1}{(n+1)} - \frac{1}{0+1} \right] \\ &= \frac{2n}{n+1} \end{aligned}$$

the thing to note here is that  $0^{-1}$  is **not** zero it is actually  $1/(0+1) = 1$ . For the tetrahedral numbers we know that it is given in Eq. 1.4 and 1.5

$$\begin{aligned} T_n &= \frac{n(n+1)(n+2)}{6} \\ \rightarrow \frac{1}{T_n} &= 6 \frac{1}{n(n+1)(n+2)} \\ &= 6(n-1)^{-3} \end{aligned}$$

so the sum is also pretty much the same

$$\begin{aligned} \sum_{1 \leq k < n+1} \frac{1}{T_k} &= \Sigma \frac{1}{T_k} \Big|_1^{n+1} \\ &= \Sigma 6(k-1)^{-3} \Big|_1^{n+1} \\ &= \Sigma \frac{6(k-1)^{-2}}{-3+1} \Big|_1^{n+1} \\ &= -3 \left[ (n)^{-2} - 0^{-2} \right] \\ &= -3 \left[ \frac{1}{(n+1)(n+2)} - \frac{1}{(0+1)(0+2)} \right] \\ &= \frac{3(n^2+3n)}{2(n+1)(n+2)} \end{aligned}$$

**Exercise 1.2.13**, Page 23, Use Summation by Parts and the Fundamental Theorem to compute  $\sum_{0 \leq k < n} H_k$ . (Hint: You can write  $H_k = H_k \cdot 1 = H_k \cdot k^0$ .) Your answer will have Harmonic numbers in it of course.

$$\Sigma H_k = \Sigma(H_k \cdot 1) = \Sigma(H_k \cdot k^0)$$

I don't know how to integrate  $H_k$  (since that is the question we are trying to answer here) but I know how to differentiate it, the derivative is just  $k^{-1}$  so we'll take  $u = H_k$  such that  $\Delta u = k^{-1}$  and so we'll take  $\Delta v = k^0$  such that  $v = k^1$ .

Next we need  $\Sigma(\Delta u \cdot Ev)$

$$\begin{aligned} \Sigma(\Delta u \cdot Ev) &= \Sigma(k^{-1} \cdot (k+1)^1) \\ &= \Sigma\left(\frac{1}{(k+1)} \cdot (k+1)k\right) \\ &= \frac{k^2}{2} \end{aligned}$$

so we have

$$\begin{aligned} \Sigma(u \cdot \Delta v) &= uv - \Sigma(\Delta u \cdot Ev) \\ \rightarrow \Sigma(H_k \cdot k^0) &= H_k k^1 - \frac{k^2}{2} \end{aligned}$$

to verify let's differentiate it

$$\begin{aligned} \Delta\left(H_k k^1 - \frac{k^2}{2}\right) &= k^{-1}(k+1)^1 + H_k k^0 - k^1 \\ &= \frac{(k+1)k}{k+1} + H_k - k \\ &= H_k \end{aligned}$$

note that the product rule here is different, there's a shift involved, see Page 22, and therefore

$$\begin{aligned} \sum_{0 \leq k < n} H_k \cdot k^0 &= \Sigma(u \cdot \Delta v)|_0^n \\ &= \left(H_k k^1 - \frac{k^2}{2}\right)\Big|_0^n \\ &= n\left(H_n - \frac{n-1}{2}\right) \end{aligned}$$

**Exercise 1.2.14**, Page 23, Use Summation by Parts and the Fundamental Theorem to compute  $\sum_{0 \leq k < n} k2^k$ . (Hint: You need the first part of Exercise 1.2.9.)

The thing to note here is that  $2^k$  is like  $e$  its derivative is itself

$$\begin{aligned}\Delta 2^k &= 2^{k+1} - 2^k \\ &= 2^k\end{aligned}$$

so we will want this to be  $\Delta v = 2^k$  such that  $v = 2^k$  and  $u = k = k^1$  such that  $\Delta u = k^0 = 1$  and so

$$\begin{aligned}\Sigma(u \cdot \Delta v) &= uv - \Sigma(\Delta u \cdot Ev) \\ &= k^1 2^k - \Sigma(k^1 2^{k+1}) \\ &= k 2^k - \Sigma(2^{k+1}) \\ &= k 2^k - 2^{k+1} \\ &= 2^k(k - 2)\end{aligned}$$

to verify let's differentiate the above

$$\begin{aligned}\Delta \{2^k(k - 2)\} &= 2^k((k + 1) - 2) + 2^k \\ &= k 2^k\end{aligned}$$

so it's correct and again the thing to note is that the product rule here now involves a shift, moving on

$$\begin{aligned}\sum_{0 \leq k < n} k 2^k &= (2^k(k - 2)) \Big|_0^n \\ &= 2^n(n - 2) - (-2) \\ &= 2^n(n - 2) + 2\end{aligned}$$

-----  
=====

**Exercise 2.1.7**, Page 30, we know that  $6 = 2^1 \cdot 3$  and  $28 = 2^2 \cdot 7$  are perfect numbers. The next ones are  $496 = 2^4 \cdot 31$  and  $8128 = 2^6 \cdot 127$  (check this!). What is the pattern in  $3, 7, 31, 127, \dots$ ? What is the pattern in the exponents  $1, 2, 4, 6, \dots$ ? Try to make



a conjecture about perfect numbers. Euclid, in his *Elements*, proved a general theorem about perfect numbers around the year 300 B.C.

$$3 = 2^2 - 1$$

$$7 = 2^3 - 1$$

$$31 = 2^5 - 1$$

$$127 = 2^7 - 1$$

so the exponents are 2, 3, 5, 7 which are all primes and are all one more than the 1, 2, 4, 6 so it looks like the formula for even perfect numbers are

$$2^{p-1} \cdot (2^p - 1)$$

where  $p$  is prime.

**Exercise 2.1.8**, Page 30, #2 and #3 are the interesting ones since the exponents of the  $2^p - 1$  are 11 and 13 and they are both prime so  $2^{10}(2^{11} - 1)$  and  $2^{12}(2^{13} - 1)$  are also prime, let's start with #2

$$\begin{aligned} 2096128 &= 2^{10}(2^{11} - 1) = 2^{10} \cdot 23 \cdot 89 \\ \rightarrow \sigma(2096128) &= \sigma(2^{10})\sigma(23)\sigma(89) \\ &= (2^{11} - 1)(23 + 1)(89 + 1) \\ &= 4421520 \\ \rightarrow s(2096128) &= \sigma(2096128) - 2096128 = 2325392 \\ &\neq 2096128 \end{aligned}$$

so even though 11, the exponent of  $2^{11} - 1$  is prime,  $2^{10}(2^{11} - 1)$  is not perfect, while

$$\begin{aligned}
33550336 &= 2^{12}(2^{13} - 1) = 2^{12} \cdot 8191 \\
\rightarrow \sigma(33550336) &= \sigma(2^{12})\sigma(8191) \\
&= (2^{13} - 1)(8191 + 1) \\
&= 8191 \cdot (8191 + 1) \\
&= 67100672 \\
\rightarrow s(33550336) &= \sigma(33550336) - 33550336 = 33550336
\end{aligned}$$

so it is a perfect number, thus the formula has to be corrected to only include  $2^p - 1$  that is prime (and  $p$  itself is also prime)

**Exercise 2.1.12**, Page 31, there is an interesting theorem about perfect numbers and sums of cubes. For example,

$$28 = 1^3 + 3^3$$

Try to make a conjecture about what is true. Ignore the first perfect number, 6; it doesn't fit the pattern.

