

# Harold Edwards Riemann's zeta function

Stefanus Koesno<sup>1</sup>

<sup>1</sup> Somewhere in California

San Jose, CA 95134 USA

(Dated: December 22, 2017)

Abstract

**Page 4.** Footnote,

$$\int_2^L \frac{1}{\log t} dt \sim \frac{L}{\log L}$$

We might be able to do this using repeated IBPs, with  $u = 1/\log t$  and  $dv = dt$

$$\begin{aligned} \int_2^L \frac{1}{\log t} dt &= \left. \frac{t}{\log t} \right|_2^L + \int_2^L \frac{1}{\log^2 t} dt \\ &= \frac{L}{\log L} + C_1 + \int_2^L \frac{1}{\log^2 t} dt \end{aligned}$$

we can either repeat the process or do the following

$$\begin{aligned} \int_2^L \frac{1}{\log^2 t} dt &< \int_2^{L^{1/2}} \frac{1}{\log^2 2} dt + \int_{L^{1/2}}^L \frac{1}{\log^2 L^{1/2}} dt \\ &< \frac{L^{1/2} - 2}{\log^2 2} + \frac{L - L^{1/2}}{\log^2 L^{1/2}} \\ &< \frac{L^{1/2}}{\log^2 2} + \frac{L}{\log^2 L^{1/2}} \end{aligned}$$

such that

$$\int_2^L \frac{1}{\log t} dt < \frac{L}{\log L} + C_1 + C_2 L^{1/2} + \frac{L}{\frac{1}{4} \log^2 L}$$

if we divide both sides by  $L/\log L$  and take the limit  $L \rightarrow \infty$

$$\begin{aligned} \lim_{L \rightarrow \infty} \frac{\int_2^L \frac{1}{\log t} dt}{L/\log L} &< 1 + 0 + \frac{\log L}{L^{1/2}} + \frac{1}{\frac{1}{4} \log L} \\ &< 1 \end{aligned}$$

recall that  $\log x$  is smaller than any  $x^{1/k}$ . But, on the other hand,

$$\begin{aligned} \int_2^L \frac{1}{\log t} dt &> \int_2^L \frac{1}{\log L} dt \\ &> \frac{L - 2}{\log L} \\ \lim_{L \rightarrow \infty} \frac{\int_2^L \frac{1}{\log t} dt}{L/\log L} &> 1 \end{aligned}$$

thus

$$\int_2^L \frac{1}{\log t} dt \sim \frac{L}{\log L}$$

A small note, this asymptotic equality is not like a normal equality, it is more like a Big Oh, say

$$f(L) \sim L + L^{1/2}$$

we can also say

$$f(L) \sim L$$

because  $L^{1/2}/L$  approaches zero when  $L \rightarrow \infty$ , so if we divide both sides by  $L$  and then take  $L \rightarrow \infty$  we get the previous asymptotic expression.

**Page 8.** The Gamma function shown here  $\Gamma(s)$  is valid for all  $s > -1$ , how is this so as

$$\int_0^\infty e^{-x} x^s dx$$

blows up for  $s < 0$  when  $x \rightarrow 0$ . The justification is as follows

$$e^{-x} x^s \leq x^s$$

for all  $x > 0$  no matter what  $s$  is so the integral is also capped by ( $s > -1$ )

$$\begin{aligned} \int_0^1 e^{-x} x^s dx &\leq \int_0^1 x^s dx = \frac{1}{s+1} x^{s+1} \Big|_0^1 \\ &= \frac{1}{s+1} \end{aligned}$$

Another way of looking at it is

$$\int_0^\infty e^{-x} x^n dx = e^{-x} \frac{1}{n+1} x^{n+1} + \int_0^\infty e^{-x} \frac{1}{n+1} x^{n+1} dx$$

so as long as  $n > -1$  we are fine.

**Page 8.** On the definition of factorial, Euler introduced the integral notation

$$n! = \int_0^\infty e^{-x} x^n dx$$

the  $e^{-x}$  is there to make sure everything is well behaved at  $\infty$ , the  $x^n$  is there to be IBP'd (differentiated) multiple times so that we get the factor  $n!$ , they also make sure that the boundary terms are zero.

But Euler had other ideas too, instead of assigning  $x^n \rightarrow u$  and differentiating it every IBP, we can instead integrate it, *i.e.* assign  $x^n dx \rightarrow dv$  like so (boundary terms are always zero and also changing  $n \rightarrow s$ )

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

repeating it  $N$  times we get

$$\int_0^\infty e^{-x} x^s dx = \frac{1}{(s+1)(s+2)\dots(s+N)} \int_0^\infty e^{-x} x^{(s+N)} dx$$

and this way we can immediately see the poles at  $s = -1, -2, \dots$ . We now expand the integral in the RHS above in an interesting way, please do not do IBP by differentiating  $x^{(s+n)}$  that will undo everything we've just done. There is, however, a curious way to expand it which only works when  $N \rightarrow \infty$ , we write

$$(s+N)! = N! \cdot (1+N)(2+N)\dots(s+N)$$

since  $s$  is just a (fixed) parameter, it's finite, and so as we take  $N \rightarrow \infty$

$$\lim_{N \rightarrow \infty} (1+N)(2+N)\dots(s+N) \sim (1+N)(1+N)\dots(1+N) = (1+N)^s$$

and thus

$$\lim_{N \rightarrow \infty} \int_0^\infty e^{-x} x^{(s+N)} dx = N!(1+N)^s$$

So, Euler's definition of a factorial was

$$s! = \lim_{N \rightarrow \infty} \frac{1 \cdot 2 \dots N}{(s+1)(s+2)\dots(s+N)} (1+N)^s$$

although whether he got it from the integral above or not I don't know.

**Page 9.**

$$\begin{aligned} \int_0^\infty e^{-nx} x^{s-1} dx &= \frac{(s-1)!}{n^s} \\ \rightarrow \sum_{n=1}^\infty \int_0^\infty e^{-nx} x^{s-1} dx &= \sum_{n=1}^\infty \frac{(s-1)!}{n^s} \\ &= (s-1)! \sum_{n=1}^\infty \frac{1}{n^s} \end{aligned}$$

so as long as  $s > 1$  the sum of the integral converges, now we want to see the other way, summing inside the integral. A few points

- Swapping sum and integral is always OK, as long as the sum is finite, for infinite sums what we are interchanging are actually limit and integral because  $\sum^\infty = \lim_{N \rightarrow \infty} \sum^N$
- Geometric series formula is only valid if the sum converges, why? because we start with

$$A = 1 + r + r^2 + \dots$$

$$rA = r + r^2 + r^3 + \dots$$

we then subtract  $A - rA = 1 \rightarrow A = 1/(1 - r)$  and so on, the problem is if  $A$  is divergent we are then doing  $\infty - \infty$  which might not be equal to one, so if we try to do  $A = 1 + 2 + 2^2 + \dots$  using the formula we get  $-1$  which is not just wrong but also not divergent (although if you read Stoppa's book you can actually do this and it's called Abel sum).

So the point is that we can't just apply the formula and then take the limit  $r \rightarrow a$ , we can only do this if  $a = 1$ .

Swapping integral and sum

$$\begin{aligned} \int_0^\infty \sum_{n=1}^\infty e^{-nx} x^{s-1} dx &= \int_0^\infty \left( \sum_{n=1}^\infty e^{-nx} \right) x^{s-1} dx \\ &= \int_0^\infty \left( \frac{1}{1 - e^{-x}} - 1 \right) x^{s-1} dx \\ &= \int_0^\infty \left( \frac{e^{-x}}{1 - e^{-x}} \right) x^{s-1} dx \\ &= \int_0^\infty \frac{x^{s-1}}{e^x - 1} dx \end{aligned}$$

the only problem we have here is when  $e^{-x} \geq 1$  and this only happens when  $x = 0$ , so we only have one point of contention, literally, by the way, this is similar to the situation for

$$\int_0^1 \frac{1}{\sqrt{x}} dx = 2$$

even though it blows up at  $x = 0$ , we do this by

$$\int_0^1 \frac{1}{\sqrt{x}} dx = \lim_{a \rightarrow 0} \int_a^1 \frac{1}{\sqrt{x}} dx$$

in our case the integral is also OK because  $e^x - 1 \geq x$  for  $x \geq 0$  and so

$$\begin{aligned} \int_0^1 \frac{x^{s-1}}{e^x - 1} dx &\leq \int_0^1 x^{s-2} dx \\ &\leq \frac{1}{s-1} x^{s-1} \Big|_0^1 \\ &\leq \frac{1}{s-1} \end{aligned}$$

since  $s > 1$ . Note that we can't just take the limit ( $s > 1$ )

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{x^{s-1}}{e^x - 1} &= \lim_{x \rightarrow 0} \frac{(s-1)x^{s-2}}{e^x} \\ &= \infty, \quad \text{since } s > 1. \end{aligned}$$

so at zero the integrand still blows up but the integral is OK.

**Page 10.** Why the minus sign? the integrand is  $\frac{(-x)^s}{e^x - 1}$ , there's a minus sign in front of  $x$ . This is due to the choice of contour. Note that we are going from  $+\infty$  to left all the way to 0 and then to the right all the way to  $+\infty$ .

If there's no minus sign and this is the cut we choose, then

$$\begin{aligned} x^s &= e^{s \log x + 0 \cdot \pi i}, & \text{going from } +\infty \text{ to } 0 \\ x^s &= e^{s \log x + 2 \cdot \pi i}, & \text{going from } 0 \text{ to } +\infty \end{aligned}$$

and when we add the two integrals we get zero

$$e^{s \log x + 0 \cdot \pi i} - e^{s \log x + 2 \cdot \pi i} = 0$$

so if we don't want the minus sign we need to choose the other cut, coming from  $-\infty$  to zero and back to  $-\infty$  so that the phases will be  $\pi i$  and  $-\pi i$ .

**Page 12.** Order of limit taking around the small semi circle, Bernoulli's number, the integral ( $x = \delta e^{i\theta}$ )

$$\int_0^{2\pi} x^{m-n-1} d\theta = \frac{\delta^{m-n-1}}{m-n-1} (e^{(m-n-1) \cdot 2\pi \cdot i} - e^{(m-n-1) \cdot 0 \cdot i})$$

if  $m-n-1 < 0$  then the delta term will blow up, but the  $e$  difference is zero no matter what, so this means we need to calculate the  $e$  difference first before taking the limit.

**Page 13.** Compact here means closed and bounded.

The whole integral is zero thanks to a corollary of Goursat theorem which says that

$$\oint_C f(z)dz = \sum_{k=1}^n \oint_{C_k} f(z)dz$$

where the contours  $C_k$ 's are all inside of  $C$ .

**Page 13.** For the integral

$$\int_{|y|=\epsilon} (-2\pi in - y)^{s-1} \frac{y}{e^y - 1} \cdot \frac{dy}{y}$$

we rewrite as

$$\int_{|y|=\epsilon} \frac{f(y)}{y-0} dy, \quad f(y) = (-2\pi in - y)^{s-1} \frac{y}{e^y - 1}$$

we need to remember that  $f(y)$  has no pole when  $y \rightarrow 0$ , so the value of the integral is just  $2\pi i f(0) = 2\pi i (-2\pi in)^{s-1}$ . Note that

$$\lim_{y \rightarrow 0} \frac{y}{e^y - 1} = \lim_{y \rightarrow 0} \frac{1}{e^y} = 1$$

after L'Hospital.

**Multiple Pages.** A note on analytic continuation. For gamma function, integral is valid only for  $s > 1$  but once we get the final form it applies for all  $s$ . It is even in the derivation of the analytic continuation

$$\int_0^\infty e^{-x} x^s dx = \int_0^1 e^{-x} x^s dx + \int_1^\infty e^{-x} x^s dx$$

for the first integral we expand  $e$

$$\begin{aligned} \int_0^1 e^{-x} x^s dx &= \int_0^1 \sum_{n=0}^\infty \frac{(-x)^n}{n!} x^s dx \\ &= \sum_{n=0}^\infty (-1)^n \int_0^1 \frac{x^{n+s}}{n!} dx \end{aligned}$$

now this integral is not valid for  $s \leq -1$ , however, at this stage we still assume  $s > -1$ , once we have the full result

$$\int_0^\infty e^{-x} x^s dx = \sum_{n=1}^\infty (-1)^{n-1} \frac{1}{n!(s+n)} + \int_1^\infty e^{-x} x^s dx$$

we say that this result is valid for all  $s$ , the derivation is not valid for all  $s$  but the final result is, that's what analytic continuation is.

The same goes with the zeta function, in the derivation on Page 10, we take the small semi circle  $\delta \rightarrow 0$  only for  $s > 1$ , once we get the final form we say it applies for all  $s$  but of course for  $s \leq 1$  we cannot take  $c \rightarrow 0$  when evaluating the integral, otherwise the integral around that semi circle blows up and we don't have an entire function anymore.

Another remark is that the radius of that small circle must be  $0 < \delta < 2\pi$ , because at  $2\pi i$  we get an infinity from  $1/(e^x - 1)$ .

**Page 15.**

$$\frac{1 + 2\psi(x)}{1 + 2\psi(1/x)} = \frac{1}{\sqrt{x}}$$

for small  $x$

$$\begin{aligned}\psi(1/x) &= \sum e^{-n^2\pi/x} \\ \lim_{x \rightarrow 0} \psi(1/x) &= \sum e^{-n^2\pi\infty} \\ &= 0\end{aligned}$$

so for small  $x$  we have

$$\begin{aligned}\lim_{x \rightarrow 0} \frac{1 + 2\psi(x)}{1 + 2\psi(1/x)} &= 1 + 2\psi(x) = \frac{1}{\sqrt{x}} \\ &\rightarrow \psi(x) = \frac{1}{2} \left( \frac{1}{\sqrt{x}} - 1 \right)\end{aligned}$$

**Page 24.** Fourier transform, I thought I was an expert LOL well, not compared to Riemann

$$\int_0^\infty J(x)x^{-s-1}dx = \int_0^\infty J(x)x^{-s}\frac{dx}{x}$$

every time you see  $\frac{dx}{x}$  you should immediately scream  $\log x$  because  $d(\log x) = \frac{dx}{x}$ , let  $w = \log x$

$$\begin{aligned}\int_0^\infty J(x)x^{-s}\frac{dx}{x} &= \int_{-\infty}^\infty J(e^w)e^{-ws}dw \\ &= \int_{-\infty}^\infty J(e^w)e^{-w(\sigma+it)}dw \\ &= \int_{-\infty}^\infty [J(e^w)e^{-w\sigma}]e^{-itw}dw\end{aligned}$$



**Page 21.** Something I'm not quite sure, "to prove the absolute convergence of  $\prod(1+a_i)$  it suffices to prove the absolute convergence of the sum  $\sum a_i$ ". To see this, take the log of the product

$$\log \prod(1+a_i) = \sum \log(1+a_i)$$

But we know that  $\log(1+x) < x$ , so if  $\sum a_i$  converges then by comparison test the log of the product as well, but if it converges then it converges to a constant and the infinite product is just the exponentiation of a finite constant and thus converges as well.

**Page 25.** Another useful application of IBP, if the integral doesn't converge (in this case due to oscillations), do IBP :)

**Page 26.** about the ambiguity of  $\log(1-s/\rho)$ , especially for  $\rho$  which are not large, this is because the convention we use is given in Page 27, *i.e.*  $\log(z)$  is defined for all  $z$  other than real  $z \leq 0$  and that for real  $z$ ,  $\log(z)$  is real.

