# ON THE DISTRIBUTION FUNCTION OF THE REMAINDER TERM OF THE PRIME NUMBER THEOREM.*

**Otto Toeplitz in Memoriam**

By Aurel Wintner.

**Introduction.** The classical result of Littlewood [1] on the distribution of primes, when expressed in terms of the standard function

$$\psi(x) = \sum_{p^m \leqq x} \log p = \sum_{n \leqq x} \Lambda(n), \text{ where } -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = \frac{\zeta'(s)}{\zeta(s)}, \qquad (\sigma > 1),$$

states that, under Riemann's hypothesis,

$$\text{(I)} \qquad \frac{\psi(x) - x}{x^{\frac{1}{2}}} = \Omega_{\pm}(\log_3 x) \qquad \text{as } x \to \infty;$$

while, according to von Koch [2]

$$\text{(II)} \qquad \frac{\psi(x) - x}{x^{\frac{1}{2}}} = O(\log^2 x) \qquad \text{as } x \to \infty.$$

While (II) means that the remainder term, $\psi(x) - x$, of Hadamard's prime number theorem, $\psi(x) \sim x$, is *at most*

$$\pm \text{ const. } x^{\frac{1}{2}}(\log x)^2$$

for *every* sufficiently large $x$ and for a certain positive constant, (I) states that this remainder term is *at least*

$$\pm \text{ Const. } x^{\frac{1}{2}} \log\log\log x$$

for *certain* sufficiently large $x$ and for another positive constant (where both signs actually occur in $\pm$ Const. for certain large values of $x$).

Correspondingly, neither (I) nor (II) implies any information concerning the asymptotic behavior of the remainder term. For instance, the upper estimate, (II) leaves open the question whether or not the even powers of the ratio on the left of (II), when measured on the proper scale of the prime

* Received November 8, 1940.

[1] Cf., e. g., A. E. Ingham, "The distribution of prime numbers," *Cambridge Tracts in Mathematics and Physics*, no. 30 (1932), pp. 86-107.

[2] Cf. *ibid.*, pp. 83-84.

number theory and then transformed into space averages (momenta), are such as to lead to *finite* asymptotic averages. Similarly, the lower estimate, (I), does not preclude the following possibility: There exists on the real axis a *bounded* interval with the property that the values of the ratio on the left of (I) are within this bounded interval for " almost all " values of the independent time variable just mentioned.

The object of the present paper is to answer these questions by proving the asymptotic counterparts of the results (I), (II) of Littlewood and of von Koch, respectively. Needless to say, neither of the theorems to be proved is implied by (I) and (II) together, although neither (I) nor (II) is implied by the two theorems to be proved. In fact, estimates from above and from below are necessarily of such a nature as to take into account possible accidental irregularities on every set clustering at infinity, even if these sets are of relative measure zero. On the other hand, such sets are irrelevant from the point of view of statistical *averages*.

The situation can be illustrated by considering that part of Littlewood's discovery which was the most surprising; namely, the fact that the remainder term changes its sign infinitely often and in such a way that the deviation of $x^{-\frac{1}{2}}\psi(x)$ from $x^{\frac{1}{2}}$ can be arbitrarily large in *either* direction. The corresponding result of the present paper goes further in this qualitative respect, since it is to the effect that the curves $y = x^{-\frac{1}{2}}\psi(x)$ and $y = x^{\frac{1}{2}}$ cross each other in arbitrarily distant $\pm\, y$-regions with non-vanishing relative frequencies, and not only infinitely often (it being understood that the relative amount of time spent in a $y$-region is measured on the scale of the proper independent time variable of the prime number theorem.) But this does not imply any explicit $\Omega$-estimate.

Riemann's hypothesis will, of course, be assumed throughout; otherwise the problems under consideration, like the problems considered by Littlewood, either do not arise at all or are known to be of a trivial nature.

It may be mentioned that the main difficulties of the problem arise from the fact that nothing is known about the Diophantine structure of the non-trivial zeros. In particular, if it were true that these zeros (or, rather, their imaginary parts) are linearly independent in the rational field, then, as pointed out previously,[3] much more than that what will now be proved could be inferred directly from the theory of infinite convolutions.

