Worksheet: Prime Number Theorem

John D Mangual

Studying the Prime Number Theorem was meant just to be a template of a much broader class. In my opinion, the arguments are not very visual, or "elementary". Many people are making the point that "elementary" proofs are more complicated then the Complex Analysis proof.

If I complain that way, and have received such a negative response, then I can argue in a visual elementary way, and the proof should new. Since all PNT problems seem to be about the hardness, I could imagine two engines for a proof:

- Stirling's formula

- Horocycle flow

So far I haven't been able to give a complete answer. Every day I'm learning a bit.

The argument's from the 1930's and 1940's all have a similar flavor. Starting from Stirling's formula for $n!$.

#1 The first step of Norbert Weiner's argument looks like Möbius inversion, but possibly the lattice point counting on a hyperbola:

$$\sum_{m=0}^{\infty} (\log m) x^m = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \Lambda(n) \, x^{-mn} = \sum_{n=0}^{\infty} \Lambda(n) \frac{x^n}{1 - x^n}$$

If I evaluate at $x = 1$, the left side looks like logarithm of factorial:

$$\log\left(\infty!\right) = \sum_{m=0}^{\infty} \left(\log m\right) = \sum \Lambda(n) \frac{1}{0}$$

Writing it this way would have been offensive to everyone at the time. We spend the rest understand the behavior at $x \approx 1$. This is my **regularization** approach to PNT. I don't think anyone else thinks of it quite this way.

I tried setting $x = 1 + \epsilon$. This is has a few derived approximations:

$$x^m = (1 + \epsilon)^m = 1 + m\epsilon + \dots \quad \text{and} \quad \frac{x^n}{1 - x^n} = \frac{(1 + \epsilon)^n}{1 - (1 + \epsilon)^n} = \frac{1 + n\epsilon}{1 - (1 - n\epsilon)} = 1 + \frac{1}{n\epsilon}$$

and if some small things get away, we can look at with a microsope and get more terms.

$$\sum_{m=0}^{\infty} (\log m)\left(1 + m\epsilon\right) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \Lambda(n) \left(1 + mn\epsilon\right) = \sum_{n=0}^{\infty} \frac{\Lambda(n)}{n}\left(1 + \frac{1}{n\epsilon}\right)$$

Algebra always does more than we tell it to. While this algebra is not correct, we can read it and know a little bit

I do not like the way the next few lines were written so, I can do it over. It's not obvious at all to me that:

$$\sum_{m=1}^{\infty} \left(\log m \right) \frac{x^m - x^{m+1}}{1-x} \stackrel{??}{=} \frac{x}{1-x} \sum_{m=1}^{\infty} \log \left(1 + \frac{1}{m} \right) x^m$$

Norbert Weiner is suggesting there is an equality here. There is also someting about Möbius transformations:

$$x \mapsto x+1 \quad \text{and} \quad x \mapsto \frac{x}{1-x} \quad \text{or} \quad m \mapsto 1 + \frac{1}{m}$$

Issues like this are not commented on in the paper. He had other things in mind.

$$\log \left(1 + \frac{1}{m} \right) - \frac{1}{m} = O \left(\frac{1}{m^2} \right)$$

He needed to keep track of exactly how bad the approximation $\log(1+m^{-1}) \approx \frac{1}{m}$ could be:

$$\frac{x}{1-x} \sum_{m=1}^{\infty} \left[\frac{1}{m} + O \left(\frac{1}{m^2} \right) \right] x^m = \frac{x}{1-x} \sum_{m=1}^{\infty} \left[\frac{1}{m} \right] x^m + \underbrace{\frac{x}{1-x} \sum_{m=1}^{\infty} \left[O \left(\frac{1}{m^2} \right) \right] x^m}_{??}$$

We don't know how large the derivative could be... but before we do that, what have we been estimating?

