

## ON LITTLEWOOD’S PROOF OF THE PRIME NUMBER THEOREM

ALEKSANDER SIMONIČ 

(Received 20 June 2019; accepted 11 July 2019; first published online 15 August 2019)

### Abstract

In this note we examine Littlewood’s proof of the prime number theorem. We show that this can be extended to provide an equivalence between the prime number theorem and the nonvanishing of Riemann’s zeta-function on the one-line. Our approach goes through the theory of almost periodic functions and is self-contained.

2010 *Mathematics subject classification*: primary 11N05; secondary 42A75.

*Keywords and phrases*: prime number theorem, almost periodic function.

### 1. Introduction

The prime number theorem (PNT) is considered one of the most important theorems in mathematics. It states that  $\pi(x) \sim x/\log x$ , where  $\pi(x)$  counts prime numbers less than or equal to  $x$ , and is equivalent to  $\psi(x) \sim x$ , where  $\psi(x) = \sum_{p^r \leq x} \log p$  is the Chebyshev function (see [5, Section I.4]).

Apart from the Selberg and Erdős elementary approach to the PNT, the essential part in all known proofs consists of knowing the zero-free region of the Riemann zeta-function  $\zeta(s)$ . Denote by  $\rho = \beta + i\gamma$  the nontrivial zeros of  $\zeta(s)$ , where  $0 \leq \beta \leq 1$  and  $\gamma \in \mathbb{R}$ . A not-very-well-known proof of the PNT was given by Littlewood in [7], where he demonstrated that it is equivalent to

$$\lim_{x \downarrow 0} \sum_{\rho} x^{1-\rho} \Gamma(\rho) = 0. \quad (1.1)$$

Here  $\Gamma(z)$  is the gamma function. The series in (1.1) is uniformly convergent and this allows us to apply the limit inside the sum. (The uniform convergence follows simply from the fact that  $\Gamma(x + iy) \ll e^{-|y|}$ , valid uniformly for  $x \leq 2$  and  $|y| \geq 1$ , and  $N(T + 1) - N(T) = O(\log T)$ , where  $N(T)$  is the number of those zeros  $\rho$  with  $0 < \gamma \leq T$  (see [9, Theorem 9.2]).)

Observe that  $\beta \neq 1$  implies (1.1) and consequently the PNT. The converse statement is well known.

**THEOREM 1.1.** *The prime number theorem implies that  $\beta \neq 1$ .*

The common proof of Theorem 1.1 goes through the formula

$$-\frac{\zeta'}{\zeta}(s) - \frac{s}{s-1} = \int_1^\infty \frac{\psi(x) - x}{x^{s+1}} dx,$$

which is valid for  $\Re\{s\} > 1$  (see [5, page 37]). While this idea is independent of any approach to the PNT, it is also very tempting to use (1.1) in the opposite direction and thus 'complete' Littlewood's proof.

The main purpose of this note is twofold: to sketch Littlewood's proof in the hope of making it more popular and to provide a proof of Theorem 1.1 using identity (1.1). Our approach goes through the theory of almost periodic functions (Definition 3.1). We should mention that this idea is not new. In [6, pages 261–262], it was used to establish equivalence of the PNT and  $\beta \neq 1$  for functions in the Selberg class. This proof is considerably more difficult than ours and uses properties of almost periodic functions, such as the uniqueness theorem, which are not so trivial as they might appear. We show that it is possible to provide all necessary details in a concise way while avoiding the concept of the Fourier series of an almost periodic function, thus making this exposition accessible also to nonspecialists.