The proofs will depend on a fact, which I proved a few years ago,[3] to the effect that the *trigonometric series* (in $t = \log x$) occurring in the explicit

[3] A. Wintner, " On the asymptotic distribution of the remainder term of the prime-number theorem," *American Journal of Mathematics*, vol. 57 (1935), pp. 534-538.

formula of Riemann and von Mangoldt is actually the *Fourier expansion* of the function which it represents (the Fourier character being meant in the sense $(B^2)$ of the theory of almost periodic functions). Due to this fact, the asymptotic distribution theory of almost periodic functions of a real variable becomes applicable and leads, without too much effort, to the asymptotic counterpart of (II), indicated above. The asymptotic counterpart of (I) lies deeper, since it depends not only on the asymptotic distribution theory of almost periodic functions of a real variable but also on a lemma concerning asymptotic averages connected with analytic, uniformly almost periodic functions of a complex variable.

Although the lemma in question regulates the increase of mean values which are represented by asymptotic averages and not, as usual, by integrals, it is only a manifestation of the maximum principle, and so one would expect it to be standard; however, it does not seem to be available in the literature. In order to avoid an interruption of the following considerations, this lemma will be established first (in a form slightly more general than necessary for the problem at hand.)

**1.** If $f(t)$ is defined for $0 \leqq t < \infty$ and is integrable $(L)$ on any bounded $t$-interval, put

$$M_t\{f(t)\} = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(t)\,dt,$$

whenever this limit exists.

LEMMA. *If a function $g = g(\sigma, t)$, defined on a strip*

$$\alpha \leqq \sigma \leqq \beta, \qquad -\infty < t < \infty,$$

*is non-negative, subharmonic and such as to satisfy, uniformly in $\sigma$, an estimate of the form*

$$g(\sigma, t) = O(|t|^C) \qquad \textit{as } t \to \pm\infty,$$

*where $C$ is a sufficiently large constant, and if $g(\sigma, t)$ is, for every fixed $\sigma$ contained in the interval $\alpha \leqq \sigma \leqq \beta$, a uniformly almost periodic function of $t$, then the least upper bound of $M_t\{g(\sigma, t)\}$ for $\alpha \leqq \sigma \leqq \beta$ is either $M_t\{g(\alpha, t)\}$ or $M_t\{g(\beta, t)\}$.*

It will be seen from the proof that, instead of the assumption of uniform almost periodicity on every fixed line $\sigma$, it is sufficient to require that, if $\sigma$ is fixed, the ratio

$$\mu \equiv \mu(T, \tau; \sigma) = \frac{1}{2T} \int_{\tau-T}^{\tau+T} g(\sigma, t)\,dt$$

should tend, as $T \to \infty$, to $M_t\{g(\sigma, t)\}$ uniformly for all $\tau$, where $-\infty < \tau < \infty$. Although this requirement, which is necessarily satisfied in the case of uniform almost periodicity, does not necessitate even a generalized type of almost periodicity, it represents a condition which need not be satisfied if the function, instead of being uniformly almost periodic, is almost periodic $(B^q)$ for some (or, for that matter, for every) value of $q$.

If the statement of the Lemma is true on every sufficiently short subinterval of $\alpha \leqq \sigma \leqq \beta$, it is true on the whole interval $\alpha \leqq \sigma \leqq \beta$. Hence, it can be assumed (after a translation) that $-\frac{1}{2}\pi < \alpha$ and $\beta < \frac{1}{2}\pi$. Then, if $\sigma$ is any point of the interval $\alpha \leqq \sigma \leqq \beta$, the subharmonic character of $g$ and the uniform $O(|t|^C)$-estimate (and, as a matter of fact, even a weaker estimate of the Phragmén-Lindelöf type) are known[4] to imply that there exists for every $\epsilon > 0$ and for every $T > 0$ a point

$$(\sigma_0, t_0) \equiv (\sigma_0(\epsilon, T; \sigma),\ t_0(\epsilon, T; \sigma))$$

which is situated on the boundary of the strip $\alpha \leqq \sigma \leqq \beta$, $-\infty < t < \infty$, and satisfies the inequality

$$\int_{-T}^{T} g(\sigma, t)\,dt - \epsilon\cdot(e^T - e^{-T})\cos\sigma \leqq \int_{-T}^{T} g(\sigma_0, t + t_0)\,dt.$$

Since $\sigma_0 \equiv \sigma_0(\epsilon, T; \sigma)$ is either $\alpha$ or $\beta$, the value of the last integral cannot exceed

$$\operatorname{Max}\Big(\int_{-T}^{T} g(\alpha, t + t_0)\,dt,\ \int_{-T}^{T} g(\beta, t + t_0)\,dt\Big),$$

where Max $(A, B)$ denotes $A$ or $B$ according as $A \geqq B$ or $A \leqq B$. Hence, division by $2T$ shows that, in terms of the above abbreviation, $\mu$, for the resulting ratios,

$$\mu(T, 0; \sigma) - \tfrac{1}{2}\epsilon T^{-1}(e^T - e^{-T})\cos\sigma \leqq \operatorname{Max}\,(\mu(T, t_0; \alpha), \mu(T, t_0; \beta)),$$

where $t_0 = t_0(\epsilon, T; \sigma)$ is a certain real function.