#1 : crazy regularized factorial

$$(\infty!)(x) := \sum_{m=0}^{\infty} (\log m) x^m \approx \frac{x}{1-x} \sum_{m=1}^{\infty} \left[\frac{1}{m} + O \left(\frac{1}{m^2} \right) \right] x^m$$

#2 : (derivative of) crazy regularized factorial

$$\xi \frac{d}{d\xi} \left[\xi \, (\infty!)(e^{-\xi}) \right] = \frac{\xi e^{-\xi}(1 - \xi - e^{-\xi})}{(1-e^{-\xi})^2} \left[\log(1 - e^{-\xi}) - \sum_{m=1}^{\infty} O \left(\frac{1}{m^2} e^{-m\xi} \right) \right]$$

$$+ \frac{\xi^2 e^{-\xi}}{1 - e^{-\xi}} \left[\frac{e^{-\xi}}{1 - e^{-\xi}} + \sum_{m=1}^{\infty} O \left(\frac{1}{m} \right) e^{-m\xi} \right]$$

What does that mean... to take a derivative of a random series? Even if it's small?

In the end... since I don't really know what he's saying and nobody's really there to explain to me, no not much progress.

Everyone is busy. I understand. It can wait.

Norbert Weiner was the grandfather of random process theory – the same collection of ideas we can use to study the weather or the stock market. What's it doing here in Prime Number Theory, which I thought would be pretty regular?

Doesn't this look like the partition function from statistical thermodynamics? I'm getting really skeptical.

Yet we proceed...

We've basically shown that:

$$\lim_{\xi \to 0} \xi \sum_{n=1}^{\infty} \Lambda(n) \frac{d}{d(n\xi)} \left(\frac{n\xi \, e^{-n\xi}}{e^{-n\xi} - 1} \right) = 1$$

or if we want to be "precise"

$$\xi \sum_{n=1}^{\infty} \Lambda(n) \frac{d}{d(n\xi)} \left(\frac{n\xi \, e^{-n\xi}}{e^{-n\xi} - 1} \right) = 1 + O(\xi \log \xi)$$

This is the Möbius inversion of the formulas on the previous page.

$$N_1(u) = \text{``}g(u)\text{''} = \frac{d}{du} \left(\frac{u}{1 - e^u} \right) = \begin{cases} O(1) & u \to 0 \\ O(ue^{-u}) & u \to \infty \end{cases}$$

This "kernel" function is behaving one way for small numbers $x \ll 1$ and one way for large number $x \gg 1$. This is not the magical part anyway.

Somehow:

$$\left[\xi \sum_{n=1}^{\infty} \Lambda(n) \frac{d}{d(n\xi)} \left(\frac{n\xi \, e^{-n\xi}}{e^{-n\xi} - 1} \right) = 1 + O(\xi \log \xi) \right] \to \left[\xi \sum_{n=1}^{1/\xi} \Lambda(n) = 1 + O_?(\xi) \right]$$

Turning the limit on the left into the limit on the right is where the bulk of the work remains.

Whether we're done really depends on who you are:

- number theorists are very comfortable manupulating sum without a "picture" of what's going on; that this discussion is "done"

- a physicist would be satisfied stopping here without any Tauberian theorem. And "intepretaion" of this discussion could be nice?

- Quantum Field Theory routinely uses divergent series of a kind that GH Hardy would have disapproved, even though he wrote an excellent book called **Divergent Series**.

- information theory? Did not exist until Claude Shannon in 1950.

- geometry? The mixing of the horocycle flow wouldn't be studied until the 1970's or 1990's ?

We're in the early 21st century and a lot as changed. And it's a little bit dangerous for me to get stuck on one problem.

References

(1) GH Hardy **Divergent Series** Oxford University Press, 1949/1973.

That went by really fast. Maybe later I can review to, very similar instances.