So if  $\rho$  is smaller than  $s$  we can have a situation where  $(1-s/\rho)$  is real and negative, but  $\log(s-\rho)$  cannot be real and negative when we set  $a > 1$  since we know that the zeros are all in the critical strip (see Section 1.9) and for  $\log(-\rho)$ , since  $\rho$  is never real we are OK since the argument to the log function is never real and negative.

**Page 27.** On the Fourier inversion of  $\frac{1}{s-\beta}$ , note that the limit of the integration has to go from  $-\infty$  to  $\infty$ , here the function is defined only from 1 to  $\infty$ , what this means is that from  $-\infty$  to 1 the value of the function is zero.

**Page 28.** Changing variable of integration from the 0 to 1 plus upper semicircle plus 1 to  $x$  into  $u = \log t$ , we see something quite strange. If we change  $[0, 1]$  using  $u = \log t$  we should get  $[-\infty, 0]$  but instead we get  $[-\infty + i\delta, 0 + i\delta]$ , in the book the integration is split into  $[i\delta - \infty, i\delta + \log x]$ .

This is actually more in line to shifting the whole original contour to be slightly above the positive  $x$  axis, rather than doing an upper semicircle. If we push it up the argument of the log has to be positive thus the addition of  $i\delta$ .

**Page 29.** Here we show that  $H(\beta) \rightarrow 0$  as  $\tau \rightarrow \infty$ , but since  $H(\beta)$  is extremely similar to  $F(\beta)$ , the difference is only in the

$$\log[1 - (s/\beta)] \longleftrightarrow \log[(s/\beta) - 1]$$

so the key difference here is that for  $H(\beta)$ , when we take  $\tau \rightarrow \infty$ , we get  $\log[1 - (s/\beta)] \rightarrow \log(1)$  while for  $F(\beta)$  we get  $\log(-1)$ . And here we see some application of Lebesgue integral that allows us to interchange limit and integration.

Looking back, this might be the exact reason why we introduce  $H(\beta)$  in the first place, we know we need to take  $\tau \rightarrow \infty$  (recall that  $\beta = \sigma + i\tau$ ), but doing this in  $F(\beta)$  causes something like  $\log(-1)$ , if we reverse the sign inside the log, we might be able to get something, but first we need to see if this log reversed function is the same as our original function, the good news is that in this case they are only separated by a constant  $i\pi$ .

**Page 30.** Pairing  $\rho$  and  $1 - \rho$ , in case you forget, this was discussed back on Page 21.

**Page 31.**  $-\zeta(0) = \frac{1}{2}$ , here we are using Abel's sum or from the Bernoulli number formula thingy

$$\zeta(2n) = (-1)^{n+1} \frac{(2\pi)^{2n}}{2} \frac{B_{2n}}{2n!}$$

and setting  $n = 0$ , a more interesting way is using Abel's sum but not on  $\zeta(0)$  directly because

$$\begin{aligned} \zeta(0) &= 1 + 1 + 1 + \dots \\ &= \frac{1}{1 - 1} \\ &= \infty \end{aligned}$$

instead we do Abel's summation on

$$\Phi(s) = (1 - 2^{1-s})\zeta(s) = 1 - \frac{1}{2^s} + \frac{1}{3^s} - \frac{1}{4^s} + \frac{1}{5^s} \dots = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^s}$$

the zero  $(1 - 2^{1-s})$  at  $s = 1$  cancels the pole of  $\zeta(s)$  at  $s = 1$ . Now, for  $\Phi(s = 0)$  we get

$$\begin{aligned} \Phi(0) &= 1 - 1 + 1 - 1 + 1 \dots \\ &= \frac{1}{1 + 1} \\ &= \frac{1}{2} \end{aligned}$$

instead of  $\frac{1}{1-1}$  (again, the pole at  $s = 1$  was canceled), thus

$$\begin{aligned} \zeta(0) &= (1 - 2^1)^{-1} \Phi(0) \\ &= (-1) \frac{1}{2} \end{aligned}$$

**Page 42.** The curious thing about the application of Jensen's theorem here is that instead of  $\xi(0)$  we get  $\xi(1/2)$ , the thing to note is that the radius used in Jensen's Theorem is  $|z| \leq R$ , here it is  $|s - \frac{1}{2}| \leq R$ , so the prerequisite is not met :)

We just need to do a simple change of variable  $w = s - \frac{1}{2}$  and thus  $|w| \leq R$

$$\begin{aligned}\xi(s) &= \xi(w + \frac{1}{2}) \\ \rightarrow f(w) &= \xi(w + \frac{1}{2})\end{aligned}$$

and we do Jensen's Theorem on  $f(w)$  instead, in this case

$$f(w=0) = \xi\left(\frac{1}{2}\right)$$

and the zeros of  $f(w)$  are now located at  $\rho - \frac{1}{2}$  where  $\xi(\rho) = 0$ , all done :)

**Page 42.** This statement, "Since all but a finite number of roots  $\rho$  satisfy the inequality"

$$\frac{1}{|\rho(1-\rho)|} = \frac{1}{|(\rho - \frac{1}{2})^2 - \frac{1}{4}|} < \frac{1}{|\rho - \frac{1}{2}|^2}$$

let  $\rho - \frac{1}{2} = a + ib$  then

$$\begin{aligned}|\rho - \frac{1}{2}|^2 - \frac{1}{4} &= (a^2 + b^2 - \frac{1}{4}) + 4a^2b^2 \\ |(\rho - \frac{1}{2})^2|^2 &= (a^2 + b^2)^2\end{aligned}$$

taking the difference

$$|(\rho - \frac{1}{2})^2 - \frac{1}{4}|^2 - |(\rho - \frac{1}{2})^2|^2 = \frac{(8a^2 - 1)(8b^2 - 1)}{16}$$

so for  $a^2 > \frac{1}{8}$  and  $b^2 > \frac{1}{8}$  we get the inequality in the book, thus only for those roots whose  $\Re(\rho - \frac{1}{2})^2 < \frac{1}{8}$  or  $\Im(\rho - \frac{1}{2})^2 < \frac{1}{8}$  we don't get the inequality in the book.

**Page 43.** Note that this Theorem and the preceding one all rely on the fact that  $R$  is sufficiently large, the inequality

$$n(R) \leq 3R \log R$$

doesn't work for all  $R$ , only for  $R$  sufficiently large.

Also this is about seeing how dense the roots of  $\xi(s)$  are, note that for example  $\sum 1/p$ , although is not as dense as  $\sum 1/n$ , still diverges.

**Page 44.** Dealing with  $u_R(s)$ , we invoke properties of harmonic functions, first a harmonic function is just a holomorphic function because it is a function that satisfies Cauchy-Riemann conditions, or in other words it satisfies Laplace's equation

$$\partial_x^2 f + \partial_y^2 f = 0$$

and because of this very reason a harmonic function cannot have a max or min within a compact domain, because if  $x$  is a max/min then it's second derivative in  $x$  must be non-zero but due to Laplace's equation above the second derivative in the  $y$  has the opposite sign and therefore it is a saddle point.

In this case the domain is not really compact because we exclude the zeros of  $\xi$ , this is why the book explains that near this point the value is near  $-\infty$ , so there's no chance to have a max point near these points and the max points should be at the boundary.

**Page 46.**  $F(s)$  is an even function of  $s - \frac{1}{2}$ . We know that  $\xi(s)$  is symmetric under  $s \rightarrow 1 - s$ , thus

$$\frac{\xi(1-s)}{\prod \left[ 1 - \frac{(1-s) - \frac{1}{2}}{\rho - \frac{1}{2}} \right]} = \frac{\xi(s)}{\prod \left[ 1 - \frac{\frac{1}{2} - s}{\rho - \frac{1}{2}} \right]}$$

However, if  $\rho$  is a zero then  $1 - \rho$  is also a zero, thus

$$\frac{\xi(s)}{\prod \left[ 1 - \frac{\frac{1}{2} - s}{\rho - \frac{1}{2}} \right]} = \frac{\xi(s)}{\prod \left[ 1 - \frac{\frac{1}{2} - s}{\frac{1}{2} - \rho} \right]} = \frac{\xi(s)}{\prod \left[ 1 - \frac{-(s - \frac{1}{2})}{-(\rho - \frac{1}{2})} \right]} = \frac{\xi(s)}{\prod \left[ 1 - \frac{s - \frac{1}{2}}{\rho - \frac{1}{2}} \right]}$$

which is all good and all but this is a symmetry under  $s \rightarrow 1 - s$ , to make it into a symmetry under  $s - \frac{1}{2}$  we do the following

$$\xi(s) = \xi\left(\left(s - \frac{1}{2}\right) + \frac{1}{2}\right)$$

so that under  $s - \frac{1}{2} \rightarrow \frac{1}{2} - s$  we get

$$\begin{aligned} \xi\left(\left(s - \frac{1}{2}\right) + \frac{1}{2}\right) &\rightarrow \xi\left(\left(\frac{1}{2} - s\right) + \frac{1}{2}\right) \\ &= \xi(1 - s) \\ &= \xi(s) \\ &= \xi\left(\left(s - \frac{1}{2}\right) + \frac{1}{2}\right) \end{aligned}$$

**Page 47.** Footnote, first the validity of the product formula for  $\sin \pi s$ , note that it is easier to show the validity of the product formula for  $\frac{\sin \pi s}{\pi s}$ , using the Theorem in Section 2.7, we see that  $\left(\frac{\sin \pi s}{\pi s}\right) / \left(\prod \left[1 - \frac{s^2}{n^2}\right]\right)$  is an even function and it does grow more slowly than  $|s|^2$ . Thus

$$\begin{aligned}\frac{\sin \pi s}{\pi s} &= c \prod \left(1 - \frac{s}{n}\right) \left(1 + \frac{s}{n}\right) \\ &= c \prod \left(1 - \frac{s^2}{n^2}\right)\end{aligned}$$

where we have paired the roots to ensure the infinite product is convergent. Next

$$\frac{(\sin \pi s)/(\pi s)}{(\sin \pi 0)/(\pi 0)} = \frac{\prod \left(1 - \frac{s^2}{n^2}\right)}{1}$$

But  $(\sin \pi 0)/(\pi 0) = 1$  thus

$$\frac{\sin \pi s}{\pi s} = \prod \left(1 - \frac{s^2}{n^2}\right)$$

The second footnote is not related to this one, it just says that everything in that page has been rigorously proven, including the product formula for  $\sin$ , the only thing not proven is the relationship between (2) and (3).

**Page 50.** Footnote, logarithmic differentiation of  $\zeta(s)$

$$\begin{aligned}\frac{d(\log \zeta(s))}{ds} &= \frac{\zeta'(s)}{\zeta(s)} = \frac{d}{ds} \left( \log \prod (1 - p^{-s})^{-1} \right) \\ &= - \sum_p \frac{d}{ds} \log(1 - p^{-s}) \\ &= - \sum_p \frac{\log p}{1 - p^{-s}}\end{aligned}$$

then

$$\frac{1}{1 - p^{-s}} = 1 + p^{-s} + p^{-2s} + \dots$$

so

$$- \sum \frac{\log p}{1 - p^{-s}} = - \sum_p \log p \sum_k p^{-ks}$$

and we are done :)

**Page 51.** formula (2) of Section 1.14 is

$$\frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{1}{s-\beta} y^s ds = \begin{cases} y^\beta & y > 1 \\ 0 & y < 1 \end{cases}$$

**Page 52.** Bottom page, about the middle equation, say  $\beta = A + iB$ , here we don't take the limit  $h \rightarrow \infty$  yet

$$\begin{aligned} \int_{a-\beta-ih}^{a-\beta+ih} &= \int_{a-A-iB-ih}^{a-A-iB+ih} \\ &= \int_{a-A-iB-ih}^{a-A-ih} + \int_{a-A-ih}^{a-A+ih} + \int_{a-A+ih}^{a-A-iB+ih} \\ &= \int_{a-A-iB-ih}^{a-A-ih} + \int_{\Re(a-\beta)-ih}^{\Re(a-\beta)+ih} + \int_{a-A+ih}^{a-A-iB+ih} \end{aligned}$$

so now we have to make sure that these side band integrals are negligible, let  $t = a - A + i\tau$

$$\begin{aligned} \int_{a-A+ih}^{a-A-iB+ih} \frac{x^t}{t} dt &= i \int_h^{(h-B)} \frac{e^{(a-A+i\tau) \log x}}{a-A+i\tau} d\tau \\ &\leq \int_h^{(h-B)} \frac{|e^{(a-A+i\tau) \log x}|}{|a-A+i\tau|} d\tau \\ &\leq e^{a-A} \int_h^{(h-B)} \frac{1}{|i\tau|} d\tau \\ &\leq e^{a-A} \log\left(1 - \frac{B}{h}\right) \end{aligned}$$

if we take the limit  $h \rightarrow \infty$  we get  $\rightarrow \log 1$  which approaches zero :) we can do the same with the lower side band and get the same result. The lesson here is that we don't need exact things, *i.e.* we don't need it to be exactly zero, approaching zero is good enough and this I see over and over in analytic number theory.

**Multiple Pages.** I find it interesting that to prove that the limit of the sum is the sum of the limit we just need to show that the latter converges but we did not need to show that they converge to the same value.