For a given $\eta > 0$, choose a positive $R = R(\eta)$ so large that

$$|\mu(T, t_0; \alpha) - M_t\{g(\alpha, t)\}| < \eta \text{ and } |\mu(T, t_0; \beta) - M_t\{g(\beta, t)\}| < \eta$$

[4] Cf. G. H. Hardy, A. E. Ingham and G. Pólya, "Notes on moduli and mean values," *Proceedings of the London Mathematical Society*, ser. 2, vol. 27 (1928), pp. 401-409 (more particularly pp. 407-408), where further references are given. The inequality referred to above is, save for the notation, identical with the case $c = 0$ of the first inequality on p. 408, *loc. cit.*

whenever $T > R(\eta)$. The existence of such a function $R$ of $\eta$ alone is assured by the assumption that the limit relation

$$\mu(T, \tau; \sigma) \rightarrow M_t\{g(\sigma, t)\}, \qquad T \rightarrow \infty,$$

holds uniformly for all values of $\tau$ (where $\tau = t_0(\epsilon, T; \sigma) \equiv t_0$), if the line $\sigma$ (where $\sigma = \alpha, \beta$) is fixed. Thus

$$\mu(T, 0; \sigma) - \tfrac{1}{2}\epsilon T^{-1}(e^T - e^{-T}) \cos \sigma \leqq \eta + \operatorname{Max}(M_t\{g(\alpha, t)\}, M_t\{g(\beta, t)\})$$

whenever $T > R(\eta)$. Hence, on letting $\epsilon \rightarrow 0$ and $\eta \rightarrow 0$ (in this order), one sees from

$$\mu(T, 0; \sigma) \rightarrow M_t\{g(\sigma, t)\}, \qquad T \rightarrow \infty,$$

that

$$M_t\{g(\sigma, t)\} \leqq \operatorname{Max}(M_t\{g(\alpha, t)\}, M_t\{g(\beta, t)\}).$$

Since this inequality holds for any point $\sigma$ of the interval $\alpha \leqq \sigma \leqq \beta$, the proof of the Lemma is complete.

**2.** By a distribution function $\phi = \phi(\alpha)$, $-\infty < \alpha < \infty$, is meant any monotone function for which $\phi(-\infty) = 0$ and $\phi(\infty) = 1$. The spectrum of a $\phi$ is defined as the (necessarily closed) set of those points $\alpha$ for which $\phi(\alpha') \neq \phi(\alpha'')$ whenever $\alpha' < \alpha < \alpha''$. Let $[\phi]$ denote the maximum ($\leqq \infty$) of the absolute values of the numbers $\alpha$ contained in the spectrum of $\phi$; so that $[\phi] < \infty$ if and only if the spectrum is a bounded set. For $k = 0, 1, 2, \cdots$, let $M_k(\phi)$ be an abbreviation for the $k$-th momentum,

$$M_k(\phi) = \int_{-\infty}^{\infty} \alpha^k d\phi(\alpha),$$

provided that

$$\int_{-\infty}^{0} \alpha^k d\phi(\alpha) = -\infty \text{ and } \int_{0}^{\infty} \alpha^k d\phi(\alpha) = \infty$$

do not hold simultaneously. Thus $M_{2k}(\phi)$ is always defined but can be $\infty$. It is easily verified that, as $k \rightarrow \infty$, the $2k$-th (non-negative) root of $M_{2k}(\phi)$ always tends to the limit $[\phi]$ (which can be $\infty$); it being understood that $[\phi] = \infty$ if, but not only if, $M_{2k}(\phi) = \infty$ from a certain $k$ onward.

For a given real-valued (measurable) function $f = f(t)$, $-\infty < t < \infty$, let $f_{(\alpha)} = f_{(\alpha)}(t)$, $-\infty < t < \infty$, denote the characteristic function of the $t$-set defined by $f(t) < \alpha$, where $\alpha$ is any fixed real number; in other words, put $f_{(\alpha)}(t) = 1$ or $f_{(\alpha)}(t) = 0$ according as $f(t)$ is or is not less than $\alpha$. If there exists a distribution function $\phi$ with the property that, unless $\alpha$ is a discontinuity point of $\phi$, the $t$-average $M_t\{f_{(\alpha)}(t)\}$, as defined at the beginning of § 1, exists and is equal to $\phi(\alpha)$, then $f = f(t)$ is said to have an asymptotic distribution function, $\phi$. It is very important (possibly, though not probably, also for the result of the present paper), that $M_t\{f_{(\alpha)}(t)\}$ is not required to exist if $\alpha$ is one of the (at most enumerable) points at which the monotone function $\phi$ has a saltus.

