Prime Number Theorem

John D Mangual

The Wiener-Ikehera Tauberian Theorem implies the Prime Number Theorem, as can be shown in various places.

None of those discussions really explain to me:

- connection to divergent series
- how we can proceed on our own

Norber Weiner's original argument is very easy to follow: Möbius inversion is what gives the series to start:

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

and then set $x = e^{-\xi}$ and $\xi \to 0$ (or $x \to 1$).

The two limiting behaviors are rather diffrent

$$\frac{x^n}{1-x^n} = \begin{cases} 1/n\epsilon & \text{as } x \to 1 \\ \epsilon^n & \text{as } x \to 0 \end{cases}$$

and the cuttoff point is when $(1-\epsilon)^n$ is getting small (near $n = 1/\epsilon$)

$$\sum_{n=1}^{\infty} \Lambda(n) \frac{\xi e^{-n\xi}}{1-e^{-n\xi}} \approx \sum_{n=1}^{1/\xi} \frac{\Lambda(n)}{n}$$

Maybe if we take the derivative of both sides:

$$\sum_{n=1}^{\infty} \Lambda(n) \, \frac{d}{d(n\xi)} \left[\frac{n\xi}{1-e^{n\xi}} \right] \approx \sum_{n=1}^{1/\xi} \Lambda(n)$$

And clearly the two sides are approximate so we are done.

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

The left side can be shown to be 1 through "elementary" arguments.

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

References

(1) David Vernon Widder **The Laplace Transform** Princeton University Press, 1948.

(2) Norbert Wiener **Tauberian Theorems** Annals of Mathematics Vol. 33, No. 1 , pp. 1-100

(3) G. H. Hardy **Divergent Series** Oxford University Press 1973

The argument in the previous section is wrong.[1]

It is very difficult to express in a vivid way - with images - why it is incorrect.

And in most situations it doesn't really matter. As for motivation, I can only speak for myself.

- why is Prime Number Theorem in a book on Laplace Transforms?
- why is Prime Number Theorem in a book in Divergent Series?

These lead me to neo-classical approaches - they will feel pretty modern to you![2]

One possibility is: **there is nothing new under the sun**. Everything is thinly dressed-up versions of the same problems since antiquity.

[1] I got really good at writing glib and suggestive proofs to hand in to graders. Who themselves are not always sure so they give you a check.

$$\sum_{n=1}^{1/\xi} a_n\, g(n\epsilon) \approx \sum_{n=1}^{1/\xi} a_n g(0)$$

Glib arguments can be acceptable, since we don't always have the time / resources to check all the cases.

[2] I tried to read the most modern papers first: there is Tao and Gowers and Green. However, that conversation presupposes knowledge I don't have and is written in language that I really don't like. They are pretty dreadful to read as are most papers in Analytic Number Theory as well as the people who write them!y The "beauty of the primes" is just a marketing term.

And that's pretty much how we feel. So let's try take some examples from `hep-th` and `math-nt`.

Hopefully, also finish a real proof of PNT.[3] What could possibly go wrong with approximations like this?

$$\sum_{n=1}^{\infty} \Lambda(n) \frac{\xi e^{-n\xi}}{1 - e^{-n\xi}} \approx \sum_{n=1}^{1/\xi} \frac{\Lambda(n)}{n}$$

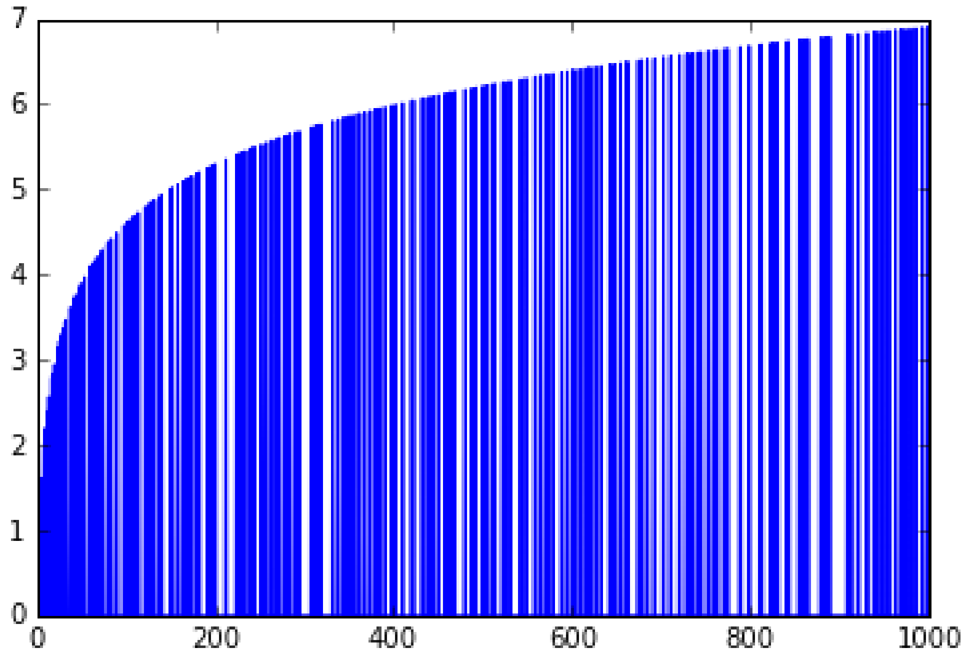
Maybe if we take the derivative of both sides:

$$\sum_{n=1}^{\infty} \Lambda(n) \, \frac{d}{d(n\xi)} \left[\frac{n\xi}{1 - e^{n\xi}} \right] \approx \sum_{n=1}^{1/\xi} \Lambda(n)$$

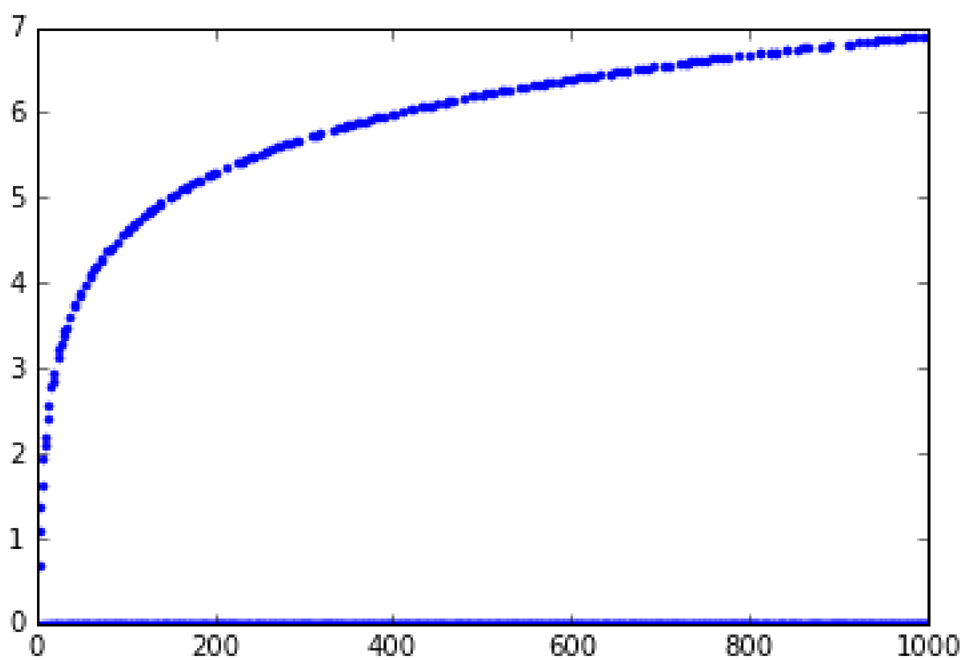
These should be satisfactory in any Physics or Engineering textbook.

[3]The terms "combinatorial" or "geometric" or "algebraic" or "elementary" have all been misleading and so I have often chosen to start from scratch.

By default plotting the van Mandolt function uses **linear interpolation** - disastrous.