**Page 55.** The integrand in the second integral is an odd function

$$i \int_{-h}^h \frac{t}{a^2 + t^2} dt$$

and it's integrated in a symmetric domain so the integral is zero :)

**Page 55.** This is the normal trig trick :)

$$\begin{aligned}\int \frac{1}{1+u^2} du &= \int \frac{1}{1+\tan^2 w} \frac{dw}{\sec^2 w}, & \frac{1}{1+\tan^2 w} &= \sec^2 w \\ &= \int dw \\ &= w = \arctan u\end{aligned}$$

**Page 56.** Inequality on first line

$$\begin{aligned}(a-c)^2 &= a^2 - 2ac + c^2 > 0 \\ a^2 + c^2 &> 2ac \\ 2(a^2 + c^2) &> a^2 + 2ac + c^2 \\ a^2 + c^2 &> \frac{(a+c)^2}{2} \\ \sqrt{a^2 + c^2} &> \frac{(a+c)}{\sqrt{2}} \\ \rightarrow |a+ic| &> \frac{(a+c)}{\sqrt{2}}\end{aligned}$$

**Page 58.** And again, just for fun

$$\begin{aligned}\cos \frac{\pi}{4} &= \cos 2\frac{\pi}{8} = 2\cos^2 \frac{\pi}{8} - 1 \\ \frac{1}{\sqrt{2}} &= 2\cos^2 \frac{\pi}{8} - 1 \\ \rightarrow \cos \frac{\pi}{8} &= \left( \frac{1 + \frac{1}{\sqrt{2}}}{2} \right)^{1/2}\end{aligned}$$

by the same token

$$\begin{aligned}\cos \frac{\pi}{4} &= \cos 2\frac{\pi}{8} = 1 - 2\sin^2 \frac{\pi}{8} \\ \rightarrow \sin \frac{\pi}{8} &= \left( \frac{1 - \frac{1}{\sqrt{2}}}{2} \right)^{1/2}\end{aligned}$$

and so

$$\begin{aligned}\tan \frac{\pi}{8} &= \left( \frac{\frac{1 - \frac{1}{\sqrt{2}}}{2}}{\frac{1 + \frac{1}{\sqrt{2}}}{2}} \right)^{1/2} = \left( \frac{1 - \frac{1}{\sqrt{2}}}{1 + \frac{1}{\sqrt{2}}} \right)^{1/2} = \left( \frac{1 - \frac{1}{\sqrt{2}}}{1 + \frac{1}{\sqrt{2}}} \times \frac{1 + \frac{1}{\sqrt{2}}}{1 + \frac{1}{\sqrt{2}}} \right)^{1/2} \\ &= \left( \frac{1 - \frac{1}{2}}{\left(1 + \frac{1}{\sqrt{2}}\right)^2} \right)^{1/2} = \frac{1}{1 + \frac{1}{\sqrt{2}}} \cdot \frac{1}{\sqrt{2}} = \frac{1}{1 + \sqrt{2}} < \frac{1}{2}\end{aligned}$$

**Page 65.** Evaluating integrals

$$\int_0^K \frac{1}{r-1} dr$$

there's a pole here at  $r = 1$ , this is almost the same as integrating  $dz/z$ , so let's make the change of variable  $z = r - 1, dz = dr$

$$\int_0^K \frac{1}{r-1} dr = \int_{-1}^{K-1} \frac{1}{z} dz = \int_{-1}^{0-\delta} \frac{1}{z} dz + \int_C \frac{1}{z} dz + \int_{0+\delta}^{K-1} \frac{1}{z} dz$$

where the contour  $C$  for the middle integral is an upper semicircle above  $z = 0$ ,

$$\begin{aligned} \int_C \frac{1}{z} dz &= \int_{\pi}^0 \frac{1}{\delta e^{i\theta}} d(\delta e^{i\theta}) \\ &= i \int_{\pi}^0 \frac{\delta e^{i\theta}}{\delta e^{i\theta}} d\theta \\ &= -i\pi \end{aligned}$$

Now for the other integrals we do the following

$$\begin{aligned} \int_{-1}^{0-\delta} \frac{1}{z} dz + \int_{0+\delta}^{K-1} \frac{1}{z} dz &= \int_1^{0+\delta} \frac{1}{-z} d(-z) + \int_{0+\delta}^1 \frac{1}{z} dz + \int_1^{K-1} \frac{1}{z} dz \\ &= - \int_{0+\delta}^1 \frac{1}{z} dz + \int_{0+\delta}^1 \frac{1}{z} dz + \int_1^{K-1} \frac{1}{z} dz \\ &= \int_1^{K-1} \frac{1}{z} dz \end{aligned}$$

so we first do  $z \rightarrow -z$  on the first integral and then swap the integration limits, for the second integral we split the integral, this way we can cancel the blowing up parts and are left with  $\log(K-1)$ .

We can also shift the pole by the usual trick

$$\int_{-1}^{K-1} \frac{1}{z} dz \rightarrow \int_{-1}^{K-1} \frac{1}{z+i\epsilon} dz = \log(K-1+i\epsilon) - \log(-1+i\epsilon)$$

taking the limit  $\epsilon \rightarrow 0$  is a bit tricky because the log is a bit ambiguous, is  $\log(-1) = i\pi$  or  $-i\pi$ ?

**Page 69.** These equations are quite easily derived using von Mangoldt's method :) For me the starting point is Section 3.7, Eq. (1) - (3) which culminates in the following equation

$$\int_0^x t^{-r} d\psi(t) = \frac{x^{1-r}}{1-r} - \sum_{\rho} \frac{x^{\rho-r}}{\rho-r} + \sum_n \frac{x^{-2n-r}}{2n+r} - \frac{\zeta'(r)}{\zeta(r)}$$



and for what we need

$$\int_0^x \left(\frac{x}{t}\right)^r d\psi(t) = \frac{x}{1-r} - \sum_{\rho} \frac{x^{\rho}}{\rho-r} + \sum_n \frac{x^{-2n}}{2n+r} - x^r \frac{\zeta'(r)}{\zeta(r)}$$

If we set  $r = 0$  we get

$$\int_0^x t^{-0} d\psi(t) = \int_0^x t^{-0} d\psi(t) = \int_0^x d\psi(t) = \psi(x) = x - \sum_{\rho} \frac{x^{\rho}}{\rho} + \sum_n \frac{x^{-2n}}{2n} - \frac{\zeta'(0)}{\zeta(0)}$$

although this still needs all those arguments about swapping sums, integrals and limits :)

To get de la Vallee Poussin's Eq (3) we need the following IBP

$$\begin{aligned} \int_0^x \log \frac{x}{t} d\psi(t) &= \log \frac{x}{t} \psi(t) \Big|_0^x + \int_0^x t^{-1} \psi(t) dt \\ &= \psi(x) \log \frac{x}{x} - \psi(0) \log \frac{x}{0} + \int_0^x t^{-1} \psi(t) dt \end{aligned}$$

Note that  $\psi(0) = 0$ , there's no limit taken here so

$$\psi(0) \log \frac{x}{0} = \lim_{r \rightarrow 0} 0 \cdot \log \frac{x}{r}$$

and therefore it is zero and from the main equation of Section 3.7 as stated above we do the following (differentiating inside the integral)

$$\begin{aligned} \frac{d}{dr} \left( \int_0^x \left(\frac{x}{t}\right)^r d\psi(t) \right) \Big|_{r=0} &= \int_0^x \left( \log \frac{x}{t} \right) d\psi(t) \\ \int_0^x t^{-1} \psi(t) dt &= x - \sum_{\rho} \frac{x^{\rho}}{\rho^2} - \sum_n \frac{x^{-2n}}{(2n)^2} - \log x \frac{\zeta'(0)}{\zeta(0)} - \frac{\zeta''(0)\zeta(0) - \zeta'^2(0)}{\zeta^2(0)} \\ \rightarrow \int_0^x t^{-1} \psi(t) dt &= x - \sum_{\rho} \frac{x^{\rho}}{\rho} - \sum_n \frac{x^{-2n}}{(2n)^2} - \log x \frac{\zeta'(0)}{\zeta(0)} + \text{const.} \end{aligned}$$

This way we don't need the explicit formula for  $\psi(x)$ .

For Hadamard's Eq (2) what we need is the footnote on Page 74. First we reuse the (modified) main formula from Page 69, shown above, *i.e.* the formula for  $\int_0^x \left(\frac{x}{t}\right)^r d\psi(t)$ ,

once we have this, we do the following

$$\begin{aligned}
\int_0^x \left(\frac{x}{t}\right)^r d\psi(t) - \int_0^x \left(\frac{x}{t}\right)^w d\psi(t) &= \left(\frac{x}{1-r} - \frac{x}{1-w}\right) - \sum_{\rho} \left(\frac{x^{\rho}}{\rho-r} - \frac{x^{\rho}}{\rho-w}\right) + \\
&\quad \sum_n \left(\frac{x^{-2n}}{2n+r} - \frac{x^{-2n}}{2n+w}\right) - \left(x^r \frac{\zeta'(r)}{\zeta(r)} - x^w \frac{\zeta'(w)}{\zeta(w)}\right) \\
&= x \frac{(r-w)}{(r-1)(w-1)} - \sum_{\rho} \frac{x^{\rho}(r-w)}{(\rho-r)(\rho-w)} \\
&\quad - \sum_n \frac{x^{-2n}(r-w)}{(2n+r)(2n+w)} - \left(x^r \frac{\zeta'(r)}{\zeta(r)} - x^w \frac{\zeta'(w)}{\zeta(w)}\right)
\end{aligned}$$

which is the formula given in the footnote, and heeding the footnote :) to get (2) we set  $r = 1, w = 0$  and do IBP on the Left Hand Side

$$\begin{aligned}
\int_0^x \frac{x}{t} d\psi(t) - \int_0^x d\psi(t) &= \frac{x}{t} \psi(t) \Big|_0^x + \int_0^x \frac{x}{t^2} \psi(t) dt - \psi(x) \\
&= \left(\cancel{\psi(x)} - \frac{x\psi(x)}{0}\right) + x \int_0^x t^{-2} \psi(t) dt - \cancel{\psi(x)} \\
&= x \int_0^x t^{-2} \psi(t) dt
\end{aligned}$$

again  $\frac{\psi(0)}{0}$  is zero because  $\psi(0) = 0$ . The RHS is straightforward, except for the first term, therefore we have to combine this with the last term because these infinities cancel out

$$\begin{aligned}
&= \left(\lim_{r \rightarrow 1} x \frac{1}{1-r} - x^r \frac{\zeta'(r)}{\zeta(r)}\right) - \sum_{\rho} \frac{x^{\rho}(1-0)}{(\rho-1)(\rho-0)} - \sum_n \frac{x^{-2n}(1-0)}{(2n+1)(2n+0)} + x^0 \frac{\zeta'(0)}{\zeta(0)} \\
&= \left(\lim_{r \rightarrow 1} x \frac{1}{1-r} - x^r \frac{\zeta'(r)}{\zeta(r)}\right) - \sum_{\rho} \frac{x^{\rho}}{(\rho-1)\rho} - \sum_n \frac{x^{-2n}}{(2n+1)2n} + \frac{\zeta'(0)}{\zeta(0)}
\end{aligned}$$

the limit is calculated by Laurent expanding  $\frac{\zeta'(r)}{\zeta(r)}$  and isolate the pole term, however, we incur an extra minus sign (and it has to for it to work, see Lemma 4 Page 289 of Apostol)

$$\begin{aligned}
\lim_{r \rightarrow 1} \frac{x}{1-r} - x^r \frac{\zeta'(r)}{\zeta(r)} &= \lim_{r \rightarrow 1} \frac{x}{1-r} - \frac{x^r}{1-r} + x^r \cdot f(r) = \lim_{r \rightarrow 1} \frac{x - x^r}{1-r} + x^r \cdot f(r) \\
&= \lim_{r \rightarrow 1} \frac{-x^r \log x}{-1} + x^r \cdot \text{const.} \\
&= x \log x + x \cdot \text{const.}
\end{aligned}$$

Combining everything

$$\begin{aligned}
 x \int_0^x t^{-2} \psi(t) dt &= x \log x - \sum_{\rho} \frac{x^{\rho}}{(\rho-1)\rho} - \sum_n \frac{x^{-2n}}{(2n+1)2n} + \frac{\zeta'(0)}{\zeta(0)} + x \cdot \text{const.} \\
 \rightarrow \int_0^x t^{-2} \psi(t) dt &= \log x - \sum_{\rho} \frac{x^{\rho-1}}{(\rho-1)\rho} - \sum_n \frac{x^{-2n-1}}{(2n+1)2n} + \frac{1}{x} \frac{\zeta'(0)}{\zeta(0)} + \text{const.}
 \end{aligned}$$

**Page 70.** The sum

$$B = \frac{1}{2} \sum \frac{1}{p^2} + \frac{1}{3} \sum \frac{1}{p^3} + \frac{1}{4} \sum \frac{1}{p^4} + \dots$$

converges, this is how, for  $k \geq 2$

$$\sum \frac{1}{p^k} < \sum_{n=2} \frac{1}{n^k} < 1$$

the last inequality can be easily shown using the integral test which shows that

$$\sum_{n=2} \frac{1}{n^k} < \int_1^{\infty} x^{-k} dx = \frac{1}{k-1}$$

$$\begin{aligned}
 B &= \frac{1}{2} \sum \frac{1}{p^2} + \frac{1}{3} \sum \frac{1}{p^3} + \frac{1}{4} \sum \frac{1}{p^4} + \dots < \frac{1}{2} \sum \frac{1}{n^2} + \frac{1}{3} \sum \frac{1}{n^3} + \frac{1}{4} \sum \frac{1}{n^4} + \dots \\
 &< \frac{1}{2} \cdot 1 + \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{3} + \dots \\
 &= \sum_{n=2} \frac{1}{n(n-1)}
 \end{aligned}$$

and using the integral test this one is also convergent :)

Here's a way that doesn't work, initially I wanted to show that we can massage  $B$  into  $\sum \frac{1}{p^s}$  with  $s > 1$ , this is because each  $\sum \frac{1}{p^k}$  is less than one, so there must be an  $s > 1$  such that

$$\frac{1}{n} \sum \frac{1}{p^n} = \frac{1}{p^s}$$

however, this  $s$  is getting smaller and smaller as  $p$  gets bigger and it approaches 1, so if we use the same  $s$  for all  $p$  we get  $\sum \frac{1}{p}$  and it doesn't work.

**Page 73.** One trick to calculate  $\int_0^x t^{k-1} \psi(t) dt$  is, one, the partial fraction formula

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

and from this we can see that to get other powers of  $x$  we do the following

$$\frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{x^{s+1}}{n^s s(s+1)} ds \rightarrow \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{x^{s+k}}{n^s s(s+k)} ds$$

splitting through partial fraction

$$\frac{x^k}{k2\pi i} \int_{a-i\infty}^{a+i\infty} \left(\frac{x}{n}\right)^s \frac{ds}{s} - \frac{n^k}{k2\pi i} \int_{a-i\infty}^{a+i\infty} \left(\frac{x}{n}\right)^{s+k} \frac{ds}{s+k}$$

which becomes

$$= \begin{cases} \frac{1}{k} (x^k - n^k) & \text{if } n \leq x \\ 0 & \text{if } n \geq x \end{cases}$$

so what we have is

$$\frac{1}{k} \sum_{n \leq x} \Lambda(n) (x^k - n^k) = \frac{1}{k} \int_0^x (x^k - t^k) d\psi(t) = \int_0^x t^{k-1} \psi(t) dt$$

the last step uses IBP which nicely cancels the minus sign and the factor  $\frac{1}{k}$ . But of course this method doesn't work for  $k = 0$ .

**Page 74.** Another trick to calculate  $\int_0^x t^{-k} d\psi(t)$  is that the same thing that was done in Chapter 3. The main thing is

$$\frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \left[ -\frac{\zeta'(s)}{\zeta(s)} \right] x^s \frac{ds}{s} = \sum_n \Lambda(n) \left\{ \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \left(\frac{x}{n}\right)^s \frac{ds}{s} \right\}$$

and whatever is inside the curly braces is 1 or 0, which translates to the integral  $\int_0^x d\psi(t)$ . Now what we want is  $\int_0^x \left(\frac{x}{t}\right)^u d\psi(t)$  which is equivalent to  $\sum_{n \leq x} \Lambda(n) \left(\frac{x}{n}\right)^u$ . So instead of 1 or 0, you want the whole curly braces to be  $\left(\frac{x}{n}\right)^u$  or 0.

This is definitely easier than the one depicted on Page 73 with partial fraction, we only need to multiply the curly braces with  $\left(\frac{x}{n}\right)^u$

$$\begin{aligned} \sum_n \Lambda(n) \left\{ \left(\frac{x}{n}\right)^u \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \left(\frac{x}{n}\right)^s \frac{ds}{s} \right\} &= \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \left[ \sum_n \Lambda(n) n^{-s-u} \right] x^{s+u} \frac{ds}{s} \\ &= \frac{x^u}{2\pi i} \int_{a-i\infty}^{a+i\infty} \left[ -\frac{\zeta'(s+u)}{\zeta(s+u)} \right] x^s \frac{ds}{s} \end{aligned}$$

but this is exactly what was done in Section 3.7. Note that we can't pull the factor  $\left(\frac{x}{n}\right)^u$  out of the sum because it involves  $n$  which is summed. So the derivation of this footnote

will just utilize the method of Section 3.7, for the full derivation see comments on Page 69.

