A QUICK DISTRIBUTIONAL WAY TO THE PRIME NUMBER THEOREM

JASSON VINDAS AND RICARDO ESTRADA

Abstract. We use distribution theory (generalized functions) to show the prime number theorem. No tauberian results are employed.

1. Introduction

We give distributional proofs of the celebrated Prime Number Theorem (in short PNT). The word distributional refers to Schwartz distributions. So, we show that

$$\pi(x) \sim \frac{x}{\log x} \ , \quad x \to \infty \ , \tag{1.1}$$

where

$$\pi(x) = \sum_{p \text{ prime}, \ p<x} 1 \ . \tag{1.2}$$

We provide two related proofs. It is remarkable that both proofs are direct and do not use any tauberian argument. Our arguments are based on Chebyshev's elementary estimate [3, p.14]

$$\pi(x) = O\left(x/\log x\right) \ , \quad x \to \infty \ , \tag{1.3}$$

and additional properties of the Riemann zeta function on the line $\Re e\, z = 1$.

We would like to point out the recent articles [6, 7], where distributional methods have been also applied to prime number theory.

2. Special functions and distributions

In this section we briefly explain some special functions and distributions related to prime numbers.

Throughout this article, the letter p stands only for a prime number. We denote by Λ the von Mangoldt function defined on the natural numbers as

$$\Lambda(n) = \begin{cases} 0 \ , & \text{if } n = 1 \ , \\ \log p \ , & \text{if } n = p^m \text{ with } p \text{ prime and } m > 0 \ , \\ 0 \ , & \text{otherwise} \ . \end{cases} \tag{2.1}$$

2000 *Mathematics Subject Classification.* Primary 46F10, 11N05. Secondary 46F20, 41A60, 46F99.

Key words and phrases. Prime number theorem, asymptotic behavior of generalized functions.

As usually done, we denote by ψ the *Chebyshev function*

$$\psi(x) = \sum_{p^m < x} \log p = \sum_{n<x} \Lambda(n) \ . \tag{2.2}$$

It follows easily from Chebyshev's classical estimate (1.3) that for some $M > 0$

$$\psi(x) < Mx \ . \tag{2.3}$$

It is very well known since the time of Chebyshev that the PNT is equivalent to the statement

$$\psi(x) \sim x \ . \tag{2.4}$$

Our approach to the PNT will be to show (2.4).

We denote by $\mathcal{D}(\mathbb{R})$ and $\mathcal{S}(\mathbb{R})$ the Schwartz spaces of the test functions consisting of smooth compactly supported functions and smooth rapidly decreasing functions, respectively, with their usual topologies. Their dual spaces, the spaces of distributions and tempered distributions, are denoted by $\mathcal{D}'(\mathbb{R})$ and $\mathcal{S}'(\mathbb{R})$, respectively. We refer to [10] for the very well known properties of these spaces.

Our proof of the PNT is based on finding the distributional asymptotic behavior of $\psi'(x)$ (the derivative is understood in the distributional sense, of course); observe that

$$\psi'(x) = \sum_{n=1}^{\infty} \Lambda(n)\, \delta(x-n) \ , \tag{2.5}$$

where δ is the well known Dirac delta distribution whose action on test functions is given by $\langle \delta(x), \phi(x) \rangle = \phi(0)$. For this goal, we shall study the asymptotic properties of the distribution

$$v(x) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n} \delta(x - \log n) \ ; \tag{2.6}$$

clearly $v \in \mathcal{S}'(\mathbb{R})$.

Consider the Riemann zeta function

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z} \ , \ \Re e\ z > 1 \ . \tag{2.7}$$

Let us first take the Fourier-Laplace transform [1, 12] of v, that is, for $\Im m\ z > 0$

$$\left\langle v(t), e^{izt} \right\rangle = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{1-iz}} = -\frac{\zeta'(1-iz)}{\zeta(1-iz)} \ , \tag{2.8}$$

a formula that Riemann obtained by logarithmic differentiation of the Euler product for the zeta function $\zeta(z) = \prod_p 1/(1-p^{-z})$.