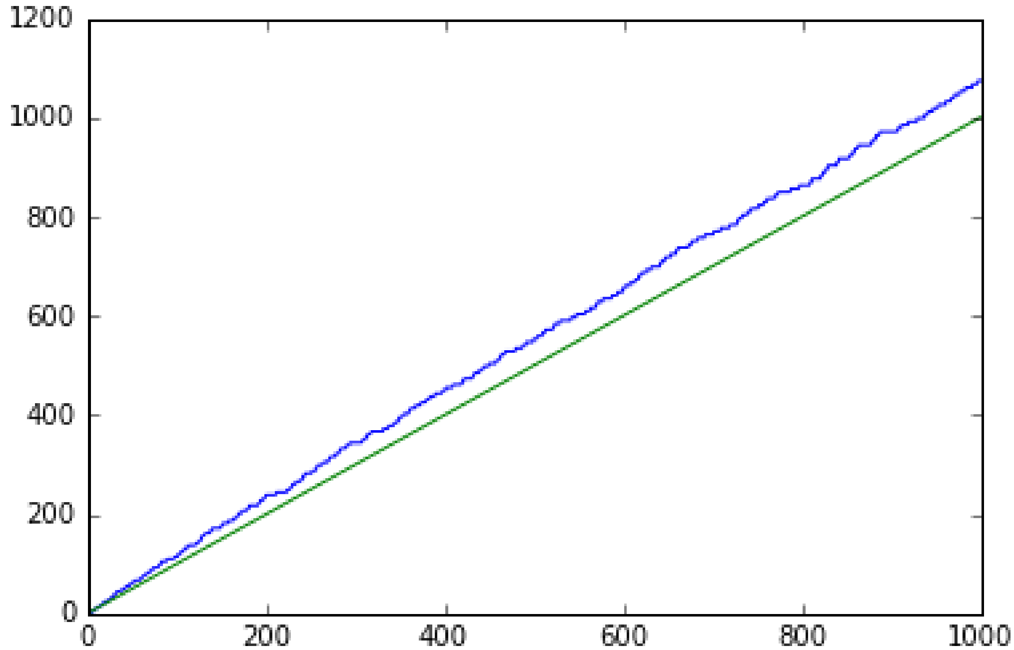


If we use "dots" we get that $\Lambda(n)$ has two parts $\Lambda(n) = 0$ when $n \neq p^k$ a prime power, and also $\Lambda(n) = \log p$ when $n = p^k$. Mysterious[4]

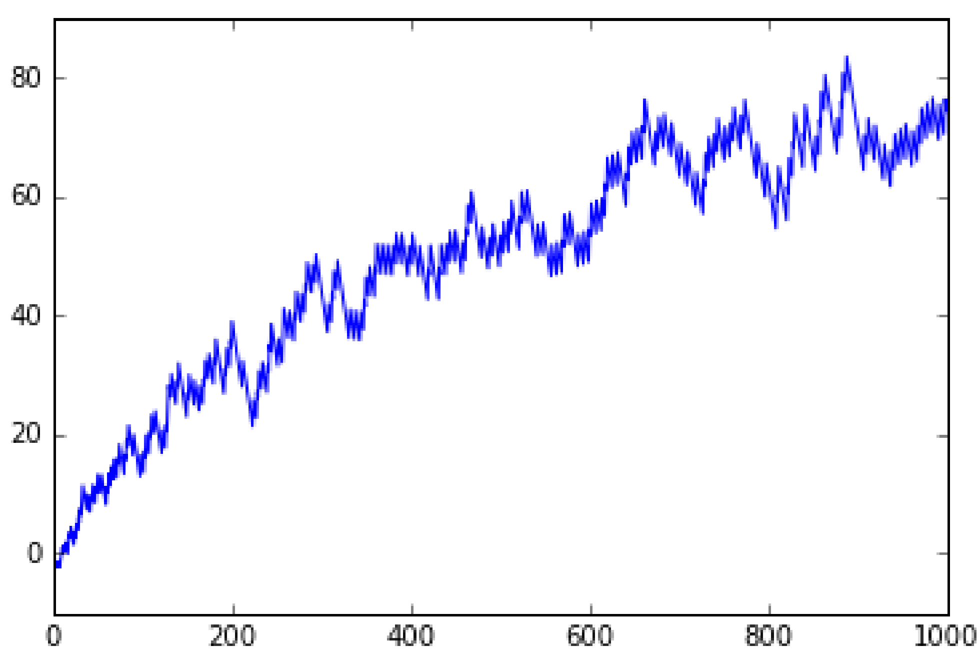


[4]It is related to $\zeta'(s)/\zeta(s) = \sum \frac{\Lambda(n)}{n^s}$.

The cumulative sum[5] $\sum_{n \leq x} \Lambda(n) \approx x$ we get almost a straight line. For comparison $y = x$.



The blue curve was consistently above the line in our range. This bias in an artifact of our choice primes $1 < p < 10^3$



[5]basically adding all the numbers. Please observe that $e^{\sum_{n \leq x} \Lambda(n)} = \mathrm{lcm}(1, 2, ;n) \approx e^x$. This is another way of stating PNT. The Least Common Multiple and the Permutation Group should play a more central role!

These pictures what the statement of PNT could mean:

$$\sum_{n\leq x} \Lambda(n) \approx x + o(x)$$

Sometimes when reading a complicated passage I'd get cynical:[6]

- why prime numbers?
- where do we consider averages over primes?
- why does it get so difficult?