**Page 82.** How do we compare  $x^{-\epsilon}$  and  $\exp[-c(\log x)^{1/2}]$ , the book claims that  $x^{-\epsilon}$  decays much faster, to see this set  $y = \log x$  thus

$$x^{-\epsilon} \rightarrow e^{-\epsilon y} \quad \text{and} \quad \exp[-c(\log x)^{1/2}] \rightarrow e^{-c\sqrt{y}}$$

so for a sufficiently large  $y$ ,  $e^{-\epsilon y}$  decays faster. Now how about the absorption into a constant  $c$ ? Since

$$e^{-\epsilon y} = O(e^{-c\sqrt{y}})$$

This means that there is a constant  $K$  such that for  $y > y_0$

$$|e^{-\epsilon y}| < K|e^{-c\sqrt{y}}|$$

and the usual trick

$$\begin{aligned} Ke^{-c\sqrt{y}} &= e^{\log K} e^{-c\sqrt{y}} \\ &= e^{-c\sqrt{y} + \log K} = e^{-c\sqrt{y}(1 - \log K/c\sqrt{y})} \end{aligned}$$

Now

$$e^{-c\sqrt{y}(1 - \log K/c\sqrt{y})} > e^{-c\sqrt{y}}$$

but  $\log K/c\sqrt{y}$  gets smaller and smaller as  $y$  goes up, so we can get the max of this value, say  $K'$  (we can set  $K' < 1$  for sufficiently large  $y$ ) and the whole thing becomes

$$e^{-c\sqrt{y}(1-K')} > e^{-c\sqrt{y}(1 - \log K/c\sqrt{y})} > e^{-c\sqrt{y} + \log K} > e^{-c\sqrt{y}}$$

thus we can choose a new  $c' = c(1 - K')$  such that

$$e^{-c'\sqrt{y}} > e^{-c\sqrt{y}}$$

**Page 83.**

$$\frac{\psi(x) - x}{x} \leq K_1 e^{-c_1(\log x)^{1/2}} \leq e^{-c'_1(\log x)^{1/2}}$$

and

$$\frac{\psi(y) - y}{y} \geq -K_2 e^{-c_2(\log y)^{1/2}} \geq -e^{-c'_2(\log y)^{1/2}}$$

combining

$$-e^{-c'_2(\log x)^{1/2}} \leq \frac{\psi(x) - x}{x} \leq e^{-c'_1(\log x)^{1/2}}$$

now we can choose  $c'' = \min\{c'_1, c'_2\}$  such that we get a bigger magnitude for the error

$$|e^{-c''(\log x)^{1/2}}| \geq \max\left\{|e^{-c'_2(\log x)^{1/2}}|, |e^{-c'_1(\log x)^{1/2}}|\right\}$$

Thus the magnitude of the relative error is

$$\left| \frac{\psi(x) - x}{x} \right| \leq e^{-c''(\log x)^{1/2}}$$

**Page 83.** The term relative error here means  $\frac{\theta(x) - x}{x}$ , so it is divided by  $x$ , thus it should be that

$$\theta(x) - x \sim x \cdot e^{-c(\log x)^{1/2}}$$

to show this, we know from its definition

$$\psi(x) = \theta(x) + \theta(x^{1/2}) + \theta(x^{1/3}) + \dots$$

so the difference between

$$\psi(x) - \theta(x) \leq \theta(x^{1/2}) \frac{\log x}{\log 2}$$

but the difference between

$$\psi(x) - x \leq x e^{-c(\log x)^{1/2}}$$

but, as stated,  $\theta(x^{1/2}) \log x \sim x^{1/2} \log x$  and this is much smaller than  $x e^{-c(\log x)^{1/2}}$  (because  $\frac{x^{1/2} \log x}{x}$  already goes to zero as  $x \rightarrow \infty$ , thus  $\theta(x^{1/2}) \log x = O(x e^{-c(\log x)^{1/2}})$ ), so the difference between  $\theta(x)$  and  $x$  is potentially

$$\begin{aligned} \left| \frac{\theta(x) - x}{x} \right| &\leq \frac{\theta(x^{1/2}) \log x}{x \log 2} + e^{-c(\log x)^{1/2}} \sim \frac{\log x}{x^{1/2} \log 2} + e^{-c(\log x)^{1/2}} \\ &\leq O(e^{-c(\log x)^{1/2}}) + e^{-c(\log x)^{1/2}} \\ &\leq e^{-c'(\log x)^{1/2}} \end{aligned}$$

However, the book wants to use the same constant  $c$  for both  $\theta(x)$  and  $\psi(x)$  (note that this  $c$  would be the  $c_0$  in the book). We can do this if we ignore  $\frac{\theta(x^{1/2}) \log x}{x \log 2}$  which means that my estimate is too conservative. I shouldn't have set  $\frac{\theta(x^{1/2}) \log x}{x \log 2} = O(e^{-c(\log x)^{1/2}})$  since  $\frac{\theta(x^{1/2}) \log x}{e^{-c(\log x)^{1/2}}} \rightarrow 0$ , this is equivalent to the usual error estimate,  $e^{-c(\log x)^{1/2}}$  is  $\epsilon$  while  $\frac{\theta(x^{1/2}) \log x}{x \log 2}$  is  $\epsilon^2$  and here we are only interested in first order error estimate only.

**Page 83 - Page 84.** Ignoring the  $\pi(x)$

$$\begin{aligned}
&= \frac{y[1 + \varepsilon(y)]}{\log y} - \frac{x[1 - \varepsilon(x)]}{\log x} + \int_x^y \frac{t[1 + \varepsilon(t)]}{(\log t)^2 t} dt \\
&= \frac{y\varepsilon(y)}{\log y} + \frac{x\varepsilon(x)}{\log x} + \left( \frac{y}{\log y} - \frac{x}{\log x} + \int_x^y \frac{1}{(\log t)^2} dt \right) + \int_x^y \frac{\varepsilon(t)}{(\log t)^2} dt \\
&= \frac{y\varepsilon(y)}{\log y} + \frac{x\varepsilon(x)}{\log x} + \int_x^y \frac{dt}{\log t} + \int_x^y \frac{\varepsilon(t)}{(\log t)^2} dt
\end{aligned}$$

In the first line we do IBP on the last integral, and going to the last line we do inverse IBP on the three terms in the brackets.

Going to Page 84 we take  $x$  as constant, thus anything containing  $x$  will be taken as const.

$$\begin{aligned}
&= \frac{y\varepsilon(y)}{\log y} + \text{const.} + \int_0^y \frac{dt}{\log t} - \int_0^x \frac{dt}{\log t} + \int_x^y \frac{\varepsilon(t)}{(\log t)^2} dt \\
&= \frac{y\varepsilon(y)}{\log y} + \text{const.} + \text{Li}(y) - \text{Li}(x) + \int_x^y \frac{\varepsilon(t)}{(\log t)^2} dt \\
&= \frac{y\varepsilon(y)}{\log y} + \text{const.} + \text{Li}(y) + \text{const.} + \int_x^y \frac{\varepsilon(t)}{(\log t)^2} dt
\end{aligned}$$

Note that for  $a > b$  we have  $\varepsilon(a) < \varepsilon(b)$  so it is now straightforward to get the upper bound.

The lower bound is a slightly trickier story :) usually if we have an integral say

$$\int_x^{y^{1/2}} \frac{\varepsilon(t)}{(\log t)^2} dt$$

to get the lower bound we minimize the integrand

$$\int_x^{y^{1/2}} \frac{\varepsilon(y^{1/2})}{(\log y^{1/2})^2} dt$$

but we can also do the following, *maximize* the integrand and then multiply by  $-1$  and that is what we do here, we also multiply the  $\frac{y\varepsilon(y)}{\log y}$  term with  $-1$

$$\begin{aligned} &= \frac{y\varepsilon(y)}{\log y} + \text{const.} + \text{Li}(y) + \int_x^{y^{1/2}} \frac{\varepsilon(t)}{(\log t)^2} dt + \int_{y^{1/2}}^y \frac{\varepsilon(t)}{(\log t)^2} dt \\ &\geq -\frac{y\varepsilon(y)}{\log y} + \text{const.} + \text{Li}(y) - \int_x^{y^{1/2}} \frac{\varepsilon(x)}{(\log x)^2} dt - \int_{y^{1/2}}^y \frac{\varepsilon(y^{1/2})}{(\log y^{1/2})^2} dt \end{aligned}$$

and we get something very similar to the upper bound except for a minus sign for the square bracket.

**Extra Note:** as have been done over and over again in the book, getting upper bounds are about getting rid of negative terms and getting lower bounds is about getting rid of positive terms, and from above we also see that we can get the upper bound and stick a negative sign to get a lower bound and I believe vice versa as well.

**Page .** Making this note just to clear my head :) Relative error between  $\text{Li}(x)$  and  $\frac{x}{\log x - A}$  usually means that

$$\frac{x/(\log x - A) - \text{Li}(x)}{\text{Li}(x)}$$

but in this case we just want the difference, *i.e.* the absolute error, not the ratio, *i.e.*  $x/(\log x - A) - \text{Li}(x)$ , while

$$\text{Li}(x) \sim \frac{x}{\log x} + \frac{x}{(\log x)^2} \quad \text{and} \quad \frac{x}{\log x - A} \sim \frac{x}{\log x} + \frac{Ax}{(\log x)^2}$$

and so

$$\begin{aligned} x/(\log x - A) - \text{Li}(x) &= \frac{x}{\log x} + \frac{Ax}{(\log x)^2} - \left( \frac{x}{\log x} + \frac{x}{(\log x)^2} \right) \\ &= (A - 1) \frac{x}{(\log x)^2} \end{aligned}$$

So if we divide this relative error by  $x/(\log x)^2$ , we get  $A - 1$ .

**Page .** const times  $x^{-1/2+\epsilon}$  is the same as just  $x^{-1/2+\epsilon}$  for sufficiently larger  $x$  :) because

$$Kx^{-1/2+\epsilon} = x^{-1/2+\epsilon+\delta}$$

where  $x^\delta = K$  or  $\delta = \log_x K$ , now say that the initial upper bound  $\leq Kx^{-1/2+\epsilon}$  is valid starting from  $x \geq x_0$ , we now need to push this starting point further up because we can make  $\delta$  as small as possible by increasing  $x$ . Although since  $\delta > 0$ ,  $\epsilon + \delta > \epsilon$



Mistake on Li

Mistake on sum over  $\gamma$ , missing factor of 2 for the first one

**Page .** Another head clearing effort :) Note that  $\mu(1) = 1$  and not zero, this was the mistake I made. This relationship can be shown directly without the integral trick, although that trick is quite sick :)

First, massage the sums into

$$\begin{aligned} \sum_{n < x} \frac{\mu(n)}{n} P\left(\frac{x}{n}\right) &= \sum_{n < x} \frac{\mu(n)}{n} \sum_{b < x/n} \frac{\Lambda(b)}{b} \\ &= \sum_{n < x} \sum_{b < x/n} \frac{\mu(n)}{n} \frac{\Lambda(b)}{b} \\ &= \sum_{n < x} \sum_{ab=n} \frac{\mu(a)}{a} \frac{\Lambda(b)}{b} \\ &= \sum_{n < x} \frac{1}{n} \sum_{ab=n} \mu(a) \Lambda(b) \end{aligned}$$

this is quite tricky, it's easier to see the other way round, going from  $\sum_{n < x} \sum_{ab=n}$  to  $\sum_{n < x} \sum_{b < x/n}$ , because for  $\sum_{n < x} \sum_{ab=n}$ , for each fixed value of  $a$  we have  $b$  going from 1 to  $x/a$ . For example, set  $x = 13$  and fix  $a = 2$  then for each of  $n = 2, 4, 6, 8, 10, 12$  we get a term from *different*  $\sum_{ab=n}$  sums with  $b = 1, 2, 3, 4, 5, 6$  while  $a = 2$ .

The reason the above is quite tricky is because we can't do algebraic manipulations on indices of the sums to get it since as mentioned above it involves rearrangement of the summands, we take summands from *different* sums and combine them.

We just need to show that the inner sum  $\sum_{ab=n} \mu(a) \Lambda(b)$  is just  $-\mu(n) \log n$  and this is quite straightforward,  $\Lambda(b)$  is zero unless  $b$  is a prime power while  $\mu(a)$  is zero unless  $a$  is square free, this means that the only non zero  $\mu(a) \Lambda(b)$  are the ones with  $ab = p_1 p_2 \cdots p_j p_k^m$ , *i.e.* only one prime factor can have powers, the others must be of power 1 the other non-zero case is  $ab = p_1 p_2 \cdots p_j p_k$ , no prime powers at all.

For the first case

$$\begin{aligned} \sum_{ab=n} \mu(a) \Lambda(b) &= \mu(n/p_k) \Lambda(p_k^{m-1}) + \mu(n/p_k^m) \Lambda(p_k^m) \\ &= (-1)^{k-1} \Lambda(p_k) + (-1)^k \Lambda(p_k) \\ &= 0 \end{aligned}$$

there are only two terms surviving from the sum because for any other case either  $\mu$  is zero or  $\Lambda$  is zero.

For the second case

$$\begin{aligned}
 \sum_{ab=n} \mu(a)\Lambda(b) &= \sum_{i=1}^k \mu(n/p_i)\Lambda(p_i) \\
 &= \sum_{i=1}^k (-1)^{k-1} \log p_i \\
 &= (-1)^{k-1} \sum_{i=1}^k \log p_i \\
 &= (-1)\mu(n) \log n
 \end{aligned}$$

since  $\mu(n) = (-1)^k$  and we are done.

## **Chapter 6**

Main outline

1. Euler-Maclaurin formula is applied to the Gamma function,  $\Pi(s)$  and the zeta function,  $\zeta(s)$ , not the Riemann symmetric functional  $\xi(s)$ . The Euler-Maclaurin expansions of the two were then used to calculate the Riemann symmetric functional  $\xi(s)$ .
2. Using the expansion above we then calculate the location of the roots of  $\xi(s)$  on the line by looking for sign change along the line.
3. We then prove Riemann hypothesis by calculating the number of roots in the strip (not just on the line), if the number of roots in the strip is the same as the one on the line then Riemann hypothesis is true.