$$\begin{aligned}
28 &= 1^3 + 3^3 \\
496 &= 1^3 + 3^3 + 5^3 + 7^3 \\
8128 &= 1^3 + 3^3 + 5^3 + 7^3 + 9^3 + 11^3 + 13^3 + 15^3
\end{aligned}$$

the pattern is given by the next exercise

**Exercise 2.1.12**, Page 31,

$$\begin{aligned}
1^3 + 3^3 + 5^3 + \dots + (2N - 1)^3 &= (1^3 + 2^3 + 3^3 + \dots + (2N)^3) - (2^3 + 4^3 + 6^3 + \dots + (2N)^3) \\
&= \frac{(2N)^2(2N + 1)^2}{4} - 2^3(1^3 + 2^3 + 3^3 + \dots + N^3) \\
&= \frac{4N^2(2N + 1)^2}{4} - 2^3 \frac{N^2(N + 1)^2}{4} \\
&= N^2(2N + 1)^2 - 2N^2(N + 1)^2 \\
&= N^2(4N^2 + 4N + 1 - 2N^2 - 4N - 2) \\
&= N^2(2N^2 - 1)
\end{aligned}$$

so  $N = 2^{(p-1)/2}$  where  $p$  is prime and  $2^p - 1$  is also prime

**Exercise 2.1.16**, Page 33, Suppose  $m$  and  $n$  are integers such that  $\sigma(m) = m + n = \sigma(n)$ . What can you say is true about  $s(m)$  and  $s(n)$ ?

$$\begin{aligned}s(m) &= \sigma(m) - m \\ &= m + n - m \\ s(m) &= n\end{aligned}$$

and

$$\begin{aligned}s(n) &= \sigma(n) - n \\ &= m + n - n \\ s(n) &= m\end{aligned}$$

so a pair of amicable numbers  $m$  and  $n$  are those such that  $s(m) = n$  and  $s(n) = m$  and therefore  $\sigma(m) = \sigma(n) = m + n$ .

### Chapter 3

Lessons from Chapter 3

- Area is equal to length if the width is one, this is the trick used over and over, once you morphed length into area you can use integrals

### Chapter 4

His notion of average order is slightly different than that of Apostol's

**Exercise 4.2.4**, Page 76, the differences are

$$\begin{aligned}\frac{1}{3} - \frac{27}{82} &= 0.0040650406504065040650406504065 \\ \frac{27}{82} - \frac{15}{46} &= 0.00318133616118769883351007423118 \\ \frac{15}{46} - \frac{12}{37} &= 0.00176263219741480611045828437133\end{aligned}$$

very small indeed, roughly one part in a thousand

**Page 80, Corollary.** *The average order of  $\sigma(n)$  is  $\zeta(2)n$ .* Note that here he's talking about average order quite differently from Apostol, his reasoning here is since

$$\begin{aligned}\sum_{k=1}^n \sigma(k) &\sim \zeta(2) \frac{n^2}{2} \\ \sum_{k=1}^n \zeta(2)k &\sim \zeta(2) \frac{n^2}{2} \\ \rightarrow \sum_{k=1}^n \sigma(k) &\sim \sum_{k=1}^n \zeta(2)k\end{aligned}$$

so since both sides are summed exactly the same each term as the same average order?

—\(^{\circ}\_{-}\circ)/—

### Interlude 1

**Exercise I1.3.4**, Page 95, from the *definition* of  $\log(x)$  as an integral over  $1/t$

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

The second term in the RHS can be massaged using the usual change of variable  $s = xt$  such that  $t = s/x$  and  $dt = ds/x$  and so

$$\begin{aligned}\int_x^{xy} \frac{1}{t} dt &= \int_1^y \frac{\cancel{x}}{s \cancel{x}} ds \\ &= \log(y)\end{aligned}$$

and so  $\log(xy) = \log(x) + \log(y)$ , we can actually do the change of variable in a slightly different way

$$\begin{aligned}\int_x^{xy} \frac{1}{t} dt &= \int_1^y \frac{x}{x} \times \frac{1}{t} dt \\ &= \int_x^{xy} \frac{1}{(xt)} d(xt), \quad s = xt \\ &= \int_1^y \frac{1}{s} ds\end{aligned}$$

it was OK to do that because here  $x$  is constant w.r.t the integration. A more interesting thing would be  $\log(x/y)$  in this case

$$\begin{aligned}\log(x/y) &= \int_1^{x/y} \frac{1}{t} dt \\ &= \int_1^x \frac{1}{t} dt - \int_{x/y}^x \frac{1}{t} dt\end{aligned}$$

here we have a subtraction because  $x/y < x$ , the next steps are the same as previous case's.

## Chapter 5

**Page 107**, lower half of page, nearing the funny trick to prove the Theorem we have

$$\begin{aligned}\pi(2n+1) &< 2 \frac{2n+1}{\log(2n)} \\ &< 2 \frac{2n+1}{\log(2n+1)} \frac{\log(2n+1)}{\log(2n)}\end{aligned}$$

and for  $n = 30$ ,  $\frac{\log(2n+1)}{\log(2n)} = 1.00403710574428$  so we can just adjust the constant 2 to 2.008... or something :) and we can certainly do approximation this way as well

$$\begin{aligned}\frac{\log(2n+1)}{\log(2n)} &= \frac{\log(2n) + \log(1 + 1/2n)}{\log(2n)} \\ &\approx 1\end{aligned}$$

But actually if you numerically calculate  $3.39 \frac{n}{\log(n)} + 1$  and  $2 \frac{2n+1}{\log(2n+1)}$  starting at  $n = 67$ ,  $3.39 \frac{n}{\log(n)} + 1 < 2 \frac{2n+1}{\log(2n+1)}$  and we can sort of prove this.

Even simpler, we can compare  $4n/\log(2n)$  and  $2(2n+1)/\log(2n+1)$ , let's take their ratio

$$\begin{aligned}\frac{2(2n+1)/\log(2n+1)}{4n/\log(2n)} &= \frac{4n+2}{\log(2n+1)} \frac{\log(2n)}{4n} \\ &= \frac{4n+2}{4n} \frac{\log(2n)}{\log(2n+1)} \\ &= \left(1 + \frac{1}{2n}\right) \frac{\log(2n)}{\log(2n) + \log\left(1 + \frac{1}{2n}\right)} \\ &= \left(1 + \frac{1}{2n}\right) \frac{1}{1 + \frac{\log\left(1 + \frac{1}{2n}\right)}{\log(2n)}}\end{aligned}$$

if  $\frac{\log(1+\frac{1}{2n})}{\log(2n)}$  is smaller than  $\frac{1}{2n}$  then the whole ratio is greater than one and  $2(2n+1)/\log(2n+1) > 4n/\log(2n)$ . So again we take the ratio

$$\begin{aligned} \frac{1/2n}{\log(1+\frac{1}{2n})/\log(2n)} &= \frac{\log(2n)}{2n \log(1+\frac{1}{2n})} \\ &= \frac{\log(2n)}{\log(1+\frac{1}{2n})^{2n}} \end{aligned}$$

Now

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$$

is the definition of  $e$  and it takes its minimum value when  $n = 1$  and as  $n$  grows it grows bigger, we can see this using the binomial expansion (there's a derivation showing exactly this by Gilles Cazalais and I'll reproduce it here), first denote

$$t_n = \left(1 + \frac{1}{n}\right)^n$$

and so

$$\begin{aligned} t_n &= \sum_{k=0}^n \binom{n}{k} \left(\frac{1}{n}\right)^k \\ &= 1 + n \left(\frac{1}{n}\right) + \frac{n(n-1)}{2!} \left(\frac{1}{n^2}\right) + \frac{n(n-1)(n-2)}{3!} \left(\frac{1}{n^3}\right) + \dots + \frac{n(n-1)(n-2)\dots 1}{n!} \left(\frac{1}{n^n}\right) \\ &= 1 + 1 + \frac{1}{2!} \left(1 - \frac{1}{n}\right) + \frac{1}{3!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) + \dots + \frac{1}{n!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{n-1}{n}\right) \end{aligned}$$

the two key observations here are

$$0 < \left(1 - \frac{k}{n}\right) < 1, \quad \text{and} \quad \left(1 - \frac{k}{n}\right) < \left(1 - \frac{k}{n+1}\right)$$

this means that  $t_n < t_{n+1}$ , in our case we only care about even  $n$ 's but the conclusion is the same  $t_{2n} < t_{2(n+1)}$ .