The tools and arguments get increasingly sophicistated, but many things are complicted:

- a chair
- an iPhone
- an airplane
- rice

All these things have an intrinsic complexity.

Why did it take Norbert Wiener to solve this? B/c there is signal processing and random processes all over the place.[7]

[6] My latest idea is that Prime Number Theorem summarizes experiences with prime numbers up to a certain level. If you're not a cryptographer, it is possible that prime number theory or the zeta function plays an implicit level.

[7] Primes are not random, but signal processing (music) is full of signals and noise which are hard to separate. And many of those basic tools are taught to *their* undergraduates.

References

(1) David Vernon Widder **The Laplace Transform** Princeton University Press, 1948.

(2) Norbert Wiener **Tauberian Theorems** Annals of Mathematics Vol. 33, No. 1 , pp. 1-100

(3) G. H. Hardy **Divergent Series** Oxford University Press 1973

The starting point varies from author to author.

#1 (Hardy) If you have a sequence $s_n = O(1)$ then:

$$\left[(1-r) \sum s_n r^n \to s \right] \longrightarrow \left[\frac{1}{x} \sum_{n \leq x} s_n \to s \right]$$

Perhaps if I am taking one arrow $\to$ to another arrow $\to$ maybe it should be denoted $\Rightarrow$ as in **category theory**.

The problem could be if $s_n = \Lambda(n)$ it is not $O(1)$ – the van Mangolt function (or simply $\log p$) is not bounded by any constant.

Nothing should stop us from writing into argument a false implication such as

$$\left[(1-r) \sum_{n=1}^{\infty} \Lambda(n)\, r^n \to x] \right] \longrightarrow \left[\frac{1}{x} \sum_{n \leq x} \Lambda(n) \to x \right]$$

as long as we are judicious.[8] Especially since the Physics, Engineering and Econ literatures are full of these (if we are lucky).

[8] "x" here means "up to experimental error" such as $x + o(x)$. Just like cooking recipes on the internet.

#2 (Widder) Textbook starts very like a machine.[9]

$$f(x) = \int_0^\infty e^{-st}\, d\alpha(t)$$

This is a Stieltjes integral. Then two asymptotics are related:

$$\left[\alpha(t) \sim \frac{At^k}{\Gamma(k+1)}\right] \Rightarrow \left[f(s) \sim \frac{A}{s^k}\right]$$

As stated I wouldn't challenge them because they seem so plausible. In the case of PNT

$$f(x) = \sum_{n=1}^{\infty} \frac{(\Lambda(n)-1)e^{-nx}}{1-e^{-nx}} = \int_0^\infty \left[\frac{te^{-xt}}{1-e^{-xt}}\right]\, dh(t)$$

setting up an integration by parts. This is not a Laplace transform but it can be rearranged into one:

$$\begin{aligned} f(x) &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \{\Lambda(n)-1\} \quad e^{-mn\,x} \\ &= \sum_{n=1}^{\infty}\left[\sum_{d|n}\big(\Lambda(d)-1\big)\right]e^{-nx} \end{aligned}$$

The coefficient $c_n \asymp n$ (I often write $c_n \propto n$).

[9]The same author wrote a very nice textbook called *Advanced Calculus* with many lively discussions.

#3 (Norbert Wiener)

The necessary and sufficient condition for the set of all translations $f(x+t)$ to be **closed** is the that the real zeros of it's Fourier transform

$$\hat{f}(t) = \lim_{A\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-A}^{A} f(x) e^{-ixt}\, dx$$

should form a set of zero measure.[10]

The **step function** cerainly has this property [11]

$$\mathbf{1}(x) = \begin{cases} 1 & \text{if } x < 0 \\ 0 & \text{if } x \geq 0 \end{cases}$$

Wiener proceeds to do a lot of trigonometry to show that:

$$\begin{aligned} g(x) \; &=_1 \; \frac{d}{dx}\left[\frac{x\, e^{-x}}{e^{-x} - 1}\right] \\ &=_2 \; \frac{d}{dx}\left(\frac{\sin x}{x}\right)^2 \end{aligned}$$

The second one behaves like a "wave packet" the first one behaves rather nicely.

[10] Sets of zero measure can be quite interesting! They may in fact be the best part. Why are we talking about translations anyway? I bet it you shift around any function along the real number line $\mathbb{R}$ you can get everything else!

[11] I think – with a small amount of care about the value of $\mathbf{1}(0)$.