List of tricks for Euler-Maclaurin

1. Trapezoidal rule.
2. Stieljes integral.
3. Bernoulli's polynomial

The first two can be combined as they are nothing but the Euler's summation formula as we shall see, of course we can just use Stieljes integral to “convert” a sum into an integral but Stieljes integral is not really an integral it's more of an identity for a sum if you think about it therefore we need the trapezoidal rule to approximate the integral. The goal here is not just to transform a sum into an integral, we want to get the error estimate which is very important in this case. The funny thing here is that the starting point is an approximation to an approximation

$$\begin{aligned}\sum_M^N f(n) &= \text{trapezoidal rule} + \frac{1}{2}[f(M) + f(N)] \\ &\sim \int_M^N f(x)dx + \frac{1}{2}[f(M) + f(N)]\end{aligned}$$

The final integral formula itself can be obtained from Euler's summation formula (see Apostol for details of derivation of Euler's summation formula)

$$\sum_{y < n \leq x} f(n) = \int_y^x f(t)dt + \int_y^x (t - [t])f'(t)dt + f(x)([x] - x) - f(y)([y] - y)$$

we focus on the integral with  $(t - [t])$

$$\begin{aligned}\int_y^x (t - [t])f'(t)dt &= - \int_y^x ([t] - t)f'(t)dt \\ &= - \int_y^x ([t] - t + \frac{1}{2} - \frac{1}{2})f'(t)dt \\ &= - \int_y^x ([t] - t + \frac{1}{2})f'(t)dt + \int_y^x \frac{1}{2}f'(t)dt \\ &= - \int_y^x ([t] - t + \frac{1}{2})f'(t)dt + \int_y^x \frac{1}{2}f'(t)dt \\ &= - \int_y^x ([t] - t + \frac{1}{2})f'(t)dt + \frac{1}{2}(f(x) - f(y))\end{aligned}$$

we now set  $y = M$  and  $x = N$  to get (and a change of dummy variable  $t \rightarrow x$  to get the same notation)

$$\begin{aligned}\sum_{M < n \leq N} f(n) &= \int_M^N f(x)dx - \int_M^N ([x] - x + \frac{1}{2})f'(x)dx + \\ &\quad f(M)([M] - M) - f(N)([N] - N) + \frac{1}{2}(f(N) - f(M)) \\ &= \int_M^N f(x)dx - \int_M^N ([t] - t + \frac{1}{2})f'(x)dx + 0 - 0 + \frac{1}{2}(f(N) - f(M))\end{aligned}$$

now we need to add  $f(M)$  to the LHS since the sum starts with an inequality and therefore  $f(M)$  is excluded while the one in the book includes  $f(M)$ , but if we are adding to the LHS we need to add the RHS as well

$$\begin{aligned}\sum_{M \leq n \leq N} f(n) &= \int_M^N f(x)dx - \int_M^N ([x] - x + \frac{1}{2})f'(x)dx + \frac{1}{2}(f(N) - f(M)) + f(M) \\ &= \int_M^N f(x)dx - \int_M^N ([x] - x + \frac{1}{2})f'(x)dx + \frac{1}{2}(f(N) + f(M))\end{aligned}$$

and we get the same formula, although I must say that the derivation here feels more natural starting with the trapezoidal rule (which is a common theme throughout the book, most the derivation was done in a more intuitive way rather than mechanically as is the case with Apostol).

The error estimate is the reason we need Bernoulli's polynomial and the reason we can use Bernoulli's polynomial is because

$$x - [x] - \frac{1}{2} = t - \frac{1}{2} = B_1$$

which to me is a coincidence :)

List of tricks for Bernoulli's polynomial

1.  $\int_x^{x+1} f(x) = \int_x^{x+1/2} f(x) + f(x+1/2)$  which is kinda obvious but not really.
2. The thing with  $p(h)$ , this is a sick trick :)

$$\begin{aligned}\sum_{n=0}^{\infty} \frac{B_n}{n!} (-h)^n &= \frac{h}{1 - e^{-h}} = \sum_{n=0}^k \frac{B_n(0)}{n!} - \frac{h^{k+1}}{k!} \int_0^{\infty} \overline{B}_k(x) e^{-hx} dx \\ \sum_{n=0}^k \frac{B_n}{n!} (-h)^n + \sum_{n=k+1}^{\infty} \frac{B_n}{n!} (-h)^n &= \sum_{n=0}^k \frac{B_n(0)}{n!} + P(h)\end{aligned}$$

where  $P(h) = -\frac{h^{k+1}}{k!} \int_0^{\infty} \overline{B}_k(x) e^{-hx} dx$ . Note that since  $\frac{h}{1-e^{-h}}$  is finite, the rest of the sum  $\sum_{n=k+1}^{\infty} \frac{B_n}{n!} (-h)^n = C(h)$  is also finite, thus

$$\begin{aligned}\sum_{n=0}^k \frac{B_n}{n!} (-h)^n + C(h) &= \sum_{n=0}^k \frac{B_n(0)}{n!} + P(h) \\ \sum_{n=0}^k \frac{1}{n!} [(-1)^n B_n - B_n(0)] h^n &= P(h) - C(h) \\ &= p(h)\end{aligned}$$

We know that  $P(h) \leq \text{const} \cdot h^{k+1}$  as stated in the book and  $C(h)$  is finite and also a power series of  $h$  starting with  $h^{k+1}$ . Therefore

$$\frac{C(h)}{h^{k+1}} = \sum_{n=0}^{\infty} \frac{B_{n+k+1}}{(n+k+1)!} (-h)^n$$

and

$$\lim_{h \rightarrow 0} \frac{p(h)}{h^{k+1}} = \lim_{h \rightarrow 0} \frac{1}{h^{k+1}} (P(h) - C(h)) = \text{const} - \frac{B_{n+k+1}}{(n+k+1)!} < \infty$$

which is finite but  $\lim_{h \rightarrow 0} \frac{p(h)}{h^{k+1}}$  is something divided by 0, for it to be bounded the numerator must be zero as well, thus  $p(h) = 0$  (recall that  $\frac{0}{0}$  can be anything and in our case it is given by  $\text{const} - \frac{B_{n+k+1}}{(n+k+1)!}$ ), from the above it can be deduced that

$$\text{const} \cdot h^{k+1} = -\frac{h^{k+1}}{k!} \int_0^{\infty} \bar{B}_k(x) e^{-hx} dx = \sum_{n=k+1}^{\infty} \frac{B_n}{n!} (-h)^n$$

and setting  $h = 1$  we get

$$\frac{1}{k!} \int_0^{\infty} \bar{B}_k(x) e^{-x} dx = \sum_{n=k+1}^{\infty} \frac{B_n}{n!} (-1)^{n+1}$$

which might be a useful identity :) Well for one thing if we set  $k = 0$

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{B_n}{n!} (-1)^{n+1} &= \int_0^{\infty} \bar{B}_0(x) e^{-x} dx = \int_0^{\infty} e^{-x} dx \\ \sum_{n=1}^{\infty} \frac{B_n}{n!} (-1)^{n+1} &= 1 \\ \rightarrow \sum_{n=0}^{\infty} \frac{B_n}{n!} (-1)^{n+1} &= 0 \end{aligned}$$

If I didn't make any mistake, this is the second Bernoulli's identity I discovered :)

A quick remark on the estimation of error as the last term omitted, note that this is always true, if you omit a term, the error you make is the term that you omit, the thing is since the Bernoulli's numbers at some point will grow out of bounds, it's just that if you use too many terms your error grows so you must know when to stop. Also, doesn't

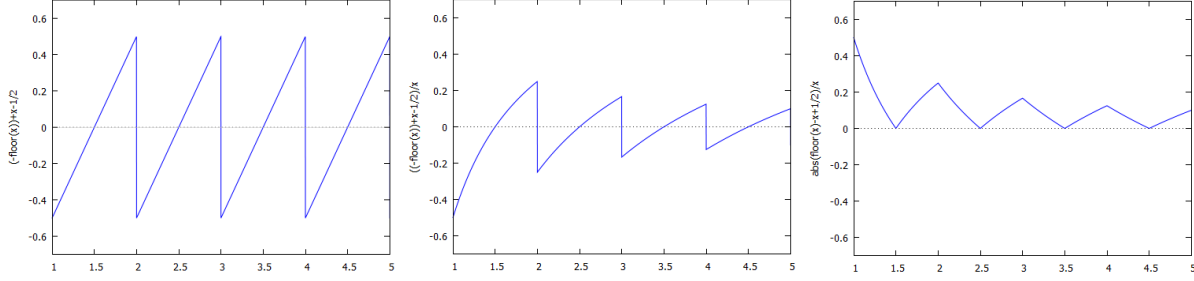


FIG. 1: Left is the plot of  $\overline{B}_n(x) = x - [x] - \frac{1}{2}$ . Middle is  $\frac{\overline{B}_n(x)}{x}$ . Right is  $\left| \frac{\overline{B}_n(x)}{x} \right|$

matter which term you stop at, we can still make the same error estimate by noting that the  $B_n$  is oscillatory and if the function is monotonically decreasing we can still make the same error estimate as the  $\int_{10}^{10.5}$ . Also since this is an asymptotic series (not a Taylor series), the series might not converge, note that the Bernoulli's numbers oscillate between positive and negative and their magnitude increases without bound.

**Page 106.** On the application of Euler-Maclaurin to  $\Pi(s)$

1. Calculate  $\log \Pi(s)$  instead of  $\Pi(s)$ . Not sure why we want to do this, one thing I can see is that  $\log \Pi(s)$  will have smaller errors, although if we exponentiate it we will get the same thing?
2. Do the following integral tricks,

$$\int_1^N = \int_1^\infty - \int_N^\infty \quad \int_N^\infty \frac{1}{x} = \int_0^\infty \frac{1}{x+N}$$

3. True to the intuitive spirit of the book, we first apply Euler-Maclaurin to  $\Pi(s)$  where  $s$  is an integer rather than going straight to the general complex  $s$ , which can't be done using Euler-Maclaurin since this formula is to transform discrete series into an integral while  $\Pi(s)$  with complex  $s$  is not a discrete series.

**Page 106.** For the first trick we have to do alternating series test to show that  $\int_1^\infty$  converges. About the alternating series test,  $\overline{B}_1(x)$  is the periodic function, *i.e.*  $\overline{B}_1(x) = x - [x] - \frac{1}{2}$ , see Fig. 1 Left. To do the alternating series test we don't actually need to do the integral. From Fig. 1 Middle, we see that in each integer interval  $M \leq x \leq M+1$ ,  $\int_M^{M+\frac{1}{2}} \frac{\overline{B}_1(x)}{x}$  is negative and  $\int_{M+\frac{1}{2}}^M \frac{\overline{B}_1(x)}{x}$  is positive. So that's your alternating series.

The alternating series test requires two conditions to be met, one the magnitude of the terms decreases and as  $n \rightarrow \infty$ ,  $a_n \rightarrow 0$ . To see that  $|a_n|$  decreases see Fig. 1 Right, this is  $\left| \int \frac{\overline{B}_1(x)}{x} \right|$  and the second condition is obvious.

**Page 108.** Mid page, in doing the integral

$$\int_0^1 \frac{t - \frac{1}{2}}{t + s} dt$$

instead of doing IBP directly or a change of variables, it did a simple yet smart trick which actually amounts to a change of variables

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

**Page 111.** In doing IBP we actually have a lot of freedom because

$$\frac{dF(x)}{dx} = f(x) \quad \longrightarrow \quad \int f(x) dx = F(x) + C$$

we have massive freedom in choosing  $C$ , this is true even when doing IBP. This is why the book says that “The function  $B_{2\nu} - \overline{B}_{2\nu}(x)$  is an antiderivative of  $-2\nu \overline{B}_{2\nu-1}(x)$ ”, it is “an” antiderivative and not “the” antiderivative :) Choosing this antiderivative simplifies a lot of things, like

**Page 112** Argument about the sign of  $B_{2\nu} - \overline{B}_{2\nu}(x)$ , this is very clever indeed, we know that the odd  $\overline{B}_{2\nu+1}(x)$  has exactly three zeros at  $x = 0, \frac{1}{2}, 1$ , therefore it has two maxima and its derivative  $\overline{B}'_{2\nu+1} = (2\nu + 1)\overline{B}_{2\nu}$  has 2 zeros at those maxima.

We also know that the  $\overline{B}_{2\nu}$  has one maxima in  $[0, 1]$  because its derivative  $(2\nu)\overline{B}_{2\nu-1}$  has a zero at  $x = \frac{1}{2}$ , so  $\overline{B}_{2\nu}$  has a maxima at  $x = \frac{1}{2}$ .

So how do you reconcile the fact that  $\overline{B}_{2\nu}$  has two zeros (not on  $x = 0$  or  $x = 1$ ) but only one maxima? This can be done if  $\overline{B}_{2\nu}$  starts and ends off positive but in the middle it dips down to negative (or vice versa), see Fig. 2 for an example of  $\overline{B}_2$ . But in this case  $\overline{B}_{2\nu}$  flips sign between positive and negative, we don’t want this, we want the sign to stay constant throughout  $[0, 1]$  to simplify things. Note that we can also do  $\overline{B}_{2\nu}(x) - \overline{B}_{2\nu}$  to accomplish this.

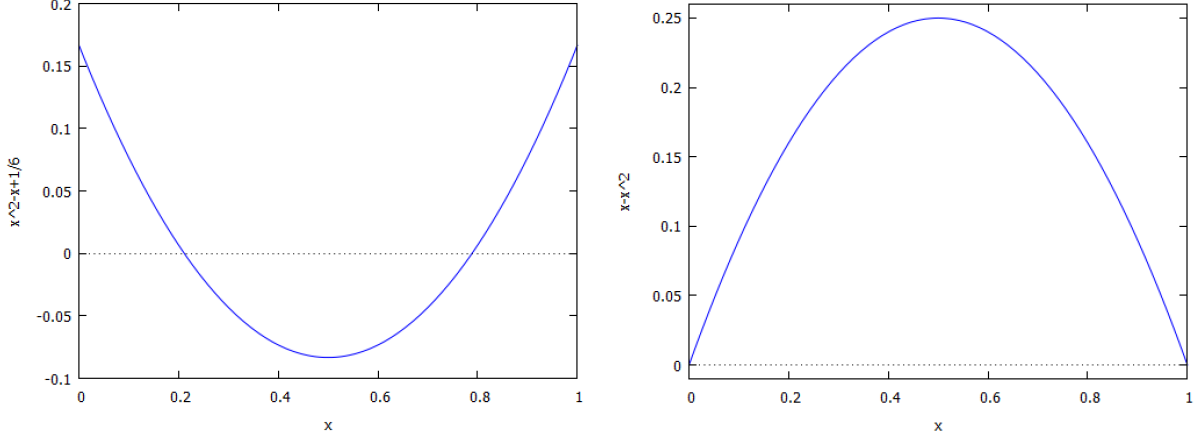


FIG. 2: Left is the plot of  $\bar{B}_2(x)$ . Right is  $B_2 - \bar{B}_2(x)$

**Page 113.** Last inequality

$$\operatorname{Re} \frac{\Pi'(\sigma + it)}{\Pi(\sigma + it)} \leq 2 \log t$$

In the derivation we start with the derivative of  $R_0$ , to do this we need to start from the very beginning at the bottom of Page 111

$$R_{2\nu-2} \leq \int_0^\infty \frac{[B_{2\nu} - B_{2\nu}(x)] dx}{2\nu(s+x)^{2\nu}}$$

we have to do the  $\frac{d}{ds}$  derivative here before we take the absolute value, otherwise we incur a minus sign and we need to do the complicated  $\frac{d|s|}{ds}$  later. So

$$\frac{d}{ds} R_{2\nu-2} \leq - \int_0^\infty \frac{[B_{2\nu} - B_{2\nu}(x)] dx}{(s+x)^{2\nu+1}}$$

now we can take the absolute value and get rid of the minus sign :) the next step is to estimate

$$\frac{1}{|s+x|} \leq \frac{1}{(|s|+x) \cos(\theta/2)}$$

as shown on Page 112, combining the two we get something like the middle of Page 112 (and also forgetting about the second integral)

$$\begin{aligned} |R'_{2\nu-2}| &\leq \int_0^\infty \frac{(-1)^{\nu+1} B_{2\nu} dx}{\cos^{2\nu+1}(\theta/2) (|s|+x)^{2\nu+1}} \\ &= - \frac{|B_{2\nu}|}{(2\nu) \cos^{2\nu+1}(\theta/2) (|s|+x)^{2\nu}} \Big|_0^\infty \\ &= \frac{|B_{2\nu}|}{(2\nu) \cos^{2\nu+1}(\theta/2) |s|^{2\nu}} \end{aligned}$$