So when  $n < \infty$ ,  $\log(1+\frac{1}{2n})^{2n} < \log(e) < 1$ . Now we need the numerator to be  $> 1$  so we need to take  $n \geq 2$  to ensure  $2n > e$  and this is good enough for us.

So in conclusion,  $\frac{1/2n}{\log(1+\frac{1}{2n})/\log(2n)} > 1$  means that  $(1+\frac{1}{2n})^{\frac{1}{1+\log(1+\frac{1}{2n})/\log(2n)}} > 1$  and that finally means that  $2(2n+1)/\log(2n+1) > 4n/\log(2n)$  and since  $3.39 \frac{n}{\log(n)} + 1 < 4n/\log(2n)$  it also means that  $3.39 \frac{n}{\log(n)} + 1 < 2(2n+1)/\log(2n+1)$ .

**Page 108**, Theorem that says

$$\frac{1}{2} \frac{x}{\log(x)} < \pi(x)$$

for  $x \geq 15$ . In fact we can say that this is true for  $x \geq 3$ , 15 numbers are not too much to check so if you just calculate all 15 of them you'll see that this inequality holds for  $x \geq 3$

**Page 108**, Lemma, this is actually Theorem 3.14 of Apostol, so here's I replicate my own explanation of why this is true, visually we can see it below with an example for  $\alpha(p)$  with  $x = 10$  and  $p = 2$ , each column shows the number of 2's in each number (we only consider even numbers here obviously)

$$\begin{array}{rcl} 2 & \longrightarrow & x/p^3 = 10/2^3 = 1 \text{ factor of } 2 \\ 2 & 2 & \longrightarrow x/p^2 = 10/2^2 = 2 \text{ factor of } 2 \\ 2 & 2 & 2 & 2 & 2 & \longrightarrow x/p^1 = 10/2^1 = 5 \text{ factor of } 2 \\ \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \\ 2 & 4 & 6 & 8 & 10 & \longrightarrow \text{a total of } 5 + 2 + 1 = 8 \text{ factors of } 2 \end{array}$$

## Interlude 2

**Page 124**, the “inverse” of  $\sin(x)$ , here

$$\begin{aligned} \sin(x) &= x + O(x^3) \\ \rightarrow \frac{1}{\sin(x)} &= \frac{1}{x} + O(x) \end{aligned}$$

and it says that he got this “by long division”, the long division he mentioned was about guessing what the outcome of a division of two polynomials by expecting the result to be  $a_0 + a_1x + a_2x^2 + \dots$ , the thing is it doesn't cover  $1/x$  so let's say here we include  $1/x$  so

$$\begin{aligned} \frac{1}{\sin(x)} &= a_{-1} \frac{1}{x} + a_0 + a_1x + \dots \\ \rightarrow 1 &= \sin(x) \frac{1}{\sin(x)} \\ &= \{x + O(x^3)\} \left\{ a_{-1} \frac{1}{x} + a_0 + a_1x + \dots \right\} \\ &= a_{-1} + a_0x + a_1x^2 + O(x^2) + O(x^3) + O(x^4) + \dots \end{aligned}$$

this means that  $a_{-1} = 1$  and the rest is  $O(x)$ , why? because here we are talking about Big Oh in  $x \rightarrow 0$  instead of  $x \rightarrow \infty$  and so  $O(x)$  gives a bigger error than  $O(x^2)$  and so on.

And we actually don't need to do long division to do this we can immediately approximate it using

$$\begin{aligned}\frac{1}{\sin(x)} &= \frac{1}{x + O(x^3)} \\ &= \frac{1}{x} \left( \frac{1}{1 + O(x^2)} \right) \\ &= \frac{1}{x} (1 - O(x^2) + O(x^4) - \dots) \\ &= \frac{1}{x} - O(x) + O(x^3) - \dots\end{aligned}$$

and we pick  $O(x)$  as the biggest error since  $x \rightarrow 0$

**Exercise I2.4.2**, Page 126, the geometric series

$$1 + \frac{1}{10} + \frac{1}{100} + \frac{1}{1000} + \frac{1}{10000} + \dots$$

is the distance where Achilles caught up with the turtle, why? because the turtle first traveled 1 km, and by the time Achilles traveled 1 km the turtle traveled 1/10 km and so on, if it's still not convincing let's calculate the time it took Achilles to catch up with the turtle, the turtle had a 1 km head start and so the time taken by Achilles to travel  $1 + d$  km the turtle spent the same amount of time to travel  $d$  km

$$\begin{aligned}\frac{d}{v_t} &= \frac{1 + d}{10v_t} \\ d &= \frac{1}{9}\end{aligned}$$

so the total distance traveled is  $1 + \frac{1}{9} = \frac{10}{9}$  which is the sum of the infinite geometric series above and the time needed for Achilles to catch up with the turtle is  $1/9v_t$  so it does depend on the turtle's speed, the thing that matters here is the ratio between Achilles's to the turtle's speed because that controls the  $r$  in the geometric series.

**Exercise I2.6.3**, Page 132, the first step is actually quite unnecessary, we need not induction :) first note that

$$\frac{1}{2}H_N = \frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \dots + \frac{1}{2N}$$



therefore

$$H_{2N} - \frac{1}{2}H_N = 1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \dots + \frac{1}{2N-1}$$

i.e. the odd harmonic series, if we repeat the process one more time

$$H_{2N} - \frac{1}{2}H_N - \frac{1}{2}H_N = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} + \dots + \frac{1}{2N-1} - \frac{1}{2N}$$

which is what we want and so we immediately see that what we need is  $H_{2N} - H_N$

**Exercise I2.6.4**, Page 132, sub-problem 1, there's No Exercise I2.9, I believe it should have been Exercise I2.2.1, but second, I stupidly explicitly computed the sum  $\sum_{n=0}^{2N} (-1)^n t^{2n}$ , since  $n$  takes on the values 0 to  $2N$ , there are odd numbers of them and the first and last  $n$  are even, so there's one less odd  $n$

$$\begin{aligned} \sum_{n=0}^{2N} (-1)^n t^{2n} &= \sum_{n=0}^N (-1)^{2n} t^{4n} + \sum_{n=0}^{N-1} (-1)^{2n+1} t^{4n+2} \\ &= t^{4N} + \sum_{n=0}^{N-1} t^{4n} - t^2 \sum_{n=0}^{N-1} t^{4n} \\ &= t^{4N} + (1 - t^2) \sum_{n=0}^{N-1} (t^4)^n \\ &= t^{4N} + (1 - t^2) \frac{1 - (t^4)^N}{1 - t^4} \\ &= t^{4N} + \frac{1 - (t^4)^N}{1 + t^2} \\ &= \frac{1}{1 + t^2} + \frac{t^{4N} + t^{4N+2} - t^{4N}}{1 + t^2} \\ &= \frac{1}{1 + t^2} + \frac{t^{4N+2}}{1 + t^2} \end{aligned}$$

2. The integral of the first term in the RHS can be done using trig substitution noting that  $1 + \tan^2(u) = \sec^2(u)$  and so let  $t = \tan(u)$ ,  $dt = \sec^2(u)du$

$$\begin{aligned} \int \frac{1}{1 + t^2} dt &= \int \frac{1}{1 + \tan^2(u)} \sec^2(u) du \\ &= \int \cos^2(u) \sec^2(u) du \\ &= u \\ &= \arctan(t) \end{aligned}$$

and so

$$\begin{aligned}\int_0^1 \frac{1}{1+t^2} dt &= \arctan(1) - \arctan(0) \\ &= \frac{\pi}{4}\end{aligned}$$

3. Property v is that if  $g(x) < f(x)$  in the range of  $a \leq x \leq b$  then  $\int_a^b g(x) < \int_a^b f(x)$  and in this case

$$\frac{t^{4N+2}}{1+t^2} < t^{4N+2}$$

and so the integral of the LHS is also smaller than the integral of the RHS and

4. The integral of the RHS is

$$\begin{aligned}\int_0^1 t^{4N+2} &= \frac{1^{4N+3}}{4N+3} = \frac{1}{4N+3} \\ \lim_{N \rightarrow \infty} \frac{1}{4N+3} &= 0\end{aligned}$$

and since  $\int_0^1 \frac{t^{4N+2}}{1+t^2} < \int_0^1 t^{4N+2}$  it is also zero.