Taking the boundary values on the real axis, in the distributional sense, we obtain the Fourier transform of v ,

$$\hat{v}(x) = -\frac{\zeta'(1-ix)}{\zeta(1-ix)} \ . \tag{2.9}$$

Notice that we are not saying that the right hand side on the last relation is a function but rather that it is a tempered distribution. We shall always interpret (2.9) as equality in the space $\mathcal{S}'(\mathbb{R})$, meaning that for each $\phi \in \mathcal{S}(\mathbb{R})$

$$\langle \hat{v}(x), \phi(x) \rangle = -\lim_{y\to 0^+} \int_{-\infty}^{\infty} \phi(x) \frac{\zeta'(1-ix+y)}{\zeta(1-ix+y)} \, \mathrm{d}x \ . \tag{2.10}$$

It is implicit in (2.9) that the Fourier transform we are using is

$$\hat{\phi}(x) = \int_{-\infty}^{\infty} e^{ixt} \phi(t) \, \mathrm{d}t \ , \quad \text{for } \phi \in \mathcal{S}(\mathbb{R}) \ ,$$

and it is defined by duality on $\mathcal{S}'(\mathbb{R})$. We emphasize that (2.10) follows automatically from the fact that v is a tempered distribution, but, alternatively, it may be directly deduced from Chebyshev's estimate (2.3) through integration by parts (when applying $-\zeta'(1-ix+y)/\zeta(1-ix+y)$ to a test function).

We discuss some properties of the distribution $\hat{v}$. From the well known properties of ζ, we conclude that on $\mathbb{R}\setminus\{0\}$ $\hat{v}$ is a locally integrable function. Indeed,

$$\zeta(z) - \frac{1}{z-1} \tag{2.11}$$

admits an analytic continuation to a neighborhood of $\Re e\, z = 1$, as one easily proves by applying the Euler-Maclaurin formula; in addition, $\zeta(1+ix)$, $x \neq 0$, is free of zeros [3, 4, 5]. It follows then that

$$\hat{v}(x) - \frac{i}{(x+i0)} \in L^1_{\mathrm{loc}}(\mathbb{R}) \ , \tag{2.12}$$

where here we use the notation $1/(x+i0)$ for the distributional boundary value of the analytic function z^{-1}, $\Im m\, z > 0$.

The property (2.12) together with Chebyshev's estimate (2.3) will be the key ingredients for the proof of the PNT given in Section 5.

The proof to be given in Section 4 makes use of additional information of the Riemann zeta function on the line $\Re e\ z = 1$; we shall take for granted that $\hat{v}$ has at most polynomial growth as $|x| \to \infty$. In fact, more than this is true: $\hat{v}(x) = O(\log^{\beta}|x|)$ as $|x| \to \infty$, for some $\beta > 0$. The reader can find the proof of this fact in [3, Chap.2] (see also [8]). Summarizing, we have that

$$\hat{v}(x) - \frac{i}{(x+i0)} \in L^1_{\mathrm{loc}}(\mathbb{R}) \ \text{ and has tempered growth} \ . \tag{2.13}$$

3. Notation from generalized asymptotics

The purpose of this section is to clarify the notation to be used in the following two sections. It comes from the theory of asymptotic behaviors of generalized functions [2, 9, 12]. Besides the notation, we do not make use of any deep result from this theory.

Let $f \in \mathcal{D}'(\mathbb{R})$, a relation of the form

$$\lim_{h\to\infty} f(x+h) = \beta\ , \quad \text{in } \mathcal{D}'(\mathbb{R})\ , \tag{3.1}$$

means that the limit is taken in the weak topology of $\mathcal{D}'(\mathbb{R})$, that is, for each test function from $\mathcal{D}(\mathbb{R})$ the following limit holds,

$$\lim_{h\to\infty} \langle f(x+h), \phi(x)\rangle = \beta \int_{-\infty}^{\infty} \phi(x)\mathrm{d}x\ . \tag{3.2}$$

The meaning of the expression $\lim_{h\to\infty} f(x+h) = \beta$ in $\mathcal{S}'(\mathbb{R})$ is clear. Relation (3.1) is an example of the so-called *S-asymptotics* of generalized functions (also called asymptotics by translation), we refer the reader to [9] and [12] for further properties of this concept.