The outline of our proof is as follows. Assume the existence of zeros  $\rho = 1 + iy$  and denote the ordinates of such zeros in the upper half-plane by  $0 < \gamma_1 \leq \gamma_2 \leq \dots$ . By symmetry,  $\gamma_{-j} := -\gamma_j$ ,  $j > 0$ , are the other zeros. Let  $\mathcal{S} \subseteq \mathbb{Z} \setminus \{0\}$  be the set of all indices of  $\gamma$  (could be finite or infinite) and let  $\{a_n\}_{n \in \mathcal{S}}$  be a sequence of complex numbers such that the corresponding series converges absolutely and, if  $\gamma_i = \gamma_j$ , then  $a_i = a_j$ . By the identity principle, the set  $\{\gamma_n\}_{n \in \mathcal{S}}$  does not have accumulation points. Define the function

$$F(x; a_n) := \sum_{n \in \mathcal{S}} a_n e^{i\gamma_n x}.$$

Then the PNT is equivalent to

$$\lim_{x \rightarrow \infty} F(x; \Gamma(1 + i\gamma_n)) = 0. \quad (1.2)$$

In Section 3 we will show that  $F(x; a_n)$  is an almost periodic function and in Section 4 that  $\lim_{x \rightarrow \infty} F(x; a_n) = 0$  implies that  $F \equiv 0$  (see Lemma 4.1) and furthermore that this implies that  $a_n = 0$  for every  $n \in \mathcal{S}$ . In view of (1.2), this would be a contradiction and the proof of Theorem 1.1 will be complete.

## 2. Littlewood's proof of the PNT

Most proofs of the PNT consist of two main parts called 'Tauberian' and 'analytic'. A Tauberian theorem deals with the question of whether it is possible to obtain a (partial) converse of an Abelian theorem. As an example, consider the statement that, for  $a_n \geq 0$ ,

$$\lim_{x \downarrow 0} x \sum_{n=1}^{\infty} a_n e^{-nx} = 1 \quad \iff \quad \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{n=1}^m a_n = 1$$

holds. The left implication is similar to the classical theorem due to Abel on the continuity of a power series at a point on the boundary of its convergence disc. The main strength lies with the right implication and this can be obtained from the celebrated Hardy–Littlewood theorem from 1914. In 1930 Karamata found a much simpler two-page proof which uses the Weierstrass approximation theorem as the only advanced tool (see [8, pages 226–229]). If we take  $a_n = \Lambda(n)$ , where  $\Lambda$  is the von Mangoldt function, then the PNT is equivalent to

$$\lim_{x \downarrow 0} x \sum_{n=1}^{\infty} \Lambda(n) e^{-nx} = 1. \tag{2.1}$$

We would like to mention here that one year earlier Ramanujan in his third letter to Hardy studied the function

$$\phi(x) := \phi_1(x) - \phi_2(x) := \sum_{n=1}^{\infty} \Lambda(n) e^{-nx} - \log 2 \sum_{n=1}^{\infty} 2^n e^{-2^n x}.$$

He claimed without proof that  $\lim_{x \downarrow 0} x \phi_2(x) = 1$  and  $\lim_{x \downarrow 0} x \phi(x) = 0$ , which would consequently imply (2.1). But, in [3, page 39], Hardy used a clever argument to show that the first limit is not only wrong but it cannot even exist, thus implying that the second limit is also wrong. On the same page Hardy wrote: ‘*I should like to say that “rigour apart, he found the Hardy–Littlewood proof”, but I cannot*’. The interested reader may find in this treatise some other examples of incorrect claims in analytic number theory by Ramanujan.

The  $\Lambda$ -function satisfies the important relation  $-\zeta'(s)/\zeta(s) = \sum_{n=1}^{\infty} \Lambda(n)n^{-s}$  for  $\Re\{s\} > 1$ . The analytical part of the proof begins with this equation together with the Mellin integral (see [9, Section 2.15]) for  $e^{-x}$  to get

$$x \sum_{n=1}^{\infty} \Lambda(n) e^{-nx} = -\frac{1}{2\pi i} \lim_{t \rightarrow \infty} \int_{2-it}^{2+it} \Gamma(s) \frac{\zeta'}{\zeta}(s) x^{1-s} ds,$$

where  $x > 0$ . Now the idea is to take a contour integral along the rectangle  $\mathcal{R}_T$  with vertices  $-1/2 \pm iT$  and  $2 \pm iT$ , where the horizontal segments do not pass through the zeros of the zeta-function. (Littlewood took  $-1$  instead of  $-1/2$  in the contour, but this is not a good choice because the gamma function has a pole at  $-1$ . We choose  $-1/2$  but any number in the interval  $(-1, 0)$  would suffice.) By the calculus of residues,

$$-\frac{1}{2\pi i} \int_{\mathcal{R}_T} \Gamma(s) \frac{\zeta'}{\zeta}(s) x^{1-s} ds = - \sum_{|\Im\{\rho\}| < T} x^{1-\rho} \Gamma(\rho) + 1 - x \frac{\zeta'}{\zeta}(0).$$