**Page 134**, Proof of Theorem about the limit test, since the book doesn't show the case for  $L > 1$  let me do it here, it's very similar, first pick a real number  $r$  between  $L$  and 1 that is bigger than 1 but smaller than  $L$ , again the choice given in the book works,  $r = (L + 1)/2$  and this means

$$\frac{|c_{n+1}|}{|c_n|} = L > r$$

and therefore

$$|c_{N+k}| > r^k |c_N|$$

but now since  $r > 1$ , the series  $|c_N| \sum_{k=0}^{\infty} r^k$  diverges and so  $\sum_{k=0}^{\infty} |c_{N+k}| = \sum_{n=N}^{\infty} |c_n|$  also diverges.

**Exercise I2.6.8**, Page 135, the series  $L(x)$  is the Euler's dilogarithm function

$$\begin{aligned}
 L(x) &= \int_0^x \frac{-\log(1-t)}{t} dt \\
 &= \int_0^x \frac{1}{t} \sum_{k=1}^{\infty} \frac{1}{k} t^k \\
 &= \int_0^x \sum_{k=1}^{\infty} \frac{1}{k} t^{k-1} \\
 &= \sum_{k=1}^{\infty} \int_0^x \frac{1}{k} t^{k-1} \\
 &= \sum_{k=1}^{\infty} \frac{x^k}{k^2}
 \end{aligned}$$

using the limit test we get

$$\begin{aligned}
 \frac{|x^{k+1}/(k+1)^2|}{|x^k/k^2|} &= \frac{|x|k^2}{(k+1)^2} \\
 \rightarrow \lim_{k \rightarrow \infty} \frac{|x|k^2}{(k+1)^2} &= |x|
 \end{aligned}$$

so the radius of convergence is  $R = 1$ .

**Exercise I2.6.11**, Page 137, this is an interesting exercise, the first step is just an integration  $\int 1/(1+x) = \log(1+x)$ , if we take the difference between the degree  $n$  Taylor approximation and the actual function  $\log(1+x)$  is (the thing to note here is that the lower bound of the integral depends on where we do the Taylor expansion, here we choose 0 because if we choose  $-1$ ,  $\log(1-1)$  blows up so the integral will blow up as well)

$$\begin{aligned}
 E_n(x) &= \frac{1}{n!} \int_0^x (x-t)^n f^{(n+1)}(t) dt \\
 &= \frac{1}{n!} \int_0^x (x-t)^n \frac{(-1)^n n!}{(1+t)^{n+1}} dt \\
 &\leq (x-0)^n \int_0^x \frac{(-1)^n}{(1+t)^{n+1}} dt
 \end{aligned}$$

from the hint

$$\frac{(-1)^n}{(1+t)^{n+1}} \leq \frac{1}{(1+t)^{n+1}} \leq 1$$

moving forward

$$\begin{aligned} E_n(x) &\leq x^n \int_0^x \frac{1}{(1+t)^{n+1}} dt \\ &= x^n \left( -\frac{1}{n(1+t)^n} \Big|_0^x \right) \end{aligned}$$

setting  $x$  to 1 we get

$$\begin{aligned} x^n \left( -\frac{1}{n(1+t)^n} \Big|_0^x \right) &= -\frac{1}{n(1+1)^n} + \frac{1}{n} \\ &= \frac{1}{n} \left( 1 - \frac{1}{2^n} \right) \\ &\rightarrow |E_n(1)| \leq \frac{1}{n} \end{aligned}$$

since  $(1 - \frac{1}{2^n}) < 1$ , but I couldn't figure out why  $|E_n(1)| \leq \frac{1}{n+1}$ , although as  $n$  gets bigger  $\frac{1}{n} (1 - \frac{1}{2^n})$  gets real close to  $\frac{1}{n+1}$  as the limit of the ratio between the two is 1.

Although the real question here is that whether we can set  $x = 1$  since the radius of convergence is  $R = 1$  (recall that at the radius itself the behavior is unknown, so it may be good or bad), I was initially surprised by this as the expansion of  $\log(1+x)$  should work if  $x \geq 1$ , it's not legit if  $x = -1$  but  $x \geq 1$  should be fine but if you do the limit test the radius is indeed 1.

**Page 139**, about convergence, we cannot rearrange terms but within each partial sums we can rearrange them because each partial sum is finite?  $\neg \setminus (^{\circ} \_o) / -$

**Exercise I2.6.12**, Page 136, please see Fig. 1. This means that

$$\sum_{n=2}^N \frac{1}{n \log(n)} > \int_2^{N+1} \frac{1}{x \log(x)} dx$$

note the upper limit of the integral, next

$$\begin{aligned} \int_2^{N+1} \frac{1}{x \log(x)} dx &= \log(\log(x)) \Big|_2^{N+1} \\ &= \log(\log(N+1)) - \log(\log(2)) \end{aligned}$$

which goes to infinity when  $N \rightarrow \infty$ .

**Page 143**, proof of Abel's Theorem Part I, "Because  $f_m(x)$  is just a polynomial, it is continuous. So, for some  $\delta > 0$ , it is true that  $|f_m(x) - f_m(1)| < \epsilon/3$  whenever  $|x-1| < \delta$ ". Recall the definition of continuity,  $f(x)$  is continuous at  $x = a$  if and only if

$$\lim_{x \rightarrow a} f(x) = f(a)$$

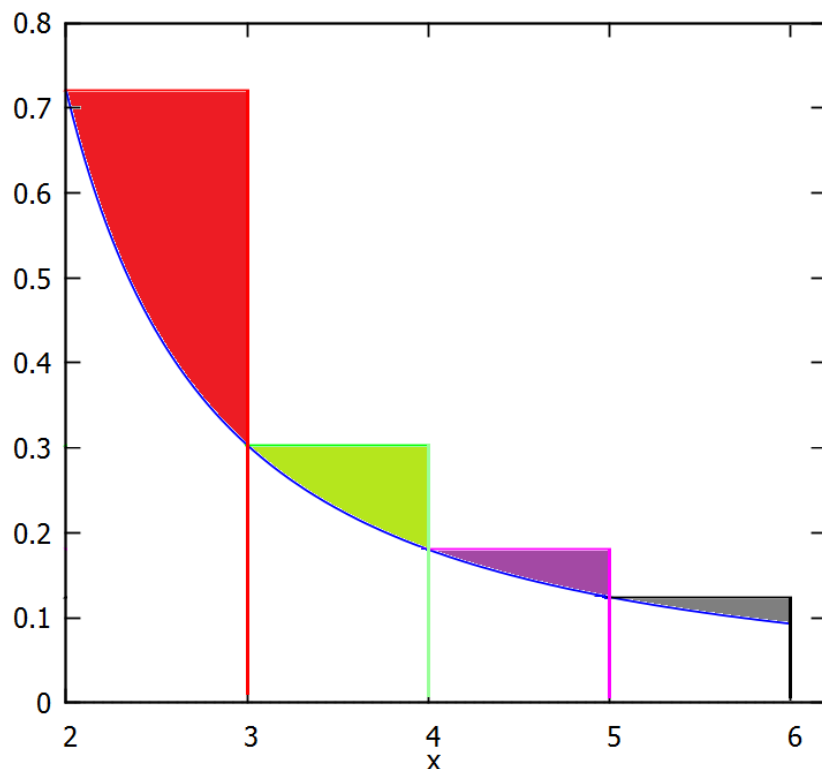


FIG. 1: Plot of  $1/x \log(x)$  and rectangles of height  $1/n \log(n)$  and width 1.

and from the definition of limit, “for every  $\epsilon$  we can find a  $\delta$  such that  $|f(x) - f(a)| < \epsilon$  when  $0 < |x - a| < \delta$ ”.