On the other hand, we may attempt to study the asymptotic behavior of a distribution by looking at the behavior at large scale of the dilates $f(\lambda x)$ as $\lambda \to \infty$. In this case, we encounter the concept of *quasiasymptotic behavior* of distributions [2, 9, 11, 12]. We will study in connection to the PNT a particular case of this type of behavior, namely, a limit of the form

$$\lim_{\lambda\to\infty} f(\lambda x) = \beta\ , \quad \text{in } \mathcal{D}'(\mathbb{R})\ , \tag{3.3}$$

Needless to say that (3.3) should be always interpreted in the weak topology of $\mathcal{D}'(\mathbb{R})$. We may also talk about (3.3) in other spaces of distributions with a clear meaning.

4. Proof of the PNT based on (2.3) and (2.13)

We begin with the distribution v given by (2.6).

Our first step is to show that

$$\lim_{h\to\infty} v(x+h) = 1, \quad \text{in } \mathcal{S}'(\mathbb{R})\ . \tag{4.1}$$

We denote the Heaviside function by $H(x)$, i.e., the characteristic function of $(0,\infty)$. Let $\phi \in \mathcal{S}(\mathbb{R})$. Consider $\phi_1 \in \mathcal{S}(\mathbb{R})$ such that $\phi = \hat{\phi_1}$; then as

$h \to \infty$

$$\begin{aligned}\langle v(x+h), \phi(x)\rangle &= \int_{-h}^{\infty} \phi(x)\mathrm{d}x + \langle v(x+h) - H(x+h)\ , \phi(x)\rangle \\ &= \int_{-h}^{\infty} \phi(x)\mathrm{d}x + \left\langle \hat{v}(x) - \frac{i}{(x+i0)}\ , e^{-ihx}\phi_1(x)\right\rangle \\ &= \int_{-h}^{\infty} \phi(x)\mathrm{d}x + \int_{-\infty}^{\infty} e^{-ihx}\phi_1(x)\left(\hat{v}(x) - \frac{i}{(x+i0)}\right)\mathrm{d}x \\ &= \int_{-\infty}^{\infty} \phi(x)\mathrm{d}x + o(1), \quad h \to \infty\ ,\end{aligned}$$

where the last step follows in view of (2.13) and the Riemann-Lebesgue lemma. This shows (4.1).

In a second essential step we pass from distributions on $(-\infty, \infty)$ to distributions on $(0, \infty)$. Specifically, we will show that

$$\lim_{\lambda\to\infty} \psi'(\lambda x) = \lim_{\lambda\to\infty} \sum_{n=1}^{\infty} \Lambda(n)\delta(\lambda x - n) = 1\ , \quad \text{in } \mathcal{D}'(0,\infty)\ . \tag{4.2}$$

By (4.1) we have $e^{t+h}v(t+h) \sim e^{t+h}$, as $h \to \infty$, in the weak topology of $\mathcal{D}'(\mathbb{R})$. This readily implies that

$$\sum_{n=1}^{\infty} \Lambda(n)\phi(\log n - h) \sim e^h \int_0^{\infty} \phi(\log x)dx\ , \quad h \to \infty\ , \tag{4.3}$$

for each test function $\phi \in \mathcal{D}(\mathbb{R})$. Now any test function $\phi_1 \in \mathcal{D}(0,\infty)$ can be written as $\phi_1(x) = \phi(\log x)$ with $\phi \in \mathcal{D}(\mathbb{R})$. Setting $\lambda = e^h$ in (4.3) we obtain (4.2).

Here comes the final step in our argument, we evaluate (4.2) at suitable test functions to deduce that $\psi(x) \sim x$. Let $\varepsilon > 0$ be an arbitrary number; find ϕ_1 and $\phi_2 \in \mathcal{D}(0,\infty)$ with the following properties: $0 \le \phi_i \le 1$, $\operatorname{supp}\phi_1 \subseteq (0,1]$, $\phi_1(x) = 1$ on $[\varepsilon, 1-\varepsilon]$, $\operatorname{supp}\phi_2 \subseteq (0, 1+\varepsilon]$, and finally, $\phi_2(x) = 1$ on $[\varepsilon, 1]$. Applying the distributional relation (4.2) to ϕ_2 and using (2.3), we obtain that

$$\begin{aligned}\limsup_{\lambda\to\infty} \frac{1}{\lambda}\sum_{n<\lambda}\Lambda(n) &\le \limsup_{\lambda\to\infty}\left(\frac{1}{\lambda}\sum_{n<\varepsilon\lambda}\Lambda(n) + \frac{1}{\lambda}\sum_{n=1}^{\infty}\Lambda(n)\phi_2\left(\frac{n}{\lambda}\right)\right) \\ &\le M\varepsilon + \lim_{\lambda\to\infty}\frac{1}{\lambda}\sum_{n=1}^{\infty}\Lambda(n)\phi_2\left(\frac{n}{\lambda}\right) \\ &= M\varepsilon + \int_0^{\infty}\phi_2(x)dx \le 1 + \varepsilon(M+1)\ .\end{aligned}$$