The first part clearly comes from the zeros of  $\zeta(s)$  within the contour, while the second and third parts come from simple poles of  $\zeta'(s)/\zeta(s)$  and  $\Gamma(s)$  at 1 and 0, respectively. We need a result which asserts that there is an increasing and unbounded sequence  $\{T_j\}_{j=1}^{\infty}$  of positive numbers such that  $\zeta'(\sigma \pm iT_j)/\zeta(\sigma \pm iT_j) = O(T_j)$ , uniformly for  $-1/2 \leq \sigma \leq 2$ , and this could be deduced from an approximate formula for  $\zeta'(s)/\zeta(s)$

(see [9, Theorem 9.6(A)], and [5, Theorem 26] for a complete proof of a much better result). We also need an estimate  $\zeta'(-1/2 + it)/\zeta(-1/2 + it) \ll |t| + 1$ , which follows from the logarithmic derivative version of the functional equation for  $\zeta(s)$ . With these estimates in hand, together with the observation that  $\Gamma(-1/2 + iy) \ll e^{-|y|}$  is still valid for  $y \in \mathbb{R}$ , we can show that

$$-\frac{1}{2\pi i} \int_{\mathcal{R}_{T_j}} \Gamma(s) \frac{\zeta'}{\zeta}(s) x^{1-s} ds = -\frac{1}{2\pi i} \int_{2-iT_j}^{2+iT_j} \Gamma(s) \frac{\zeta'}{\zeta}(s) x^{1-s} ds + O\left(x\sqrt{x} + e^{-T_j} \left(\frac{x^2\sqrt{x}(1-\log x) - 1}{2x\log x} T_j - x\sqrt{x}\right)\right).$$

Taking  $j \rightarrow \infty$  in the above formula,

$$x \sum_{n=1}^{\infty} \Lambda(n) e^{-nx} = - \sum_{\rho} x^{1-\rho} \Gamma(\rho) + 1 - x \frac{\zeta'}{\zeta}(0) + O(x\sqrt{x}),$$

which finally gives

$$\lim_{x \downarrow 0} \left( x \sum_{n=1}^{\infty} \Lambda(n) e^{-nx} + \sum_{\rho} x^{1-\rho} \Gamma(\rho) \right) = 1. \tag{2.2}$$

Equation (2.2) was already announced in Hardy and Littlewood's influential paper [4], but used for different purposes. It is clear now that the combination of (2.1) and (2.2) produces (1.1).

### 3. Almost periodic functions

The theory of almost periodic functions was initiated by H. Bohr in 1925 and turned out to be very useful in the study of differential equations and Fourier analysis (see [1, 2]). The space of such functions has remarkable properties. It includes the space of periodic functions and is a vector space. Moreover, the limit of every uniformly convergent sequence of almost periodic functions is also almost periodic (see Theorem 3.2 below). The following definition is due to Bochner.

**DEFINITION 3.1.** Let  $f(x)$  be a continuous function on  $\mathbb{R}$  with complex values. We say that  $f(x)$  is (uniformly) almost periodic if for every sequence  $\{x_n\}_{n=1}^{\infty} \subset \mathbb{R}$  there exists a subsequence  $\{x_{1,n}\}_{n=1}^{\infty}$  such that  $f(x + x_{1,n})$  converges uniformly.

We denote the space of all such functions by  $\mathfrak{A}$ . While the next theorem is fundamental, the common proof is somewhat longer than our proof (see [1, pages 1–5]). The reason is that we use Bochner's definition instead of Bohr's original one.