The next question is whether the  $\delta$ ’s are all the same, here we have three  $\delta$ ’s, so if they are all different we just pick the biggest one because the definition of limit/continuity only requires that for every  $\epsilon$  we can find a  $\delta$  and so the  $\delta$  is a free variable.

## Chapter 6

**Page 147**, Euler’s proof of the quadratic harmonic series, from Eq 6.2 we can see that we have every quadratic numbers in the denominators,  $4, 9, 16, \dots$  therefore we will be able to make any quartic and other even powers, but it’s almost impossible to make odd powers

**Page 154**, two typos, first mid of page, the substitution is reversed it should’ve been  $x = 2iz$  instead of the other way around stated in the book, bottom of page, formula for  $\cot(z)$ , an overall factor of  $\frac{1}{2^n}$  is missing.

## Chapter 8

**Exercise 8.2.1**, Page 199, first we expand

$$\begin{aligned}
 1 - 2^{1-s} &= 1 - e^{(1-s)\log(2)} \\
 &= 1 - \left( \sum_{n=0}^{\infty} \frac{1}{n!} [(1-s)\log(2)]^n \right) \\
 &= \sum_{n=1}^{\infty} \frac{1}{n!} (-1)^{n+1} (s-1)^n \log^n(2)
 \end{aligned}$$

we now calculate the expansion for  $1/(1 - 2^{1-s})$ , we use long division for this, or we can use geometric series if  $|s| > 1$

$$\begin{aligned}
 1 &= (1 - 2^{1-s}) \times \frac{1}{(1 - 2^{1-s})} \\
 &= \left( \sum_{n=1}^{\infty} \frac{1}{n!} (-1)^{n+1} (s-1)^n \log^n(2) \right) \times \left( \sum_{k=-1}^{\infty} a_k (s-1)^k \right)
 \end{aligned}$$

and so

$$\begin{aligned}
 1 &= \log(2)a_{-1} \\
 0 &= \log(2)a_0 - \frac{1}{2!} \log^2(2)a_{-1} \\
 0 &= \log(2)a_1 - \frac{1}{2!} \log^2(2)a_0 + \frac{1}{3!} \log^3(2)a_{-1} \\
 0 &= \log(2)a_2 - \frac{1}{2!} \log^2(2)a_1 + \frac{1}{3!} \log^3(2)a_0 - \frac{1}{4!} \log^4(2)a_{-1}
 \end{aligned}$$

and so  $a_{-1} = 1/\log(2)$ ,  $a_0 = 1/2$ ,  $a_1 = (1/12)\log(2)$ ,  $a_2 = 0$ ,  $a_3 = \dots$

Now, as the book says that  $\Phi(s) = \log(2) + O(s-1)$ , since if we Taylor expand around  $s = 1$  we know that  $\Phi(1) = \log(2)$ , so the first term is indeed  $\log(2)$  and since we are expanding around  $s = 1$  the correction terms are of order  $O(s-1)$  (note that here  $O(s-1)$  is bigger than  $O((s-1)^m)$  since  $s \rightarrow 1$ ). Now from the definition

$$\begin{aligned}
 \zeta(s) &= \frac{1}{1 - 2^{1-s}} \Phi(s) \\
 &= \left( \frac{1}{\log(2)} \frac{1}{s-1} + \frac{1}{2} + \frac{\log(2)}{12} (s-1) + \dots \right) (\log(2) + O(s-1)) \\
 &= \frac{1}{s-1} + O(1)
 \end{aligned}$$

again recall here we are interested in  $s \rightarrow 1$  and so  $O(1)$  is the biggest error, all other Big Oh are in  $O((s-1)^m)$ ,  $m \geq 1$

**Exercise 8.2.2**, Page 202, let's apply Euler operators using Maxima, so

$$x \frac{d}{dx} \frac{-x}{(1+x)^2} = x \frac{d}{dx} \sum_{n=1}^{\infty} (-1)^n n x^n$$

$$\frac{x(x-1)}{(x+1)^3} = \sum_{n=0}^{\infty} (-1)^n n^2 x^n$$

and so  $\Phi(-2) = 0$  and therefore  $\zeta(-2) = 0$ , continuing  $\mathcal{E}$

$$\mathcal{E}^2 \left( \frac{-x}{(1+x)^2} \right) = -\frac{x(x^2 - 4x + 1)}{(x+1)^4}$$

$$\mathcal{E}^3 \left( \frac{-x}{(1+x)^2} \right) = \frac{x(x^3 - 11x^2 + 11x - 1)}{(x+1)^5}$$

$$\mathcal{E}^4 \left( \frac{-x}{(1+x)^2} \right) = -\frac{x(x^4 - 26x^3 + 66x^2 - 26x + 1)}{(x+1)^6}$$

$$\mathcal{E}^5 \left( \frac{-x}{(1+x)^2} \right) = \frac{x(x^5 - 57x^4 + 302x^3 - 302x^2 + 57x - 1)}{(x+1)^7}$$

$$\mathcal{E}^6 \left( \frac{-x}{(1+x)^2} \right) = -\frac{x(x^6 - 120x^5 + 1191x^4 - 2416x^3 + 1191x^2 - 120x + 1)}{(x+1)^8}$$

and the  $\Phi(s)$ 's are  $-1$  times the above and so and connecting it to the next exercise

$$\begin{aligned}\zeta(-1) &= (1-2^2)^{-1} \Phi(-1) = -\frac{1}{12} = -\frac{1}{2} \times \frac{1}{6} \\ \zeta(-2) &= (1-2^3)^{-1} \Phi(-2) = 0 \\ \zeta(-3) &= (1-2^4)^{-1} \Phi(-3) = \frac{1}{120} = -\frac{1}{4} \times -\frac{1}{30} \\ \zeta(-4) &= (1-2^5)^{-1} \Phi(-4) = 0 \\ \zeta(-5) &= (1-2^6)^{-1} \Phi(-5) = -\frac{1}{252} = -\frac{1}{6} \times \frac{1}{42} \\ \zeta(-6) &= (1-2^7)^{-1} \Phi(-6) = 0 \\ \zeta(-7) &= (1-2^8)^{-1} \Phi(-7) = \frac{1}{240} = -\frac{1}{8} \times -\frac{1}{30}\end{aligned}$$

**Exercise 8.2.3**, Page 203, the conjecture here is that according to the list above,  $\zeta(-2n) = 0$  and  $\zeta(-(2n+1)) = -\frac{1}{2(n+1)} \times B_{2(n+1)}$ ,  $n \geq 0$