Note that the minus sign in the second line is crucial as it negates the minus sign from computing the lower limit of the integral :) if we set  $\nu = 1$  we get

$$|R'_0| \leq \frac{B_2}{2 \cos^3(\theta/2) |s|^2}$$

which is what we have on Page 114. The second line on Page 114 is

$$\left| \frac{\Pi'(s)}{\Pi(s)} - \log s \right| \leq \frac{1}{2|s|} + \frac{B_2}{2 \cos^3(\theta/2) |s|^2}$$

for sufficiently large  $s$  and this is the key, sufficiently large because

$$\frac{1}{2|s|} \geq \frac{B_2}{2 \cos^3(\theta/2) |s|^2}$$

thus

$$\begin{aligned} \left| \frac{\Pi'(s)}{\Pi(s)} - \log s \right| &\leq \frac{1}{2|s|} + \frac{1}{2|s|} \\ &\leq \frac{1}{|s|} \end{aligned}$$

this is the second line, to get to the third line, note that

$$\left| \frac{\Pi'(s)}{\Pi(s)} - \log s \right| = \left| \operatorname{Re} \left( \frac{\Pi'(s)}{\Pi(s)} - \log s \right) + i \cdot \operatorname{Im} \left( \frac{\Pi'(s)}{\Pi(s)} - \log s \right) \right|$$

thus because we are ignoring the imaginary part

$$\begin{aligned} \left| \operatorname{Re} \left( \frac{\Pi'(s)}{\Pi(s)} - \log s \right) \right| &= \left| \operatorname{Re} \frac{\Pi'(s)}{\Pi(s)} - \operatorname{Re} \log s \right| \leq \left| \frac{\Pi'(s)}{\Pi(s)} - \log s \right| \\ &\rightarrow \left| \operatorname{Re} \frac{\Pi'(s)}{\Pi(s)} - \operatorname{Re} \log s \right| \leq \frac{1}{|s|} \end{aligned}$$

which is the third line, also since  $1 \leq \sigma \leq 2$ ,  $\operatorname{Re} \log s$  is always positive because  $s$  is on the right half of the complex plane thus  $\operatorname{Re} \log s \leq \log |s|$  and

$$\begin{aligned} \operatorname{Re} \frac{\Pi'(s)}{\Pi(s)} - \log s &\leq \left| \operatorname{Re} \frac{\Pi'(s)}{\Pi(s)} - \operatorname{Re} \log s \right| \\ \rightarrow \operatorname{Re} \frac{\Pi'(s)}{\Pi(s)} - \log |s| &\leq \frac{1}{|s|} \\ \rightarrow \operatorname{Re} \frac{\Pi'(s)}{\Pi(s)} &\leq \log |s| + \frac{1}{|s|} \end{aligned}$$

setting  $s = (\sigma + it)/2$

$$\begin{aligned} \operatorname{Re} \frac{\Pi'((\sigma + it)/2)}{\Pi((\sigma + it)/2)} &\leq \log \left| \frac{\sigma + it}{2} \right| + \frac{1}{|s|} \\ &\leq \log \left| 1 + i \frac{t}{2} \right| + \frac{1}{|s|} \end{aligned}$$

now since  $1 \leq \sigma \leq 2$  we set  $\sigma = 2$ , for the very last step we want to show that  $\log t \geq |1 + i\frac{t}{2}|$  and  $\log t \geq \frac{1}{|s|}$ , the latter is obvious for sufficient large  $s$ , the former though

$$\begin{aligned} \left|1 + i\frac{t}{2}\right| &= \left(1 + \frac{t^2}{4}\right)^{1/2} \\ &= t \left(\frac{1}{t^2} + \frac{1}{4}\right)^{1/2} \\ &\leq t \end{aligned}$$

since whatever is in the bracket is less than 1 (again for sufficiently large  $s$ ) and so we get the final result :)

**Page 114.** Application of Euler-Maclaurin to  $\zeta(s)$ . This is quite straightforward except that we cannot start the sum  $\sum n^{-s}$  from  $n = 1$  because then the series becomes

$$\sum_{n=1}^{\infty} n^{-s} = \int_1^{\infty} x^{-s} dx + \frac{1}{2} 1^{-s} + \frac{B_2}{2} s + \cdots + \frac{B_{2\nu}}{(2\nu)!} s(s+1) \cdots (s+2\nu-2) + R_{2\nu}$$

so the series decreases only by virtue of the decrease the Bernoulli's numbers.

**Page 117.** About the calculation of  $n^{-\frac{1}{2}-it}$  as  $n^{-1/2}[\cos(t \log n) + i \sin(t \log n)]$  as

$$18 \log 2 = 4\pi - \theta$$

$$18 \log 3 = 6\pi + (\pi/3) - \theta$$

$$18 \log 5 = 9\pi + (\pi/4) - \theta$$

at first I was thoroughly confused by this but then I understood what he's trying to do, because we are calculating sin and cos, we want to massage  $t \log n$  in terms of special angles whose sin, cos we know the value of like  $\pi, \pi/3, \pi/4$ , the second reason is we want to minimize  $\theta$  to minimize the error. For example,  $18 \log 5$  is  $\sim 28.97$  so it's  $9\pi + (\pi/4)$  minus some change. This is because we know sin, cos of  $9\pi$  and  $\pi/4$ . Thus

$$\begin{aligned} \sin(9\pi + (\pi/4) - \theta) &= \sin(9\pi + (\pi/4)) \cos(\theta) + \sin(\theta) \cos(9\pi + (\pi/4)) \\ &= -\sin(\pi/4) \cos(\theta) + \sin(\theta) \cos(\pi/4) \end{aligned}$$

and if we minimize  $\theta$ , since more than likely we calculate sin, cos using the series expansion, the smaller  $\theta$  is the better.

**Page 119.**  $\xi(s)$  is real along the vertical line  $\sigma = \frac{1}{2}$ , this is due to its symmetry  $\xi(1-s) = \xi(s)$  and  $\xi(1-s) = \Pi((1-s)/2)((1-s)-1)\pi^{-(1-s)/2}\zeta(1-s)$

$$\begin{aligned}
\xi\left(1 - \left(\frac{1}{2} + it\right)\right) &= \Pi\left(\frac{1 - (\frac{1}{2} + it)}{2}\right) \left(1 - \left(\frac{1}{2} + it\right) - 1\right) \pi^{-(1 - (\frac{1}{2} + it))/2} \zeta\left(1 - \left(\frac{1}{2} + it\right)\right) \\
&= \Pi\left(\frac{\frac{1}{2} - it}{2}\right) \left(\frac{1}{2} - it - 1\right) \pi^{-(\frac{1}{2} - it)/2} \zeta\left(\frac{1}{2} - it\right) \\
&= \Pi\left(\frac{\frac{1}{2} + it}{2}\right) \left(\frac{1}{2} + it - 1\right) \pi^{-(\frac{1}{2} + it)/2} \zeta\left(\frac{1}{2} + it\right) \\
&= \overline{\Pi\left(\frac{\frac{1}{2} + it}{2}\right)} \overline{\left(\frac{1}{2} + it - 1\right)} \overline{\pi^{-(\frac{1}{2} + it)/2}} \overline{\zeta\left(\frac{1}{2} + it\right)} \\
&= \overline{\xi\left(\frac{1}{2} + it\right)}
\end{aligned}$$

But  $\xi\left(1 - \left(\frac{1}{2} + it\right)\right) = \xi\left(\frac{1}{2} + it\right)$ , therefore

$$\xi\left(\frac{1}{2} + it\right) = \overline{\xi\left(\frac{1}{2} + it\right)}$$

and thus  $\xi\left(\frac{1}{2} + it\right)$  is real for any real  $t$ . This is assuming  $\zeta(\bar{s}) = \overline{\zeta(s)}$ . If  $\zeta(s)$  is real all along the real line then we can use the reflection principle but I'm not sure because  $\zeta(s)$  has a pole at  $s = 1$ , also for  $s$  real and  $s > 1$ ,  $\zeta(s)$  is positive but for  $0 < s < 1$ ,  $\zeta(s)$  is negative, see Apostol's Theorem 13.11. For  $\sigma > 1$  we know that  $\zeta(\bar{s}) = \overline{\zeta(s)}$ , how about for the rest  $\sigma \leq 1$ ?

Let me state the answer first and how I got it and then go for the difficulties I had. The analytically continued  $\zeta(s)$  can be expanded in Laurent series about  $s = 1$ , note that Laurent expansion is not Taylor series, it doesn't have radius on convergence since it also includes poles, so the singularities are included. To have  $\zeta(\bar{s}) = \overline{\zeta(s)}$  we must have the coefficient of the Laurent series to be real

$$\begin{aligned}
\zeta(s) &= \sum_{k=-1}^{\infty} a_k (s-1)^k \\
\rightarrow \overline{\zeta(s)} &= \sum_{k=-1}^{\infty} \overline{a_k (s-1)^k} \\
&= \sum_{k=-1}^{\infty} a_k (\bar{s}-1)^k
\end{aligned}$$

we can get to the last line only if the  $a_k$ 's are all real, which will be our task to show that it is so.

For  $s$  real and  $s > 1$ , we know that  $\zeta(s)$  is real, this can mean two things, either all the  $a_k$ 's are all real, or that the imaginary parts cancel out at the end, since  $s$  is real we can split

$$\begin{aligned}\zeta(s) &= \Re \left\{ \sum_{k=-1}^{\infty} a_k (s-1)^k \right\} + i \Im \left\{ \sum_{k=-1}^{\infty} a_k (s-1)^k \right\} \\ &= \sum_{k=-1}^{\infty} a_k^{\text{Re}} (s-1)^k + i \sum_{k=-1}^{\infty} a_k^{\text{Im}} (s-1)^k \\ \rightarrow 0 &= \sum_{k=-1}^{\infty} a_k^{\text{Im}} (s-1)^k\end{aligned}$$

To show that each of the  $a_k^{\text{Im}}$  we need a result (from algebra?) that says that a polynomial that is zero for all  $x \in [c, d]$ ,  $d > c$ , must be exactly zero, *i.e.* each coefficient is zero. While it seems intuitive I was looking for a way show it more formally somehow. The other complication here is that we start with  $\frac{1}{s-1}$  instead with a constant. But an even bigger problem is that this is not just a polynomial of degree  $n$ , this is an infinite power series. So my meandering led me to a few things. First, let's start with the solution. What we have is an infinite power series, multiplying both sides with  $(s-1)$

$$0 = \sum_{k=0}^{\infty} a_{k-1}^{\text{Im}} (s-1)^k$$

the good news is that a Taylor series is also an infinite power series so we can think of this as a Taylor expansion of a function

$$\begin{aligned}f(s) &= f(1) + f'(1)(s-1) + \frac{f''(s)}{2!}(s-1)^2 + \frac{f'''(s)}{3!}(s-1)^3 + \dots \\ &= \sum_{k=0}^{\infty} a_{k-1}^{\text{Im}} (s-1)^k \\ \rightarrow f^{(k)}(1) &= a_{k-1}^{\text{Im}}\end{aligned}$$

but in the interval  $s \in (1, \infty]$ ,  $f(s) = 0$ , this means  $f(s)$  is a constant, which also means that its derivative is also constant, *i.e.* 0, which means that the second derivative must be zero as well, and so on and so  $f'(s) = f''(s) = \dots = f^{(k)}(s) = 0$ . Which means that each

$a_{k-1}^{\text{Im}} = 0$ . Therefore

$$\overline{\zeta(s)} = \zeta(\bar{s})$$

I initially wanted to show that this is the case through two things, one, that we know the sum  $\sum_{\rho} \frac{1}{\rho}$  given in chapter 3 to be real, thus the imaginary parts have to cancel. But this doesn't mean that they come in conjugate pairs, for example we have two complex numbers,  $ae^{i\theta_a}$  and  $be^{i\theta_b}$  then

$$\begin{aligned} \frac{1}{ae^{i\theta_a}} + \frac{1}{be^{i\theta_b}} &= a^{-1}e^{-i\theta_a} + b^{-1}e^{-i\theta_b} \\ &= (a^{-1}\cos(\theta_a) + b^{-1}\cos(\theta_b)) + i(a^{-1}\sin(\theta_a) + b^{-1}\sin(\theta_b)) \end{aligned}$$

and as long as  $b^{-1} = -\frac{\sin(\theta_a)}{a\sin(\theta_b)}$  we have the RHS to be real but they don't form a conjugate pair.

Then I wanted to use the von-Mangoldt's  $\psi(x)$  that includes  $\sum_{\rho} \frac{x^{\rho}}{\rho}$ , again the LHS,  $\psi(x)$ , is real but again, they don't need to form conjugate pairs to cancel out, they just need to cancel out like above, just messier algebra

$$\begin{aligned} \frac{x^{\sigma_1+it_1}}{\sigma_1+it_1} + \frac{x^{\sigma_2+it_2}}{\sigma_2+it_2} &= \frac{x^{\sigma_1+it_1}}{\sigma_1^2+t_1^2} + \frac{x^{\sigma_2+it_2}}{\sigma_2^2+t_2^2} = \frac{x^{\sigma_1+it_1}}{\rho_1^2} + \frac{x^{\sigma_2+it_2}}{\rho_2^2} \\ &= \frac{\rho_2^2 x^{\sigma_1} \cos(t_1 \log x) + \rho_1^2 x^{\sigma_2} \cos(t_2 \log x) + i(\rho_2^2 x^{\sigma_1} \sin(t_1 \log x) + \rho_1^2 x^{\sigma_2} \sin(t_2 \log x))}{\rho_1^2 \rho_2^2} \end{aligned}$$

by adjusting  $t_1$  or  $t_2$  we can cancel the imaginary part but they don't have to be a conjugate pair.

### **Erroneous Excursions**

So I initially wanted to show the above by showing that if a polynomial of degree  $N$  is zero in the interval  $[c, d]$  then it must be exactly zero, although the proof below should be OK, it doesn't apply to an infinite power series, which can also be seen from how the proof goes.

**Polynomial Proposition.** Suppose that a polynomial of degree  $N$ ,  $\sum_{n=0}^N b_n y^n$  is zero for all  $y \in [c, d]$ . We can think of the polynomial as a function  $f(y) = \sum_{n=0}^N b_n y^n$  then for this interval the function  $f(y)$  is constant (which is zero in this case). This also means that the derivative  $f^{(1)}(y)$

$$f^{(1)}(y) = \sum_{n=1}^N n b_n y^{n-1} = 0$$

is also zero in this interval (sans the endpoints). Repeating the argument  $N - 1$  times we get  $f^{(N-1)}(y) = N!b_N = 0$ , which means that the highest coefficient  $b_N$  is zero. Going back to the beginning, repeating the arguments over again we get  $b_{N-1} = 0, b_{N-2} = 0, \dots b_0 = 0$ , *i.e.* each of the coefficient to be zero, thus the whole polynomial is exactly zero.

We can apply the same argument for a polynomial that is equal to a constant instead of zero,  $\sum_{n=0}^N b_n y^n = C$ . We just need to set  $b_0 \rightarrow b_0 - C$ , again, each coefficient must be zero, including  $b_0 - C = 0$ , thus the polynomial is just a constant  $b_0 = C$ .

We can employ the same strategy for a polynomial that starts with a negative power  $\sum_{k=-M}^N b_k y^k$ , since the whole thing is equal zero we just need to multiply the whole thing with  $y^M$ , if the interval contains zero we just need to split the interval  $[c, 0), (0, d]$  to avoid multiplying by zero and focus on one of them since if it is exactly zero in one of the intervals, then it is zero everywhere.

**Another Fundamental Proposition.** We can use the above differentiation strategy to show a weaker form of the fundamental theorem of algebra, *i.e.* a polynomial of degree  $n$  has at most  $n$  real roots. The fundamental theorem asserts there are exactly  $n$  complex roots but we are only interested in reality :)

We can use the above together with induction and an observation that to have more than one zeros we must have extremum points.