**THEOREM 3.2.** *The following three properties hold:*

- (1) *continuous periodic functions are almost periodic;*
- (2) *if  $f_1, f_2 \in \mathfrak{A}$  and  $a_1, a_2 \in \mathbb{C}$ , then  $a_1 f_1 + a_2 f_2 \in \mathfrak{A}$ ;*
- (3) *if  $\{f_n\}_{n=1}^{\infty} \subset \mathfrak{A}$  and  $f_n \rightarrow f$  uniformly, then  $f \in \mathfrak{A}$ .*

**PROOF.** We will provide a proof of the first property only. The second property is trivial and a proof of the third property is straightforward if one exploits Cantor’s diagonal process.

Let  $f$  be a continuous periodic function with period  $\omega > 0$ . Take an arbitrary sequence  $\{x_n\}_{n=1}^\infty \subset \mathbb{R}$  and define a sequence  $\{x'_n\}_{n=1}^\infty$  by

$$x'_n := \min\{x_n - k\omega \geq 0 : k \in \mathbb{Z}\}.$$

Then  $f(x + x_n) = f(x + x'_n)$  and  $\{x'_n\}_{n=1}^\infty \subset [0, \omega)$ . It follows that there exists a subsequence  $\{x_{1,n}\}_{n=1}^\infty$  of  $\{x_n\}_{n=1}^\infty$  such that  $\{x'_{1,n}\}_{n=1}^\infty$  converges to some  $x_0 \in [0, \omega)$ . Because  $f$  is continuous on a compact set  $[0, 2\omega]$ , it is also uniformly continuous there. This means that for every  $\varepsilon > 0$  there exists  $\delta > 0$  such that  $|x'_{1,n} - x_0| < \delta$  implies that  $|f(x + x'_{1,n}) - f(x + x_0)| < \varepsilon$  for every  $x \in [0, \omega]$ . But then for every  $\varepsilon > 0$  there exists  $N$  such that  $|f(x + x_{1,n}) - f(x + x_0)| < \varepsilon$  for every  $n > N$  and  $x \in \mathbb{R}$ . Therefore,  $f(x + x_{1,n}) \rightarrow f(x + x_0)$  uniformly and  $f$  is thus almost periodic.  $\square$

#### 4. Proof of Theorem 1.1

Like a periodic function, an almost periodic function has the property that the existence of its limit at infinity characterises it completely.

**LEMMA 4.1.** *Let  $f(x)$  be an almost periodic function. Then  $f(x) \equiv C \in \mathbb{C}$  if and only if  $\lim_{x \rightarrow \infty} f(x) = C$ .*

**PROOF.** If  $f(x) \equiv C$ , then  $\lim_{x \rightarrow \infty} f(x) = C$ . On the contrary, assume that  $f(x)$  is not a constant function. Then there exist  $x_0$  and  $y_0$  such that  $f(x_0) \neq f(y_0)$ . Define  $m := \frac{1}{3}|f(x_0) - f(y_0)| > 0$ . Because  $f \in \mathfrak{A}$ , there exists a strictly increasing sequence  $\{h_n\}_{n=1}^\infty$  with  $h_n \rightarrow \infty$  such that  $f(x + h_n)$  converges uniformly. This means that there exists  $N$  such that  $|f(x + h_N) - f(x + h_n)| < m$  for all  $n > N$  and  $x \in \mathbb{R}$ . For  $n > N$ , define  $x_n := x_0 + h_n - h_N$  and  $y_n := y_0 + h_n - h_N$ , so that  $x_n \rightarrow \infty$  and  $y_n \rightarrow \infty$ . Then

$$\begin{aligned} 3m &= |f(x_0) - f(y_0)| \\ &\leq |f(x_0) - f(x_n)| + |f(x_n) - f(y_n)| + |f(y_n) - f(y_0)| \\ &< 2m + |f(x_n) - f(y_n)|; \end{aligned}$$

therefore,  $|f(x_n) - f(y_n)| > m$ . This means that  $\lim_{x \rightarrow \infty} f(x)$  could not exist. So, if  $\lim_{x \rightarrow \infty} f(x) = C$ , then  $f$  must be a constant function and it is clear that this constant is  $C$ .  $\square$