This rather stalwart opening brings back some bad memories for me. Back when I took a very beginning PDE[12] course, I was asked to show that:

$$L^2([0, 2\pi]) \text{ is spanned by } \{\sin n\, x\}_{n \in \mathbb{Z}}$$

I said this space was complete. It's not. You cannot get $\cos x$ this way. There went my A+.

References

(1) David Vernon Widder **The Laplace Transform** Princeton University Press, 1948.

(2) Norbert Wiener **Tauberian Theorems** Annals of Mathematics Vol. 33, No. 1 , pp. 1-100

(3) G. H. Hardy **Divergent Series** Oxford University Press 1973

[12] Partial Differential Equations

These proofs are completely arbitrary! One after the other, each more ridiculous than the last. And that is all there is.

Could there be more unifying recent ideas coming from other areas ?

- thermodynamics, entropy
- the geodesic flow, hyperbolic geometry
- enumerative combinatorics, generalizing $\binom{2n}{n}$
- ergodic theory of dynamical systems
- newer develoments in the theory of primes
- modular forms

The answer is certainly yes in at least one of these cases, even if it's just a rearrangement of an old proof.

Connections to Other Divergent Series

I think a less ambitious divergent series is to show:

$$1 + 2 + 3 + 4 + \cdots = -\frac{1}{12}$$

If we subtract a series from four times itself:

$$\begin{aligned} c &= 1 + 2 + 3 + 4 + \dots \\ 4c &= 4 + 8 + 12 + \dots \\ -3c &= 1 - 2 + 3 - 4 + 5 - 6 = \frac{1}{(1+1)^2} = \frac{1}{4} \end{aligned}$$

and we conclude $c = -\frac{1}{12}$. Then there is **zeta regularization** and **cutoff regularization**.

Can these senseless elementary-school tricks, become serious mathematics? GH Hardy addresses both of these methods in his book on Divergent Series (he knew Ramanujan personally).

The way I learned Euler Maclaurin summation is through a kind of symbol calculus:

$$\begin{aligned} f(0) + f(1) + \cdots + f(n) &= \left(1 + e^{\Delta} + e^{2\Delta} + e^{3\Delta} + \dots\right) f \\ &= \left[\frac{e^{(n+1)\Delta} - 1}{e^{\Delta} - 1}\right] f(0) \end{aligned}$$

Yet what is the formula for the integral.

$$\frac{1}{N} \sum f(i) = \sum_{k=0}^{nN} e^{k\Delta/N} f(0) = \left[\frac{e^{\Delta n(1+1/N)} - 1}{e^{\Delta/N} - 1}\right] f(0)$$

so I don't remember anymore. The generating function for the Bernoulli numbers:

$$g(t) = \frac{t}{e^t - 1} = \sum_{m=0}^{\infty} B_m \frac{t^m}{m!}$$

is also the weight for Lambert summation (rather $g'(t)$).

$$\frac{1}{\xi} \approx \prod_{n=1}^{\infty} p^{\frac{\xi e^{-p\xi}}{1-e^{-p\xi}}} \approx \prod_{n=1}^{1/\xi} p^{1/n}$$

and set $x = 1/t$ with $t \to 0$. Exercise! Um... right.

Whenever we add evenly spaced numbers (or by rescaling sum over the integers) are are evaluating a "weighted" zeta function. And then we talk about "the" sum when you add a function $f(\cdot)$ over all of $\mathbb{Z}$.

This generalizes in a way that nobody ever works out: to number fields, lattices in higher dimensions, Todd classes, other quadratures, L-functions, etc.

From here on the strategy is:

- pick a problem
- realized someone worked on it and solved already

and step 1. In order to not be too critical, these are

- new technique (last 5-15 yrs)
- which have classical counterparts (or not).
- known to experts
- not widely known

and nobody is 100% sure which bin to put any given problem.

Perhaps better not to think too much about this aspect.[13]

[13]The best one can do, I think is review. I think collectively I have

Ex #1 Let's sum all the imginary numbers:

$$\frac{1}{4}\sum_{(a,b)\neq 0}(a^2+b^2) = \left[\prod_{p\equiv 1(4)}\left(1-p\right)\right]^{-2}\left[\prod_{p\equiv 3(4)}\left(1-p^2\right)\right]^{-1}$$

In fact we can sum all the positive integers:

$$\sum_{a\geq 0} a = -\tfrac{1}{12} = \prod_{p}(1-p)^{-1}$$

and other amazing artifacts await!