Induction base is the case of polynomial of degree one,  $a_1 x + a_0 = 0$ , thus there is only one zero at  $x = -a_0/a_1$ . Inductive hypothesis is that up to degree  $N$  the number of roots is at most the degree of the polynomial. Now we tackle  $N + 1$ . Between any two consecutive zeros there must be an extremum between them, *i.e.* the derivative must have a zero between them. But the derivative of a polynomial of degree  $N + 1$  is a polynomial of degree  $N$  and by the inductive hypothesis this polynomial can have at most  $N$  zeros, since the number of zeros is just the number of extrema plus one, the original polynomial of degree  $N + 1$  can at most have  $N + 1$  zeros. The argument still applies in even the case where we have multiple roots, in this case the roots are at the extremum itself.

As to why they always have roots complex roots, intuitively you can see that since a polynomial is analytic, it obeys the Riemann-Cauchy condition, *i.e.* it's a harmonic function, therefore a maxima in the real line must be a minima in the imaginary line,

since a holomorphic function cannot have true maxima/minima, it only has saddle points thanks to Riemann-Cauchy equation.

Other ways I was trying to show that  $a_k^{\text{Im}} = 0$ , is to calculate the analytically continued integral explicitly as a function of  $s$ . There are at least a couple of difficulty with this

1. We need to set the value of  $s$  before doing the integral  $\int_{+\infty}^{+\infty} \frac{(-x)^{s-1}}{e^x-1}$ . We can't just integrate it and then set  $s$ , this is especially true for the small circle integral as we will see.
2. Even if we can do the integral first, doing the integral itself doesn't really give anything useful, one problem is that the final result must be independent of the small circle radius  $c$ , but to show that this is so is not straightforward.

Let's see what happens when we integrate the small circle, we can actually do this with the help of the Bernoulli numbers

$$\begin{aligned}
\lim_{\epsilon \rightarrow 0} \int_{0+\epsilon}^{2\pi-\epsilon} \frac{(-x)^s}{e^x-1} \frac{dx}{x} &= \lim_{\epsilon \rightarrow 0} \int_{0+\epsilon}^{2\pi-\epsilon} (-x)^{s-2} \frac{x}{e^x-1} dx \\
&= \lim_{\epsilon \rightarrow 0} \int_{0+\epsilon}^{2\pi-\epsilon} (-x)^{s-2} \sum_{n=0}^{\infty} \frac{B_n x^n}{n!} dx \\
&= \lim_{\epsilon \rightarrow 0} \sum_{n=0}^{\infty} (-1)^s \int_{0+\epsilon}^{2\pi-\epsilon} \frac{B_n x^{s+n-2}}{n!} dx \\
&= \lim_{\epsilon \rightarrow 0} \sum_{n=0}^{\infty} (-1)^s i \int_{0+\epsilon}^{2\pi-\epsilon} \frac{B_n (ce^{i\theta})^{s+n-2}}{n!} (ce^{i\theta}) d\theta \\
&= \lim_{\epsilon \rightarrow 0} \sum_{n=0}^{\infty} (-1)^s \frac{B_n c^{s+n-1}}{n!(s+n-1)} (e^{i(2\pi-\epsilon)(s+n-1)} - e^{i(0+\epsilon)(s+n-1)}) \\
&= \sum_{n=0}^{\infty} (-1)^s \frac{B_n c^{s+n-1}}{n!(s+n-1)} (e^{i2\pi(s+n-1)} - 1)
\end{aligned}$$

Here we see the first problem, if we set  $s$  to be a negative integer before the integration we will get the values of  $\zeta(-n)$  but if we do it after the integration we will just get zero since  $e^{2\pi i(s+n-1)} = 1$  if  $s+n-1$  is an integer.

Say  $s$  is not an integer so we can still do the integral, the problem is that this is a function of  $c$  (and it is obvious that the complex conjugate of this integral is not the same as the integral of  $\bar{s}$  because conjugating  $e^{2\pi i(s+n-1)}$  will not only turn  $s \rightarrow \bar{s}$  but also adding a negative sign in front of the  $i$ ).

To remove the dependency on  $c$ , we can integrate the rest of the integrals  $\int_c^\infty \frac{(-x)^{s-1}}{e^x-1}$ , however, if we expand it using Bernoulli's numbers, the final integral we get will be of the form  $\int_c^\infty x^{s+n-2}$  and the upper limit will give us issues.

**Page 119, Footnote.**  $Z(t)$  is real when  $t$  is real, we can see this easily from

$$\xi\left(\frac{1}{2} + it\right) = \left[ e^{\operatorname{Re} \log \Pi(s/2)} \pi^{-1/4} \frac{-t^2 - \frac{1}{4}}{2} \right] Z(t)$$

since the LHS is real and the first term on the RHS is real,  $Z(t)$  must be real as well.

**Page 120.** The goal here is to express everything in terms of powers of  $\left(\frac{1}{t}\right)$ .

About the symmetry of  $\frac{\xi'(s)}{\xi(s)}$ , for simplicity let's denote  $\frac{\xi'(s)}{\xi(s)} = F(s)$ ,

$$\int_{1\frac{1}{2}}^{1\frac{1}{2}+iT} F(s)ds = i \int_0^T F(1\frac{1}{2} + it)dt$$

while on the opposite side we have

$$\int_{-\frac{1}{2}+iT}^{-\frac{1}{2}} F(s)ds = i \int_T^0 F(-\frac{1}{2} + it)dt$$

next is to use the symmetry  $\xi(s) = \xi(1-s)$  but we need to be careful in using the symmetry for  $F(s)$  because for the derivative  $\xi'(s)$

$$\left. \frac{d}{ds} \xi(s) \right|_{s=w} = - \left. \frac{d}{ds} \xi(s) \right|_{s=1-w}$$

this is analogous to our familiar friend the cosine function, even though it's symmetric with respect to the  $y$  axis it's derivative is the opposite of one another across the vertical axis

$$\left. \frac{d}{dx} \cos(x) \right|_{x=y} = - \left. \frac{d}{dx} \cos(x) \right|_{x=-y}$$

Therefore, since  $\xi(s) = \xi(1-s)$  we have  $\frac{\xi'(1-s)}{\xi(1-s)} = F(1-s) = -F(s)$  so the second integral becomes

$$\begin{aligned} i \int_T^0 -F(1 - (-\frac{1}{2} + it))dt &= i \int_0^T F(1\frac{1}{2} - it)dt = i \int_0^T F(\overline{1\frac{1}{2} + it})dt \\ &= i \int_0^T \overline{F(1\frac{1}{2} + it)}dt \\ &= i \left( \overline{\int_0^T F(1\frac{1}{2} + it)dt} \right) \end{aligned}$$



Now, what we originally want was  $\text{Im}[\int_C F(s)]$ , since there is a factor of  $i$  in front of the integral what we really want is the real part, but

$$\text{Re} \left( \overline{\int_0^T F(1\frac{1}{2} + it)dt} \right) = \text{Re} \left( \int_0^T F(1\frac{1}{2} + it)dt \right)$$

and so the two integral along the vertical segments are equal.

We now need to show that the integral on the two horizontal segments are equal, the steps are similar, on the right segment we have

$$\int_{1\frac{1}{2}}^{\frac{1}{2}} F(x + iT)dx$$

while on the left segment we have

$$\int_{\frac{1}{2}}^{-\frac{1}{2}} F(x + iT)dx$$

using  $F(1 - s) = -F(s)$

$$\begin{aligned} \int_{\frac{1}{2}}^{-\frac{1}{2}} F(x + iT)dx &= \int_{\frac{1}{2}}^{-\frac{1}{2}} -F(1 - x - iT)dx, & y = 1 - x \\ &= \int_{\frac{1}{2}}^{1\frac{1}{2}} F(y - iT)dy = \int_{\frac{1}{2}}^{1\frac{1}{2}} F(\overline{y + iT})dy \\ &= - \left( \overline{\int_{1\frac{1}{2}}^{\frac{1}{2}} F(y + iT)dy} \right) \end{aligned}$$

in this case since there's no  $i$  multiplying it we want the imaginary part of the integral and the overall minus sign helps us because the complex conjugate adds an extra minus sign to the imaginary part thus

$$\text{Im} \left[ - \left( \overline{\int_{1\frac{1}{2}}^{\frac{1}{2}} F(y + iT)dy} \right) \right] = \text{Im} \left[ \int_{1\frac{1}{2}}^{\frac{1}{2}} F(y + iT)dy \right]$$

and the integrals along the two segments are equal.

**Branches and Cuts.** I initially wanted to show the above symmetry using the Fundamental Theorem of Calculus. Since the integrand is nothing but  $\frac{d \log \xi(s)}{ds}$ , if we integrate

it we just get back  $\log \xi(s)$ . But this is a dangerous thing to do since  $\log \xi(s)$  has many branch points, one for each of the zeros of  $\xi(s)$ .

A pertinent example would be  $\log \left( \frac{z+1}{z-1} \right)$

**Brief (Pedestrian) Review of Branches and Cuts.** We have branches because we have multi-valued functions. If we restrict the function to be single valued, it is said that we are choosing a branch. For example, how do we create single-valued functions out of  $f(z) = z^{1/3}$ ?

One way is by creating three different single-valued functions,  $f_1(z) = |z|^{1/3} e^{i\bar{\theta}/3}$ ,  $f_2(z) = |z|^{1/3} e^{i\bar{\theta}/3 + 2\pi i/3}$  and  $f_3(z) = |z|^{1/3} e^{i\bar{\theta}/3 + 4\pi i/3}$  where  $z = |z|e^{i\theta}$  and  $\bar{\theta} = \theta \pmod{2\pi}$ .

The first,  $f_1(z)$  will generate numbers with angle  $0 \leq \arg(f_1(z)) < \frac{2\pi}{3}$ , the second  $f_2(z)$  generates  $\frac{2\pi}{3} \leq \arg(f_2(z)) < \frac{4\pi}{3}$  and the third  $f_3(z)$  produces numbers with  $\frac{4\pi}{3} \leq \arg(f_3(z)) < \frac{6\pi}{3}$ . This way we cover all possible values of  $z^{1/3}$ .

Rather than splitting the function into three single valued ones we can split the domain into three complex planes. Let's take a simpler example of  $z^{1/2}$ . We can again create two single valued functions  $f_R(z) = |z|^{1/2} e^{i\bar{\theta}/2 - \pi i/2}$  whose output is in the right half plane and  $f_L(z) = |z|^{1/2} e^{i\bar{\theta}/2 + \pi i/2}$  whose output is in the left half plane.

But wouldn't it be simpler if we can just designate  $f(z) \rightarrow |z|^{1/2} e^{i\theta/2}$  while still maintaining the output to be always say for instance on the right half of the plane? This looks a lot more natural since taking a square root is (primarily) just multiplying the exponent by  $\frac{1}{2}$ .

We can do this by setting the domain to be  $z = |z|e^{i\theta}$  where  $\theta$  is now restricted to  $\theta \in (-\pi, \pi)$ , since the interval is open, this is equivalent to excluding the negative real axis. With this choice of domain our  $f_R$  simply becomes  $f_R(z) = |z|^{1/2} e^{i\theta/2}$  which is very natural (compared to  $f_R(z) = |z|^{1/2} e^{i\bar{\theta}/2 - \pi i/2}$ ).

What we are doing is essentially creating two copies of the complex plane, one for  $f_R$  and another for  $f_L$ . In their own planes,  $f_R$  and  $f_L$  take the same form  $f_L(z) = f_R(z) = |z|^{1/2} e^{i\theta/2}$ . However, the argument of  $z$  is now different in these two planes.

In the  $f_R$  plane,  $\theta \in (-\pi, \pi)$  and in the  $f_L$  plane  $\theta \in (-3\pi, -\pi)$  (we are excluding the negative real axis). These choices of angles make sense, by limiting  $\theta \in (-\pi, \pi)$  we essentially limit  $-\frac{\pi}{2} < \arg(f_R(z)) < \frac{\pi}{2}$ , which is the right half of the plane, and

$\theta \in (-3\pi, -\pi)$  means  $-\frac{\pi}{2} < \arg(f_L(z)) < -\frac{3\pi}{2}$  which is the left half of the plane.

To make things clearer, in the  $f_R$  plane we start the angle from slightly above the negative real axis with  $\pi$ , we then go clockwise, the positive real axis has the angle zero and when we reach slightly below the negative real axis we get  $-\pi$ . For the  $f_L$  plane we start slightly above the negative real axis with  $-\pi$  going clockwise, the positive real axis is now at angle  $-2\pi$  and when we are at slightly below the negative real axis we are at  $-3\pi$ .

Now, just because  $\theta \in (-3\pi, -\pi)$  in the  $f_L$  plane, it doesn't mean that  $w_L = f_L(z)$  cannot have  $\arg(w_L)$  bigger than  $-\pi$ , for example,  $f_L(e^{-i\pi}) = e^{-i\pi/2}$ , this is fine since this is the angle of  $w_L$  and not of  $z$ , we only restrict the angle of  $z$ .

From the illustration above it is clear that if we start with the  $f_R$  plane at slightly above the negative real axis and go clockwise one round and still continues to more negative angles  $< -\pi$  we are crossing into the  $f_L$  plane. The negative real axis that we exclude is called a branch cut. Intuitively it is where we glue together the two planes as illustrated above. So if we cross it, our single valued function will evolve into another (different) one, this is why we are not supposed to cross the branch cut whether when we're doing integral or not.

For our first example  $z^{1/3}$ , instead of defining three separate single-valued functions like before, we can create three separate planes which can be affected by restricting the angle of  $z$ .

Recall that  $f_1(z) = |z|^{1/3}e^{i\bar{\theta}/3}$ , *i.e.* the output of this function has angle  $0 \leq \arg(f_1(z)) < \frac{2\pi}{3}$  a.k.a the first 120° pizza slice and the other two  $f_2$  and  $f_3$  has output in the subsequent 120° pizza slices respectively. We can instead have  $f_{1,2,3}$  all have the same form  $|z|^{1/3}e^{i\theta/3}$  by defining a branch cut  $\theta \in (0, 2\pi)$  which excludes the positive real axis.

Let's verify the outcome, well,  $f_1$  is obvious,  $f_2$  we get by going to the second plane which is  $\theta \in (2\pi, 4\pi)$  so that  $\frac{2\pi}{3} < \arg(f_2) < \frac{4\pi}{3}$  which is the second 120° pizza slice and to get  $f_3$ , again, we cross the branch cut a second time from  $f_2$  plane going to  $f_3$  plane, this time we restrict  $\theta \in (4\pi, 6\pi)$  so that  $\frac{4\pi}{3} < \arg(f_2) < \frac{6\pi}{3}$  which is the last 120° pizza slice. If we cross the branch cut again, we will end up in  $f_1$ 's plane but this time we have

$\theta \in (6\pi, 8\pi)$  with the resulting output  $\frac{6\pi}{3} < \arg(f_1) < \frac{8\pi}{3}$  which is the first pizza slice again :)

So that's the story on how we got branch cuts and how it relates to multi-valued functions. In essence

1. If we have a multi-valued function, we need to make it single-valued.
2. We can either explicitly generate different single-valued functions or restrict the domain using branch cuts.
3. Branch cuts are the way to go, this way we create multiple complex planes, one for each of the single valued functions, and in this way all of the single-valued functions have the same form.
4. We must not cross branch cuts because then we are going into a different plane and therefore a different function (although the function has the same form).
5. Branch cuts need not be straight lines and they don't have to be on the real axis, they can be some wavy lines as long as they end on a branch point and not cross each other.
6. The final goal is to make the function single-valued and analytic, so it's not about assigning angles to  $z$ , we do it to make the function well-behaved, we'll see an example later.