**Exercise 8.3.2**, Page 209, Assuming  $s > 0$  is real, change variables by  $x = t^s$  in the integral (8.4), this means  $dx = st^{s-1}dt$

$$\begin{aligned}\Gamma(s) &= \int_0^{\infty} e^{-t} t^{s-1} dt \\ &= \int_0^{\infty} e^{-x^{1/s}} \frac{dx}{s}\end{aligned}$$

now change  $s \rightarrow 1/s$

$$\Gamma(1/s) = s \int_0^\infty e^{-x^s} dx$$

using the “normal” Gamma function relationship

$$\Gamma(s+1) = s\Gamma(s)$$

now substituting for  $s \rightarrow 1/s$

$$\begin{aligned}\Gamma(1+1/s) &= \frac{1}{s}\Gamma(1/s) \\ &= \frac{1}{s} \int_0^\infty e^{-x^s} dx\end{aligned}$$

so even for the  $1/s$  Gamma, the relation is still the same

$$\Gamma(1/s+1) = \frac{1}{s}\Gamma(1/s)$$

and so

$$\begin{aligned}\Gamma(5/2) &= \frac{3}{2}\Gamma(3/2) \\ &= \frac{3\sqrt{\pi}}{4}\end{aligned}$$

and  $\Gamma(7/2) = 15\sqrt{\pi}/8$ ,  $\Gamma(9/2) = 105\sqrt{\pi}/16$  and so the general formula is

$$\Gamma(n+1/2) = \frac{(2n-1)!!}{2^n}$$

and now

$$\begin{aligned}\Gamma(1/2) &= 2\Gamma(3/2) \\ &= \sqrt{\pi}\end{aligned}$$

and

$$\begin{aligned}\Gamma(-1/2) &= -2\Gamma(1/2) \\ &= -2\sqrt{\pi}\end{aligned}$$

Note that the goal of this whole exercise is to massage the Gamma function into a Gaussian integral so that we can evaluate it.



**Page 211**, Eq (8.7), there's a typo, it should have been  $s = -(2k-1)$ , it's for negative odd integers that we have simple poles

**Exercise 8.4.1**, Page 212, What are the singular parts for the Laurent expansions of  $\Gamma(s)$ , here we can use (8.5)

$$\Gamma(s) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{1}{s+k} + \int_1^{\infty} e^{-t} t^{s-1} dt$$

so near  $s = -2n+1$  we have

$$\Gamma(s) = \frac{(-1)^{2n-1}}{(2n-1)!} \frac{1}{(s+(2n-1))} + O((s+(2n-1)))$$

and so

$$\Gamma(s)^{-1} = \frac{(2n-1)!}{(-1)^{2n-1}} (s+(2n-1)) + O((s+(2n-1))^2)$$

and around  $s = -2n+1$

$$F(s) + G(s) = \frac{B_{2n}}{(2n)!} \frac{1}{(s+(2n-1))} + O(1)$$

and so

$$\begin{aligned} \zeta(-2n+1) &= \Gamma(s)^{-1} [F(s) + G(s)] \\ &= \left( \frac{(2n-1)!}{(-1)^{2n-1}} (s+(2n-1)) + O((s+(2n-1))^2) \right) \left[ \frac{B_{2n}}{(2n)!} \frac{1}{(s+(2n-1))} + O(1) \right] \\ &= \frac{(-1)^{2n-1} B_{2n}}{2n} + O(1) \\ &= -\frac{B_{2n}}{2n} + O(1) \end{aligned}$$

**Page 213**, lower half of the page, the formula for  $\zeta(2n)$  is given in Chapter 6, Page 154 from the expansion of  $z \cot(z)$

## Chapter 10

**Page 234**, Theorem of Eq (10.5), the *proof* of this theorem is actually not a proof, more like a validation check, what it actually does is that first, it shows that the Mellin transform of  $\Psi(x)$

$$-s\mathcal{M}\Psi(s) = \frac{\zeta'(s)}{\zeta(s)}$$

and then we show the Mellin transform of each term in  $\Psi(s)$  to correspond to the terms in  $\frac{\zeta'(s)}{\zeta(s)}$ , but it did not show how  $\Psi(x)$  was derived in the first place.

**Page 236**, last sentence, this is related to Eq (10.5), we know that  $\Psi(x)$  is real and so

$$\begin{aligned}\overline{\Psi(x)} &= \Psi(x) \\ \rightarrow \sum_{\rho} \frac{x^{\bar{\rho}}}{\bar{\rho}} &= \sum_{\rho} \frac{x^{\rho}}{\rho}\end{aligned}$$

and therefore the zeros  $\rho$  must come in pairs of  $\rho$  and  $\bar{\rho}$

**Page 244 to Page 247**, proof of Theorem, Eq (10.11), again this proof doesn't show how Riemann derived the explicit formula for  $\Pi(x)$ , it only shows that it is consistent by showing the relationship on Page 246

$$\mathcal{E}\Psi(x) = \Pi(x) + \mathcal{E}(\Pi(x) \cdot \log(x))$$

since we know the explicit formula for  $\Psi(x)$ , the first and crucial step is of course showing

$$\log(\zeta(s)) = s\mathcal{M}\Pi(s)$$

of Eq (10.12)

So to summarize, we have two explicit formulae, one from Von Mangoldt and one from Riemann, but the main formula is of course Riemann's explicit formula for  $\Pi(x)$  and  $\pi(x)$ , so here's how everything is connected to the zeta function

- First, we show that the zeta function can be represented as the Hadamard's Product
- Next, we show that the log derivative of the conventional zeta function is just a sum of Von Mangoldt's functions,  $\Psi(x)$
- Using this equality we now substitute the Hadamard's product for the conventional zeta function and show that Von Mangoldt's function  $\Psi(x)$  can be explicitly formulated by taking the "inverse" Mellin transform of the log derivative of Hadamard's product and this is how the Von Mangoldt's function  $\Psi(x)$  is related to the zeros of the zeta function
- With all this set, we now relate the Von Mangoldt's function to the Riemann prime counting function  $\Pi(x) = \sum_k \frac{1}{k} \pi(x^{1/k})$ , the relationship is quite complicated  $\mathcal{E}\Psi(x) = \Pi(x) + \mathcal{E}(\Pi(x) \cdot \log(x))$

- And this is how the Riemann prime counting function is related to the zeros of the zeta function, and using the generalized Mobius inversion we get  $\pi(x)$  as a sum of  $\Pi(x)$  and now we can express the original prime counting function  $\pi(x)$  in terms of the zeros of the zeta function

**Exercise 10.4.1**, Page 246, for  $\alpha = 1, \rho$ , or  $2n$ , show that

$$x \frac{d}{dx} \left( \frac{x^\alpha}{\alpha} \right) = -\text{Li}(x^\alpha) + x \frac{d}{dx} (\text{Li}(x^\alpha) \cdot \log(x))$$

I think it's easier to start with the second term of the RHS

$$\begin{aligned} x \frac{d}{dx} (\text{Li}(x^\alpha) \cdot \log(x)) &= x \log(x) \frac{d}{dx} \text{Li}(x^\alpha) + x \text{Li}(x^\alpha) \frac{d}{dx} \log(x) \\ &= x \log(x) \frac{d}{dx} \left( \int_0^{x^\alpha} \frac{dt}{\log(t)} \right) + x \text{Li}(x^\alpha) \frac{1}{x} \\ &= x \log(x) \frac{1}{\log(x^\alpha)} \cdot \frac{d}{dx} (x^\alpha) + \text{Li}(x^\alpha) \\ &= \frac{\cancel{\alpha \log(x)} x^\alpha}{\log(x^\alpha)} + \text{Li}(x^\alpha) \\ &= x^\alpha + \text{Li}(x^\alpha) \end{aligned}$$