I was trying to show that Gaussian primes – in the ring $\mathbb{Z}[i]$ are (almost) rotationally symmetric. Here is the picture of that:
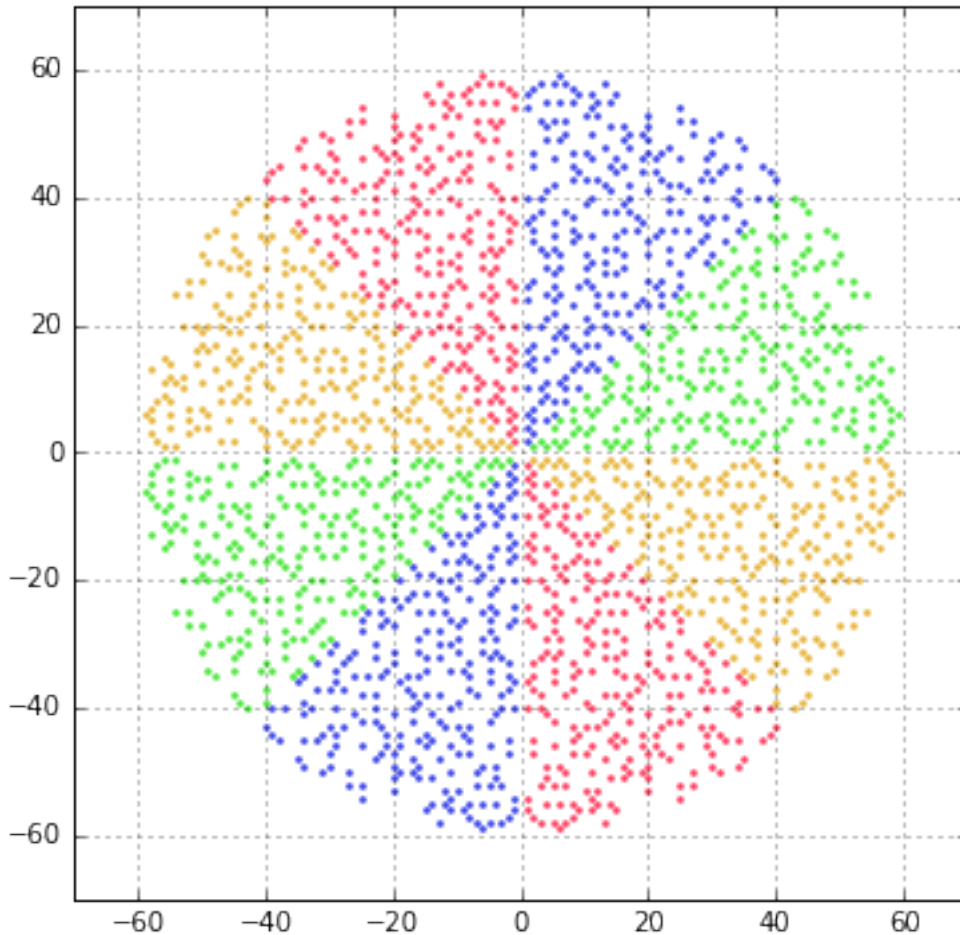
Number in $\mathbb{Z}[i]$ can factor that do not factor in $\mathbb{Z}$. For example, 5 is prime but

$$5 = (2 + i)(2 - i)$$

and we know that all primes $p = 4k + 1$ factor. Lookup a large prime number and try again:

$$7109 = 70 \times 70 + 47 \times 47 = (70 + 47\,i)(70 - 47\,i)$$

When I asked online[1] (someone else asked) I was referred to graduate-level textbook **Algebraic Number Theory** by Serge Lang. It's really dull and I didn't learn much.

The Tauberian Theorem used is so severe, it's almost meaningless. Let ϕ be a function of "bounded variation".

$$f(s) = \int_0^\infty e^{-sx}\,d\phi(x)$$

This doesn't look like a Dirichlet series, but we can set:

$$\phi(x) = \sum_{n \leq x} a_n\,n^{-s} \quad \text{then} \quad f(s) = \sum_{n \leq x} a_n e^{-nx} n^{-s}$$

We look for cases when $f(s)$ converges for $\mathrm{Re}(s) > 1$. All Dirichlet series look convergent to me. Hopefully $f(s) < \infty$ – is this too much to ask for?

---

[1] https://mathoverflow.net/questions/133410/hecke-equidistribution

4

We asked everything behave good for numbers to the right of a certain line, $z = s + it$ with $s > 1$. Then we ask for some good (relatively OK) behavior at $s = 1 + \epsilon + it$:

$$h_\epsilon(t) = f(s) - \frac{1}{s-1}$$

The statement of the Tauberian theorem looks really tautological (doesn't say anything):

$$\left[\lim_{\epsilon \to 0} \left(f(1 + \epsilon + it) - \frac{1}{\epsilon + it} \right) = f(1 + it) - \frac{1}{it} \right] \to \left[\lim_{x \to \infty} e^{-x}\phi(x) = 1 \right]$$

I am almost willing to use complex analysis at this point. Especially since I was concerned about points in $\mathbb{Z}[i]$.