**PROOF OF THEOREM 1.1.** All three properties in Theorem 3.2 guarantee that  $F(x; a_n)$  is an almost periodic function. If  $\lim_{x \rightarrow \infty} F(x; a_n) = 0$ , then Lemma 4.1 implies that  $F(x; a_n) = 0$  for all  $x \in \mathbb{R}$ . Take some  $n_0 \in \mathcal{S}$  and let  $\mathcal{S}_0$  be a subset of  $\mathcal{S}$  such that  $n \in \mathcal{S}_0$  if and only if  $\gamma_n = \gamma_{n_0}$ . Then

$$F(x; a_n)e^{-i\gamma_{n_0}x} = |\mathcal{S}_0|a_{n_0} + \sum_{n \in \mathcal{S} \setminus \mathcal{S}_0} a_n e^{i(\gamma_n - \gamma_{n_0})x}.$$

Because we can change the order of summation and integration, the above equation implies that

$$I(X) := \frac{1}{X} \int_0^X F(x; a_n) e^{-i\gamma_{n_0} x} dx = |\mathcal{S}_0| a_{n_0} + \frac{1}{iX} R(X),$$

where

$$R(X) := \sum_{n \in \mathcal{S} \setminus \mathcal{S}_0} \frac{a_n}{\gamma_n - \gamma_{n_0}} (e^{i(\gamma_n - \gamma_{n_0})X} - 1).$$

We also observe that  $\lim_{X \rightarrow \infty} I(X) = |\mathcal{S}_0| a_{n_0}$  since

$$|R(X)| \leq \frac{2}{d} \sum_{n \in \mathcal{S}} |a_n| < \infty,$$

where  $d := \min_{n \in \mathcal{S} \setminus \mathcal{S}_0} \{|\gamma_n - \gamma_{n_0}|\} > 0$ . But then  $F(x; a_n) \equiv 0$  implies that  $a_n = 0$  for all  $n \in \mathcal{S}$ . Consequently, the limit (1.2) does not hold and the Riemann zeta-function does not have zeros with real parts equal to one.  $\square$

Finally, we point out that it is possible to construct the theory of almost periodic functions through trigonometric polynomials  $T(x) = \sum_{k=1}^n c_k e^{i\lambda_k x}$ , where  $c_k$  are complex numbers and  $\lambda_k$  are real numbers (see [2]). Then we could say that  $f(x)$  is an almost periodic function when for every  $\varepsilon > 0$  there exists a trigonometric polynomial  $T_\varepsilon(x)$  such that  $|f(x) - T_\varepsilon(x)| < \varepsilon$  for every  $x \in \mathbb{R}$ . By this definition, our function  $F(x; a_n)$  is of course almost periodic, but the author could not find a similar argument as in the proof of Lemma 4.1 by using this definition. It is equivalent to Bochner's, but the proof is somehow longer than the proof of Theorem 3.2.

### Acknowledgement

The author is grateful to Tim Trudgian for helpful and constructive remarks on the manuscript.

### References

- [1] A. S. Besicovitch, *Almost Periodic Functions* (Dover, New York, 1955).
- [2] C. Corduneanu, *Almost Periodic Functions* (Chelsea, New York, 1989).
- [3] G. H. Hardy, *Ramanujan. Twelve Lectures on Subjects Suggested by his Life and Work* (Cambridge University Press, Cambridge, UK–Macmillan, New York, 1940).
- [4] G. H. Hardy and J. E. Littlewood, 'Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes', *Acta Math.* **41**(1) (1916), 119–196.
- [5] A. E. Ingham, *The Distribution of Prime Numbers*, Cambridge Mathematical Library (Cambridge University Press, Cambridge, 1990).
- [6] J. Kaczorowski and A. Perelli, 'On the prime number theorem for the Selberg class', *Arch. Math. (Basel)* **80**(3) (2003), 255–263.
- [7] J. E. Littlewood, 'The quickest proof of the prime number theorem', *Acta Arith.* **18** (1971), 83–86.
- [8] E. C. Titchmarsh, *The Theory of Functions* (Oxford University Press, Oxford, 1958).

- [9] E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function*, 2nd edn (The Clarendon Press–Oxford University Press, New York, 1986).

ALEKSANDER SIMONIČ, School of Science,  
The University of New South Wales (Canberra), ACT,  
Australia  
e-mail: [aleks.simonic@gmail.com](mailto:aleks.simonic@gmail.com)