It is known [5] that if $f$ is uniformly almost periodic, it possesses an asymptotic distribution function, $\phi$, and that the spectrum of this $\phi$ is identical with the (bounded) interval representing the closure of the values attained by $f(t)$ for $-\infty < t < \infty$; so that, in particular, $[\phi]$ is the least upper bound of $|f(t)|$.

If $f(t)$; $f_1(t), f_2(t), \cdots, -\infty < t < \infty$, are measurable functions, $f_n$ is said to tend, as $n \rightarrow \infty$, to $f$ in relative measure if, for every fixed $\epsilon > 0$,

$$\limsup_{T\rightarrow\infty} \frac{1}{2T} \int_{-T}^{T} \{1 - |f(t) - f_n(t)|_{(\epsilon)}\} dt \rightarrow 0 \text{ as } n \rightarrow \infty,$$

where the integrand, { }, is the characteristic function of the $t$-set defined by $|f(t) - f_n(t)| \geqq \epsilon$ (the function $|f(t) - f_n(t)|_{(\epsilon)}$ representing the characteristic function of the $t$-set defined by $|f(t) - f_n(t)| < \epsilon$). It is known [6] that if $f_n$ tends to $f$ in relative measure, and if every $f_n$ has an asymptotic distribution function, $\phi_n$, then $f$ has an asymptotic distribution function, $\phi$, and that $\phi_n \rightarrow \phi$. It is understood that $\phi_n \rightarrow \phi$ means that $\phi_n(\alpha) \rightarrow \phi(\alpha)$ holds at every $\alpha$ which is not a discontinuity point of $\phi$.

It follows, in particular, that every relatively almost periodic function $f$ has an asymptotic distribution function; $f$ being defined to be relatively almost periodic if it is measurable and such that there exists a sequence of uniformly almost periodic functions $f_1 f_2, \cdots$ which tend to $f$ in relative measure. This implies that $f$ has an asymptotic distribution function whenever it is almost periodic ($B^q$) for some $q$; in fact, convergence in the mean (of any index $q$) necessitates convergence in relative measure. On the other hand, straightforward examples show that $f$ can be relatively almost periodic without being almost periodic ($B$).

**3.** Let the function $h = h(t)$ be defined for $0 < t < \infty$ by

$$e^{\frac{1}{2}t}h(t) = \psi(e^t) - e^t, \qquad (e^t > 1);$$

[5] A. Wintner, *Spektraltheorie der unendlichen Matrizen*, Leipzig, 1929, pp. 267-272; cf. B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 48-88, more particularly p. 77.

[6] Cf. B. Jessen and A. Wintner, *loc. cit.*[5], pp. 75-76.

so that the lower and upper estimates mentioned at the beginning of the Introduction become

(I) $$h(t) = \Omega_{\pm}(\log\log t)$$

and

(II) $$h(t) = O(t^2),$$

respectively, while the prime number theorem appears in the form

$$h(t) = o(e^{\frac{1}{2}t}); \qquad (t \to \infty).$$

Under Riemann's hypothesis, let

$$\tfrac{1}{2} \pm i\gamma_1, \cdots, \tfrac{1}{2} \pm i\gamma_n, \cdots, \text{ where } 0 < \gamma_1 < \cdots \leqq \gamma_n \leqq \cdots,$$

denote the sequence of the complex zeros of the $\zeta$-function. The signs of equality are not excluded, a multiple zero, if any, being reckoned in accordance with its order.

Since

(1) $$\gamma_n \sim 2\pi n/\log n, \qquad (n \to \infty),$$

the trigonometric series

(2) $$\omega(t) = -\sum_{n=1}^{\infty}\left(\frac{e^{i\gamma_n t}}{\frac{1}{2} + i\gamma_n} + \frac{e^{-i\gamma_n t}}{\frac{1}{2} - i\gamma_n} - 2\frac{\sin\gamma_n t}{\gamma_n}\right)$$

has a uniform majorant for all $t$ and defines, therefore, a uniformly almost periodic function, $\omega(t)$. On the other hand, the explicit formula of Riemann and von Mangoldt states [7] that the trigonometric series