**Page 248**, top of page, on inverting  $\Pi(x)$ , this is almost the same strategy as the one used for Euler's product of the zeta function, first

$$\frac{1}{2}\Pi(x^{1/2}) = \frac{1}{2}\pi(x^{1/2}) + \frac{1}{4}\pi(x^{1/4}) + \frac{1}{6}\pi(x^{1/6}) + \dots$$

subtracting this from  $\Pi(x)$

$$\begin{aligned} \Pi(x) - \frac{1}{2}\Pi(x^{1/2}) &= \left( \pi(x) + \cancel{\frac{1}{2}\pi(x^{1/2})} + \frac{1}{3}\pi(x^{1/3}) + \dots + \cancel{\frac{1}{6}\pi(x^{1/6})} + \dots \right) - \\ &\quad \left( \cancel{\frac{1}{2}\pi(x^{1/2})} + \frac{1}{4}\pi(x^{1/4}) + \cancel{\frac{1}{6}\pi(x^{1/6})} + \dots \right) \end{aligned}$$

removes any term  $\frac{1}{k}\pi(x^{1/k})$  with  $k$  a multiple of two, so we kind of see the pattern, we need to keep subtracting  $\Pi(x^{1/p})$  where  $p$  is prime, but, as the example above shows,  $\frac{1}{6}\pi(x^{1/6})$  is already removed by  $-\Pi(x^{1/2})$ , subtracting  $\Pi(x^{1/3})$  will doubly remove  $\frac{1}{6}\pi(x^{1/6})$ , therefore we need to add this back and note that the first term of  $\Pi(x)$ , *i.e.*  $\pi(x)$  is never removed with these subtractions

$$\pi(x) = \Pi(x) - \frac{1}{2}\Pi(x^{1/2}) - \frac{1}{3}\Pi(x^{1/3}) + \frac{1}{6}\Pi(x^{1/6}) + \dots$$

but it's a bit hard to keep track these subtractions and addition, but this is how we get the inversion. An easier (and more historically accurate) way is to use the generalized Mobius inversion formula as follows

$$G(x) = \sum_{n=1}^{\infty} \frac{1}{n} F(x^{1/n}) \quad \longleftrightarrow \quad F(x) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} G(x^{1/n})$$

but how do we see that this is correct, the easiest way is back substitution

$$\begin{aligned} G(x) &= \sum_{n=1}^{\infty} \frac{1}{n} F(x^{1/n}) \\ &= \sum_{n=1}^{\infty} \frac{1}{n} \left( \sum_{m=1}^{\infty} \frac{\mu(m)}{m} G\left((x^{1/n})^{1/m}\right) \right) \\ &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{\mu(m)}{mn} G(x^{1/(nm)}) \end{aligned}$$

we now rearrange the sum as follow, group all  $G(x^{1/(nm)})$  with the same  $mn = c$  and so

$$\begin{aligned} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{\mu(m)}{mn} G(x^{1/(nm)}) &= \sum_{c=1}^{\infty} G(x^{1/c}) \sum_{d|c} \frac{\mu(d)}{d \left(\frac{c}{d}\right)} \\ &= \sum_{c=1}^{\infty} G(x^{1/c}) \sum_{d|c} \frac{\mu(d)}{c} \\ &= \sum_{c=1}^{\infty} G(x^{1/c}) \frac{1}{c} \sum_{d|c} \mu(d) \\ &= \sum_{c=1}^{\infty} G(x^{1/c}) \frac{1}{c} \delta_{c,1} \\ &= G(x) \end{aligned}$$

so it works, the only thing left to fill is that why  $\sum_{d|c} \mu(d) = \delta_{c,1}$  where  $\delta_{c,1}$  is kronecker delta, there's a proof of this in a direct way in Apostol but let's not do it that way, let start with the Euler's product

$$\sum_{n=1}^{\infty} \frac{1}{n} = \prod_p \left(1 - \frac{1}{p}\right)^{-1}$$

taking  $s = 1$  in the zeta function, taking the reciprocal

$$\begin{aligned}\frac{1}{\sum_{n=1}^{\infty} \frac{1}{n}} &= \prod_p \left(1 - \frac{1}{p}\right) \\ &= \left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{3}\right) \left(1 - \frac{1}{5}\right) \dots \\ &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n}\end{aligned}$$

because each  $\frac{1}{p}$  carries a minus sign in front of it, and there's only one factor for each prime so there won't be anything of the form  $\frac{1}{3^2}$  or  $\frac{1}{5^9}$ , *i.e.* anything with duplicate primes. And since the initial product is over all primes we cover all  $n$  with distinct primes and so the  $\sum_{n=1}^{\infty}$  in the last line and also for a single prime the sign is always minus and for an odd number of prime as well and for an even number of primes the sign will be positive, this is the definition of  $\mu(n)$ . More generally

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \quad \longleftrightarrow \quad \frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$$

the proof of which is exactly the same for the case of  $s = 1$ .

So what does it have to do with  $\sum_{d|c} \mu(d) = \delta_{c,1}$ ? well, let's multiply them

$$\begin{aligned}\left(\prod_p \left(1 - \frac{1}{p^s}\right)^{-1}\right) \left(\prod_p \left(1 - \frac{1}{p^s}\right)\right) &= \sum_{n=1}^{\infty} \frac{1}{n^s} \sum_{m=1}^{\infty} \frac{\mu(m)}{m^s} \\ &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{\mu(m)}{(mn)^s} \\ &= \sum_{c=1}^{\infty} \frac{a_n}{c^s} = \sum_{c=1}^{\infty} \sum_{d|c} \frac{\mu(d)}{c^s} \\ &= \sum_{c=1}^{\infty} \frac{\left(\sum_{d|c} \mu(d)\right)}{c^s}\end{aligned}$$

where  $a_n = 1$  where  $n = 1$  and  $a_n = 0$  otherwise, we now only need to compare the coefficient for each  $1/c^s$  and in this case it's very easy

$$\begin{aligned}a_1 &= \sum_{d|1} \mu(d) \\ &\rightarrow 1 = \mu(1)\end{aligned}$$

which is consistent and for all other  $a_{n>1} = 0 = \sum_{d|c>1} \mu(d) = 0$  which means that  $\sum_{d|c} \mu(d) = \delta_{c,1}$

**Page 248**, middle of page, “a consequence of Prime Number Theorem says that  $\sum_n \mu(n)/n = 0$ ”, this result can be obtained without the prime number theorem, it is in fact a very clever trick by Mobius as follows.

With the above trick Mobius showed that  $\sum_{n=1}^{\infty} \frac{\mu(n)}{n} = 0$ , here’s how. First, just like above the following relationship is also true

$$G(x) = \sum_{n=1}^{\infty} a_n F(x) \quad \longleftrightarrow \quad F(x) = \sum_{n=1}^{\infty} b_n G(x)$$

express  $G(x)$ ,  $F(x)$  in terms of power series

$$\begin{aligned} \sum_{j=0}^{\infty} g_j x^j &= \sum_{n=1}^{\infty} a_n \left( \sum_{m=1}^{\infty} b_m \left( \sum_{i=0}^{\infty} g_i x^i \right) \right) \\ &= \sum_{i=0}^{\infty} \left( \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_n b_m \right) g_i x^i \\ &= \sum_{i=0}^{\infty} \left( \sum_{c=1}^{\infty} \sum_{d|c} a_{\frac{c}{d}} b_d \right) g_i x^i \end{aligned}$$

again comparing coefficients of same power of  $x$  on both sides we get that

$$\begin{aligned} 1 &= \sum_{c=1}^{\infty} \sum_{d|c} a_{\frac{c}{d}} b_d = \sum_{c=1}^{\infty} \sum_{d|c} a_c b_{\frac{c}{d}} \\ &\rightarrow \sum_{d|c} a_c b_{\frac{c}{d}} = \delta_{c,1} \end{aligned}$$

this can be immediately generalized into

$$G(x) = \sum_{n=1}^{\infty} a_n F(x^n) \quad \longleftrightarrow \quad F(x) = \sum_{n=1}^{\infty} b_n G(x^n)$$

because by the same token as before (but not expanding each function into power series)

$$\begin{aligned}
G(x) &= \sum_{n=1}^{\infty} a_n \left( \sum_{m=1}^{\infty} b_m G((x^n)^m) \right) \\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_n b_m G(x^{nm}) \\
&= \sum_{c=1}^{\infty} G(x^c) \sum_{d|c} a_d b_{\frac{c}{d}} \\
&\rightarrow \sum_{d|c} a_d b_{\frac{c}{d}} = \delta_{c,1}
\end{aligned}$$

with the same conclusion (and again we are rearranging the double sum into a sum for a fixed exponent for  $x$  and a sum for the factors of  $c$ ). This is more or less the generalization of our previous case involving  $\Pi(x)$  and  $\pi(x)$ , note that since  $\sum_{c|d}$  involves all factors of  $c$  we can swap  $a_{\frac{c}{d}} b_d \leftrightarrow a_d b_{\frac{c}{d}}$ . The series  $\sum_{n=0}^{\infty} \frac{f_n}{n^s}$  is normally called Dirichlet series in the literature.