Talking about branch points, in the above examples of  $z^{1/2}$  and  $z^{1/3}$ , the branch points are  $z = 0$  and  $z = \infty$ . What is a branch point? A branch point can be defined (at least) two ways. The first not so pedestrian definition is that it is a point  $z_0$  where for any  $\epsilon > 0$ ,  $f(z)$  is multivalued for all points in its neighborhood, *i.e.* all  $z$  such that  $|z - z_0| \leq \epsilon$ .

The more pedestrian definition is, a branch point is a point where the multi-valued function stopped being multi-valued. An even more pedestrian definition is that a branch point is a point we use to define other points, this is what we usually do to calculate things. Let's see how this goes.

The above way of restricting the angle,  $\theta \in (0, 2\pi)$  for example, is not the best way of expressing how to calculate with a branch cut. A better way is to first pick a branch point

and then use it as a reference to define other points in the domain. Let's take  $(z - 1)^{1/2}$  for example.

The branch point here is  $z = 1$  because here the function stops being multi-valued. We can then express all other points with respect to  $z = 1$ . Note that numbers in the complex plane are like vectors, so if we choose  $z = 1$  as a reference, all other points can be expressed as  $z = 1 + re^{i\alpha}$  where  $r = |z - 1|$  and  $\alpha$  is the angle we need to define to make things work.

To define this angle  $\alpha$ , we first need to choose a branch cut. Say we choose the cut to be the real axis with  $\text{Re}(z) > 1$ , it doesn't mean that the angle has to be  $(0, 2\pi)$ . We can still assign  $(-\pi, \pi)$  even though this means that the angle sort of cuts through the cut, this is fine.

The problem with  $(-\pi, \pi)$  is that at  $z = 0$ ,  $(z - 1)^{1/2}$  will have a discontinuity, because at  $z = 0^+$  we get  $(0^+ - 1)^{1/2} = 1 + e^{i\pi/2}$  while at  $z = 0^-$  we have  $(0^- - 1)^{1/2} = 1 + e^{-i\pi/2}$ . This was my big misunderstanding that I thought the angle assignments must not cut across the cut, rather it must make the function well behaved.

To see an example of angle assignment that cuts through the branch cut but still makes the function well behaved we go to the next non-trivial example,  $\log\left(\frac{z+1}{z-1}\right)$ .

This time we have two branch points,  $z = 1$  and  $z = -1$ . The most common choices of branch cuts are a cut from  $-1$  to  $+1$  or a cut from  $-1$  to  $-\infty$  and  $+1$  to  $+\infty$ . We'll go through both of them.

For the first choice, we define the angles like shown, to make

As an application of what we learned above let's see what happens when we have a multi-valued function inside an integral and as a result of an integral. We'll start with a simple example and move on to more complicated ones.

Analytic continuation, we can continue  $\log_{(-\pi, \pi)}$  with  $\log_{(0, 2\pi)}$  because they have the same value in the upper half plane.

The way we do branch cuts

Example of multi-valued function in an integral and as a result of an integral (Fundamental Theorem).

It's not always straightforward as to what to do to split the angle, especially when we

have many branch points, a pertinent example would be  $\log\left(\frac{z+1}{z-1}\right)$

$z^p$  has branch points at 0 and infinity, at infinity because  $z^p = e^{p \log z}$  and  $\log z$  has branch points at 0 and infinity through  $\log \frac{1}{z} = -\log z$

The branches dictate the angle assignment, not the other way round

Each branch is on its own sheet and single valued. Connected through the cut, this is why we cannot go across the cut since then we are going to a different branch of the function which is a different function. Give example.

Can only use fundamental if the function is fundamental, thus must choose a branch and a cut.

The expansion of  $(z^2 - 1)^{\frac{1}{2}}$  is  $z - \frac{1}{2z} + \dots$

So you don't want  $\zeta$  to be 0 along  $C$  but you only want  $\text{Re}\zeta(s)$  to be zero. But beyond  $\text{Re}(s) > 1$  there's no zero.

Baclund's estimate of  $N(T)$ , towards the end it has

$$|\text{Re}\zeta(2 + iT)| \geq 1 - \frac{1}{2^2} - \frac{1}{3^2} - \frac{1}{4^2} - \dots$$

This is because

$$\begin{aligned} \text{Re}\zeta(2 + iT) &= \sum_{n=1}^{\infty} n^{-2} \cos(T \log n) \\ &= 1 + \sum_{n=2}^{\infty} n^{-2} \cos(T \log n) \\ &\geq 1 + \sum_{n=2}^{\infty} n^{-2} (-1) \end{aligned}$$

as  $\cos \geq -1$ .

**Page .** We don't have to use Taylor expansion, we just need to massage it into a power series.

**Page .** So the contour used for this saddle point is a peculiar one, see Fig. 3, it starts from infinity,  $a + \infty \cdot e^{i\pi/4}$  to  $a - \frac{1}{2}|a|e^{i\pi/4}$ , vertically down to the line  $\text{Im}(x) = (2N + 1)\pi$  and to infinity again on this same line. The goal is to show that only the integral on the segment  $L_1$  is significant.

To make sense of this integral, first do the following variable substitution,  $x = a + ke^{i\pi/4}$ ,  $dx = dk$  with  $a = i(2\pi t)^{1/2}$  and  $k$  starting from  $\frac{1}{2}|a|$ , again see Fig. 3, note also

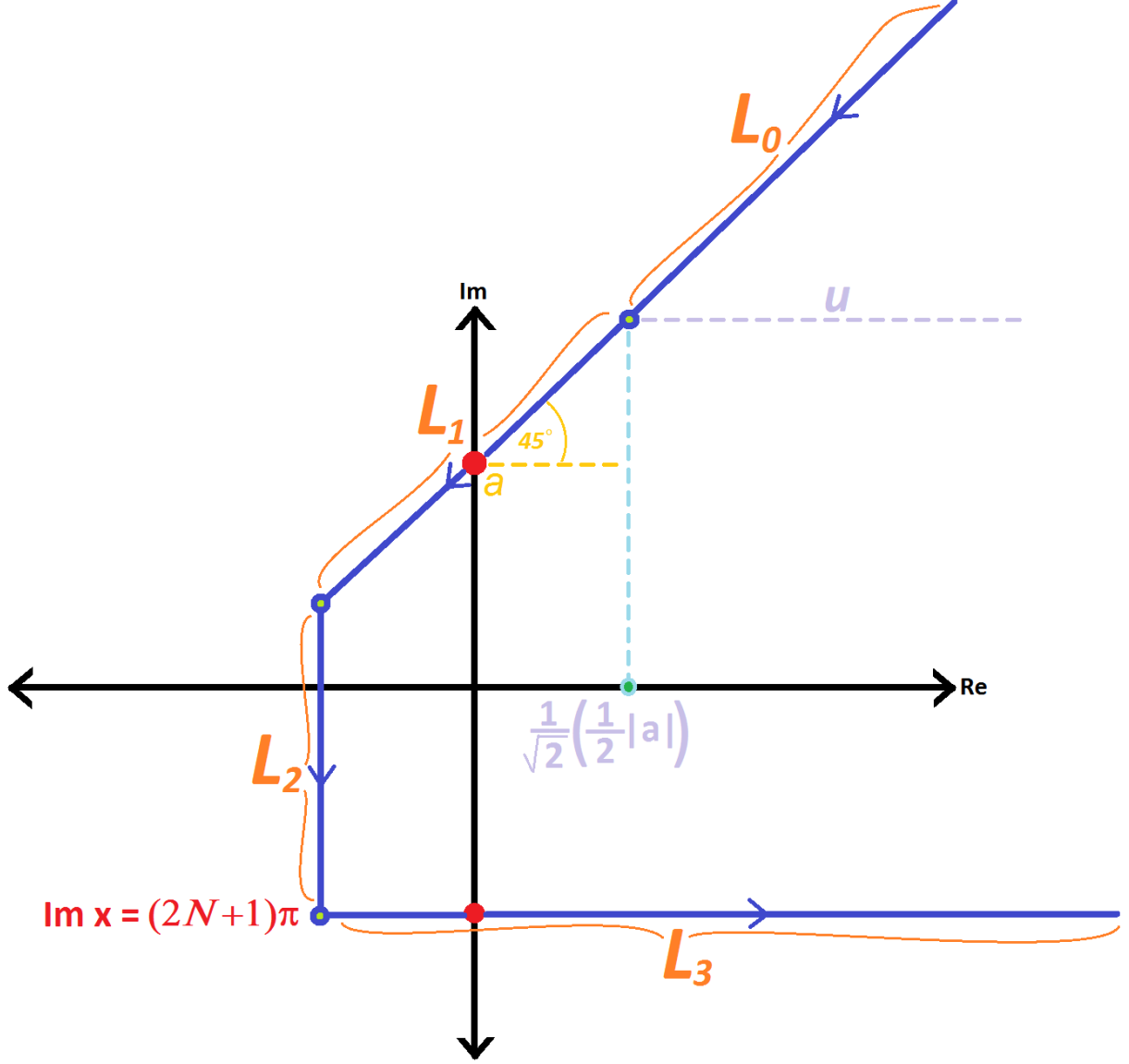


FIG. 3: Contour of integration used for the saddle point method.

that we are taking the modulus, so we can flip the integration direction to go from  $a$  to infinity rather than from infinity to  $a$ , therefore the integral becomes

$$\begin{aligned}
 \left| \int_{L_0} \frac{dx}{e^x - 1} \right| &= \left| \int_{-L_0} \frac{dx}{e^x - 1} \right| \leq \int_{L_0} \frac{|dx|}{|e^x - 1|} \\
 &= \int_{\frac{1}{2}(2\pi t)^{1/2}}^{\infty} \frac{|dk|}{|e^x - 1|} \\
 &= \int_{\frac{1}{\sqrt{2}}(\pi t)^{1/2}}^{\infty} \frac{dk}{|e^{a+ke^{i\pi/4}} - 1|}
 \end{aligned}$$

Now we use the less famous triangle inequality  $|a - b| \geq |a| - |b|$

$$\begin{aligned} |e^{a+ke^{i\pi/4}} - 1| &\geq |e^{a+ke^{i\pi/4}}| - 1 \\ &= |e^a| |e^{k \cos(\pi/4) + ik \sin(\pi/4)}| - 1 \\ &= e^{\frac{1}{\sqrt{2}}k} - 1 \end{aligned}$$

We can now do another variable change  $u = \frac{1}{\sqrt{2}}k, dk = \sqrt{2}du$  and the integral becomes

$$\begin{aligned} \int_{\frac{1}{\sqrt{2}}(\pi t)^{1/2}}^{\infty} \frac{dk}{e^{\frac{1}{\sqrt{2}}k} - 1} &= \int_{\frac{1}{\sqrt{2}}(\pi t)^{1/2}}^{\infty} \frac{\sqrt{2}du}{e^u - 1} \\ &= \int_{\frac{1}{2}(\pi t)^{1/2}}^{\infty} \frac{\sqrt{2}du}{e^u - 1} \\ &= \int_{\frac{1}{2}(\pi t)^{1/2}}^{\infty} \frac{e^{-u}\sqrt{2}du}{1 - e^{-u}} \end{aligned}$$

The way I figured out this integral is by noting that integration along  $k$  is somewhat like integration along the dotted line  $u$  in Fig. 3 but we need to of course take care of the fact that (from the triangle they form) the line  $u$  is  $\frac{1}{\sqrt{2}}k$ .

**Page .** Dealing with log in the complex plane is a bit tricky, we usually have to designate a cut to avoid certain numbers having multiple log values, for example, 1 can be represented by  $e^{i0}$  or  $e^{i2\pi}$ , the log of which are 0 and  $i2\pi$  respectively. If we made the cut along the positive real axis we prevent  $e^{i0}$  to becoming  $e^{i2\pi}$ , if the cut is on the negative real axis we prevent  $e^{i\pi}$  from becoming  $e^{-i\pi}$  while the positive real axis is set to be  $re^{i0}$ , we can set other conventions of course even with the same cut.

But in this case we are not bothered by this, I initially wrongly thought that the reason the starting point of  $L_2$  cannot have the min  $\text{Im}(\log(-x))$  was because in this case  $-x$  would be in the south east quadrant and the integration contour  $C_M$  dictates that we make the cut on the positive real axis, so we have  $0 \leq \theta < 2\pi$  starting from the positive real axis going counter clockwise. In this south east quadrant the argument of  $-x$  would be near  $2\pi$ . While the end point of  $L_2$  would have  $-x$  in the north east quadrant, thus the argument of  $-x$  would be the smallest.

Then I thought maybe the starting point of  $L_2$  has the smallest argument because

$$\text{Im}(\log(-x)) = \text{Im}(i\pi + \log(x))$$



so to minimize  $\text{Im}(\log(-x))$  is the same as minimizing the argument of  $x$  and since the starting point of  $L_2$  is in the north west quadrant it has the smallest argument, again assuming the cut is in the positive real axis.

But I was wrong, this is just about a lower bound, we don't care about cut here. If you see the derivation he just wanted the most negative argument of all, since the starting point of  $L_2$  has  $-x$  in the south east quadrant, its argument is negative while its end point in the north east quadrant would have a positive argument so we chose the one with the negative argument, that's all.

Lesson of the day, doing integrals without doing integrals.

We Taylored  $\Psi(p + y)$  around  $p$

$$\begin{aligned}\Psi(p + y) &= \Psi^{(0)}(p) + \Psi^{(1)}(p)(y + p - p) + \frac{1}{2!}\Psi^{(2)}(p)(y + p - p)^2 + \dots \\ &= \sum_{k=0}^{\infty} \frac{\Psi^{(k)}(p)}{k!} y^k\end{aligned}$$

“... and in which sum is finite”, “which” here refers to  $\sum_{j=0}^{\infty} A_j \omega^j$

So the goal here is to calculate  $b_0 c_0 + b_1 c_1 + \dots + b_K c_K$ , but the way Riemann did it is not to calculate this sum at all, the way he did it was to calculate an approximation to this sum by expressing it in terms of a power series of  $\omega^j$ . I say an approximation because  $b_K$  contains  $\omega_K$  while we are only mucking around with  $\omega^\nu$  where  $K$  is large enough such that  $b_{K+1}, b_{K+2}, \dots$  does not contain  $\omega^\nu$ .

This is what confused me when he expressed  $G(y)$  as  $\sum_{j=0}^{\infty} A_j \omega^j$  and he was only concerned about  $A_0, A_1, \dots, A_{\nu-1}$  that are not dependent on  $K$ . Why care only up to  $A_{\nu-1}$ , the goal here is to calculate  $b_0 c_0 + b_1 c_1 + \dots + b_K c_K$ , we need the  $A$ 's so that we can get the coefficient of negative powers of  $y$  so that we can combine them with positive powers of  $y$  in  $\sum_{k=0}^{\infty} \frac{\Psi^{(k)}(p)}{k!} y^k$  to get the coefficient of  $y^0$ , but the coefficient of  $y^0$  is exactly  $b_0 c_0 + b_1 c_1 + \dots + b_K c_K$ , it includes  $\omega^K$ , so I figured that what we are doing here is an approximation of  $b_0 c_0 + b_1 c_1 + \dots + b_K c_K$ .

$$(2y)^{n-1}(n+1)b_{n+1}$$

res f, only is f is holomorphic, no branch cut nothing