This Dirichlet series has bounded variation. Doesn't sound like very much Number Theory. So I backtracked to this obscure textbook on Laplace transforms, and got the softer proof in the previous section.

These statements look very different and we're being told they have the structure:

- as $p \to \infty$ among only $p = 4k + 1$ prime numbers the angle $\theta = \tan^{-1}(b/a)$ with $p = (a + bi)(a - bi)$ is evenly distributed in $\theta \in [0, 2\pi]$.

- the density of primes is $\#(\text{ digits })^{-1}$

These Tauberian theorems are very difficult and don't quantify in a nice way. However, they do give a window into regularization, which I like. If we prove using Complex Analysis have we reguarlized in some way?

I can only remember a few of these issues at a time, but PNT starts off as a discussion of $\zeta(1)$ but we have used $\zeta(-1)$, possibly $\zeta(2)$ and $\zeta(0)$ in this process. And I can see modern generalizations of that.

References

(1) Serge Lang **Algebraic Number Theory** (Graduate Texts in Mathematics #110) Springer, 1986.

For me Farey Fractions make total sense, whenever we study the convergence of Fourier series. Here's a function:

$$f(x) = \lfloor 0.5 - x \rceil = \sum_{n=0}^{\infty} \frac{1}{n} \sin(2\pi x) \text{ where } \lfloor a \rfloor = a - \{a\}$$
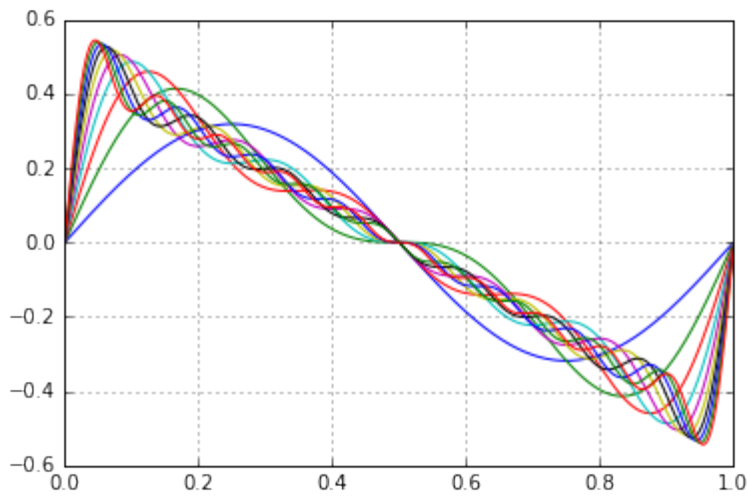
We can subtract off the main part and it looks like a jumprope. The waves cut the axis in the same pattern as the Farey fractions, almost.
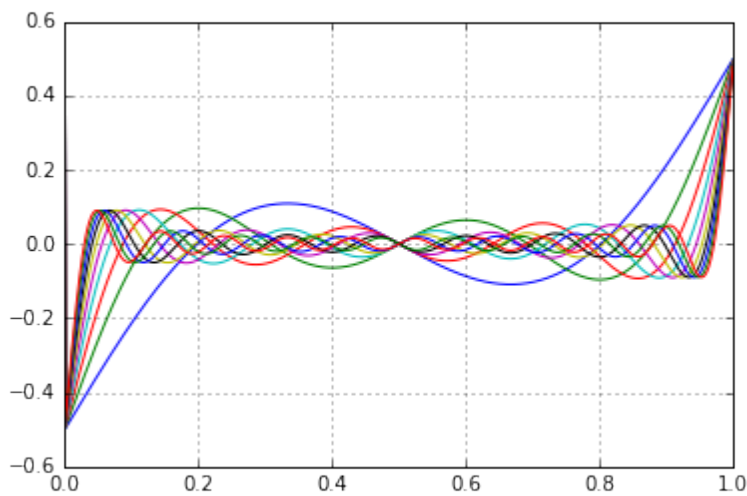
We could use high frequences. There is an artifct near the jump at $x \in \mathbb{Z}$ (any whole number).