Mobius then cleverly applied this good finding on our beloved geometric series by setting  $a_n = 1$  and  $b_n = \mu(n)$  for all  $n$

$$\frac{x}{1-x} = \sum_{n=1}^{\infty} x^n \quad \longleftrightarrow \quad x = \sum_{n=1}^{\infty} \mu(n) \frac{x^n}{1-x^n}$$

again since as shown above that  $\sum_{d|c} \mu(d) = \delta_{c,1}$ . He then noticed that if  $x \rightarrow 1$  we have  $1-x = \epsilon$  and

$$\begin{aligned}
1-x^n &= 1-(1-\epsilon)^n \\
&= 1-(1-n\epsilon + O(\epsilon^2)) \\
&= n\epsilon + O(\epsilon^2)
\end{aligned}$$

where we just expanded  $(1-\epsilon)^n$  using the binomial theorem, therefore

$$\begin{aligned}
\frac{x^n}{1-x^n} &= \frac{1+O(\epsilon)}{n\epsilon + O(\epsilon^2)} \\
&= \frac{1}{n\epsilon} (1+O(\epsilon)) \left( \frac{1}{1+O(\epsilon)} \right) \\
&= \frac{1}{n\epsilon} (1+O(\epsilon)) (1+O(\epsilon)) \\
&= \frac{1}{n\epsilon} + O(1)
\end{aligned}$$

therefore

$$x = \sum_{n=1}^{\infty} \mu(n) \frac{x^n}{1 - x^n}$$

$$1 + \epsilon = \sum_{n=1}^{\infty} \mu(n) \left( \frac{1}{n\epsilon} + O(1) \right)$$

multiplying both sides by  $\epsilon$

$$\epsilon + \epsilon^2 = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} + O\left(\epsilon \sum_{n=1}^{\infty} \frac{\mu(n)}{n}\right)$$

the thing to note here is that the above is true for any  $\epsilon$ , this means that  $\sum_{n=1}^{\infty} \frac{\mu(n)}{n}$  is finite, but if it's finite since  $\epsilon \rightarrow 0$  the only choice is  $\sum_{n=1}^{\infty} \frac{\mu(n)}{n} = 0$

**Page 249 and Page 250**, proof of Theorem, Eq (10.15), to get the last line on Page 249

$$\begin{aligned} \operatorname{Re}\{\log(z)\} &= \operatorname{Re}\{\log(re^{i\theta})\} \\ &= \operatorname{Re}\{\log(r) + i(\theta + 2\pi n)\} \\ &= \operatorname{Re}\{\log(r)\} \\ &= \operatorname{Re}\{\log(|z|)\} \\ &= \log(|z|) \end{aligned}$$

the actual kick here is on Page 250, where “suppose that  $t$  is . . . a zero of the form  $1 + i\gamma$ ” and it leads to a contradiction so this covers the case for  $\operatorname{Re}(\rho) = 1$ , how about the other boundary  $\operatorname{Re}(\rho) = 0$ , since we have shown that there's no zero with  $\operatorname{Re}(\rho) = 1$ , by symmetry,  $\Lambda(1 - s) = \Lambda(s)$ , there is no zero with  $\operatorname{Re}(\rho) = 0$  either.

Also, here  $\epsilon \rightarrow 0$  from  $> 1$ , the other thing is that we try to take the limit  $\epsilon \rightarrow 1$  from below the argument for the Big Oh still works so we will still have a contradiction except for the fact that  $\zeta(s)$  is only defined for  $\operatorname{Re}(s) > 1$  so the limit must be taken from above

#### **Interlude 4**

**Page 258**, Hensel's Lemma, this is my alternative proof, also I found the statement that  $x_{j+1}$  to be unique is not quite right, even for the example of  $a = -1$  and  $p = 5$ ,  $x_1 = 2$ , I can get  $x_2 = 32$  and  $32^2 = 1024 = -1 + 25 \cdot 41$ , in his example his  $x_2 = 7$  with



$7^2 = 49 = -1 + 25 \cdot 2$  so I believe what he meant by unique is that it is unique modulo  $p^{j+1}$ .

Now here is my proof, say there's a solution to  $x_j^2 \equiv a \pmod{p^j}$ , if there's another solution at all, it must be of the form

$$\begin{aligned} x_{j+1} &= x_j + mp^j \\ \rightarrow x_{j+1}^2 &= (x_j + mp^j)^2 \\ &= x_j^2 + 2x_jmp^j + m^2p^{2j} \\ &= (a + np^j) + 2x_jmp^j + m^2p^{2j} \end{aligned}$$

for the RHS to be  $a \pmod{p^{j+1}}$ ,  $n + 2x_jm$  must be a multiple of  $p$  and this is actually guaranteed by Bezout's lemma, here's how. What we need is

$$n + 2x_jm = pw$$

if  $\gcd(2x_j, p) = 1$  then we are done since Bezout guarantee that there would be an  $m'$  and a  $w'$  such that

$$1 = (2x_j)m' + pw'$$

multiplying both sides by  $n$  we get our solution, but what if  $\gcd(2x_j, p) = d > 1$ ? since  $p$  is prime, the only possible values for  $d$  is  $d = p = 2$  or  $d = p$ , first let's take  $d = p \neq 2$ , this means that  $p|x_j$  since  $p \nmid 2$

$$\begin{aligned} x_j^2 &= a + p^j n \\ p(px_j'^2) &= a + p(p^{j-1}n) \end{aligned}$$

but this means that  $a \equiv 0 \pmod{p}$ , which is not allowed by the assumption of the lemma, *i.e.*  $a \not\equiv 0 \pmod{p}$  right at the beginning of the lemma, so the only other alternative is  $d = p = 2$ , but this case is excluded because the lemma only concerns odd  $p$  :) but why won't  $p = 2$  work? if you look back at our crucial equation

$$n + 2x_jm = 2w$$

this means that  $n$  has to be even and so it's not guaranteed that as long as  $x_j^2 \equiv a \pmod{2^j}$  we'll have another solution  $x_{j+1}^2 \equiv a \pmod{2^{j+1}}$ , only on certain cases this will hold, but the case for odd  $p$  will *always* hold.

Coming back to uniqueness modulo  $p^{j+1}$ , from Bezout's lemma we know that if  $\gcd(a, b) = d$  there exist  $x, y$  such that  $ax + by = d$ , it so happens that these  $x, y$  are not unique, any other

$$x' = x + c \cdot \frac{b}{d} \quad y' = y - c \cdot \frac{a}{d}$$

also work where  $c$  is any integer. Applying it to our case, what we have is

$$\begin{aligned} 1 &= 2x_j(-m') + pw' \\ \rightarrow n &= 2x_j(-m) + pw \end{aligned}$$

and so any  $-M' = -m' - c \cdot p \rightarrow -M = (-m - ncp)$  is also a solution for any integer  $c$ , note that this  $M$  plays its role in

$$\begin{aligned} x_{j+1} &= x_j + Mp^j \\ &= x_j + mp^j + ncp^{j+1} \end{aligned}$$

and so it differs from another  $x_{j+1}$  by a multiple of  $p^{j+1}$