Using now ϕ_1, we easily obtain that

$$1 - 2\varepsilon \le \liminf_{\lambda\to\infty}\frac{1}{\lambda}\sum_{n<\lambda}\Lambda(n)\ .$$

Since ε was arbitrary, we conclude that $\psi(\lambda) \sim \lambda$ and the PNT follows immediately.

5. Proof of the PNT based on (2.3) and (2.12)

In this section we present a variant of the proof discussed in Section 4. In fact, we show how to avoid the use of the growth properties of $\zeta(z)$ on $\Re e\ z = 1$. We begin by observing that it is enough to establish (4.1). Indeed, once (4.1) is obtained, one can proceed identically as in the Section 4 and derive the PNT merely from Chebyshev's estimate. Therefore, we shall derive (4.1) from (2.3) and (2.12). In view of (2.12) and the argument from the previous section involving the Riemann-Lebesgue lemma, we can still deduce that for each test function ϕ with $\operatorname{supp} \hat{\phi}$ compact

$$\lim_{h\to\infty} \langle v(x+h), \phi(x)\rangle = \int_{-\infty}^{\infty} \phi(x)\mathrm{d}x \ . \tag{5.1}$$

The set of test functions having this property is dense in $\mathcal{S}(\mathbb{R})$. Then, if one were able to show that $v(x+h) = O(1)$ in $\mathcal{S}'(\mathbb{R})$, that is, that the set of translates of v is a weakly bounded set, then (4.1) would follow from the Banach-Steinhaus theorem and the convergence over a dense subset of $\mathcal{S}(\mathbb{R})$. We now show this last property. Let $g(x) = e^{-x}\psi(e^x)$. Because of (2.3), we have that $g(x+h) = O(1)$ in the weak topology of $\mathcal{S}'(\mathbb{R})$. Consequently, we also have that $g'(x+h) = O(1)$ in $\mathcal{S}'(\mathbb{R})$. Hence, $v(x+h) = g'(x+h)+g(x+h) = O(1)$ in $\mathcal{S}'(\mathbb{R})$, as required. The boundedness of $v(x+h)$ together with (5.1) imply the PNT.

References

[1] H. Bremermann, *Distributions, Complex Variables and Fourier Transforms*, Addison-Wesley, Reading, Massachusetts, 1965.

[2] R. Estrada, R. P. Kanwal, *A distributional approach to asymptotics. Theory and applications,* Second edition, Birkhäuser, Boston, 2002.

[3] A. E. Ingham, *The distribution of prime numbers,* Cambridge University Press, Cambridge, 1990.

[4] A. Ivić, *The Riemann zeta-function. Theory and applications,* Dover Publications, New York, 2003.

[5] J. Korevaar, *Tauberian theory. A century of developments*, Springer-Verlag, Berlin, 2004.

[6] J. Korevaar, *Distributional Wiener-Ikehara theorem and twin primes,* Indag. Math. (N.S.) **16** (2005), 37–49.

[7] J. Korevaar, *Prime pairs and the zeta function,* J. Approx. Theory **158** (2009), 69–96.

[8] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen,* Vol.1, Second edition, Chelsea Publishing Co., New York, 1953.

[9] S. Pilipović, B. Stanković, A. Takači, *Asymptotic behaviour and Stieltjes transformation of distributions*, Teubner-Texte zur Mathematik, Leipzig, 1990.

[10] L. Schwartz, *Théorie des Distributions,* Hermann, Paris, 1966.

[11] J. Vindas, *Structural Theorems for Quasiasymptotics of Distributions at Infinity,* Pub. Inst. Math. (Beograd) **84**(98) (2008), 159–174.

[12] V. S. Vladimirov, Yu. N. Drozzinov, B. I. Zavialov, *Tauberian theorems for generalized functions,* Kluwer Academic Publishers, Dordrecht, 1988.

Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803-4918, U.S.A.

E-mail address: jvindas@math.lsu.edu, restrada@math.lsu.edu