(3) $$-2\sum_{n=1}^{\infty}\gamma_n^{-1}\sin\gamma_n t$$

is convergent (for $0 < t < \infty$) and represents a function, say $f(t)$, for which the sum $f(t) + \omega(t)$ differs from the reduced remainder term,

$$\frac{\psi(x) - x}{x^{\frac{1}{2}}} \equiv \frac{\psi(e^t) - e^t}{e^{t/2}},$$

of the prime number theorem only in a Dirichlet phenomenon (at the discontinuity points) and in the trivial additive terms

$$\log 2\pi = \frac{\zeta'(0)}{\zeta(0)} \text{ and } \tfrac{1}{2}\log(1 - x^{-2}) = \sum_{n=1}^{\infty}\frac{1}{-2nx^{2n}}$$

(respectively introduced by the pole, $s = 0$, and the real zeros, $s = -2n$, of the $\zeta$-function). Since $\psi(x)$, where $x = e^t$, has at every prime power, $x = p^m$, the saltus $\log p$ and is otherwise continuous, it follows that

(4) $$h_0(t) = f(t) + \omega(t) - e^{-\frac{1}{2}t}\{\log 2\pi + \tfrac{1}{2}\log(1 - e^{-2t})\} \text{ for } 0 < t < \infty,$$

[7] Cf. A. E. Ingham, *loc. cit.*[1], pp. 76-78.

where

$$(5)\quad h_0(t)=\frac{h(t+0)+h(t-0)}{2},\ h(t)-h_0(t)=\begin{cases}0 \text{ unless } t=m\log p,\\ \tfrac{1}{2}p^{-\frac{1}{2}m}\log p \text{ if } t=m\log p;\end{cases}$$

so that

$$(6)\qquad h_0(t)-f(t)-\omega(t)\to 0 \text{ and } h(t)-h_0(t)\to 0 \text{ as } t\to\infty.$$

The function $f(t)$ has thus far been considered as defined by the odd trigonometric series (3) only for $0<t<\infty$. The number-theoretical meaning of this series for negative $t$ is well known [8] but will be immaterial in what follows.

**4.** The mere fact that the trigonometric series (3) represents the principal part of the reduced remainder term, $(\psi(x)-x)/x^{\frac{1}{2}}$, of the prime number theorem does not involve any information as to the behavior of the function $f(t)$, where $x=e^t$ (the situation is well illustrated by the immensity of the gap between the estimates (I) — (II), § 3). For instance, it is easy to see that a trigonometric series which is convergent for $-\infty<t<\infty$ can represent a function (even continuous) which has neither an asymptotic distribution function [9] nor Fourier averages, $M_t\{e^{i\lambda t}f(t)\}$.

It was suggested by an apparent parallelism between certain problems in celestial mechanics [10] on the one hand and the "wobbly" terms of the explicit formula of Riemann and von Mangoldt on the other hand, that, by leaving aside for a moment the aspect of the Abschätzungen and replacing it by a point-of-view in theoretical astronomy, it would perhaps be possible to obtain some new insight into the irregularities of the prime number distribution. In this direction, it was possible to prove [11] that, after a suitable extension of $f(t)$ and $h(t)$ for negative $t$ (cf. the beginning of § 5):

There exists for the function defined at the end of § 3 an anharmonic analysis, i. e., that the time averages $M_t\{e^{i\lambda t}f(t)\}$, where $-\infty<\lambda<\infty$, exist;

---

[8] Cf. A. E. Ingham, *loc. cit.*[1], pp. 80-81.

[9] H. Poincaré, *Oeuvres*, vol. 1 (1928), pp. 164-166; cf. A. Wintner, "Ueber die kleinen numerischen Divisoren in der Theorie der allgemeinen Störungen," *Mathematische Zeitschrift*, vol. 31 (1929), pp. 434-440.

[10] A. Wintner, "Sur l'analyse anharmonique des inégalités séculaires fournies par l'approximation de Lagrange," *Rendiconti Reale Accademia dei Lincei*, ser. 6, vol. 11 (1930), pp. 464-467, and "Ueber eine Anwendung der Theorie der fastperiodischen Funktionen auf das Levi-Civitasche Problem der mittleren Bewegung," *Annali di Matematica*, ser. 4, vol. 10 (1931-32), pp. 277-282; cf. also "Almost periodic functions and Hill's theory of lunar perigee," *American Journal of Mathematics*, vol. 59 (1937), pp. 795-802, and "On an ergodic analysis of the remainder term of mean motions," *Proceedings of the National Academy of Sciences*, vol. 26 (1940), pp. 126-129.

[11] A. Wintner, *loc. cit.*[3].

That the values of these Fourier averages are precisely those which one would expect on the basis of the formal series (3), i. e., that the prime numbers $p_1\, p_2, \cdots$ (by means of which (4) is representable, via (5), in finite terms) and the complex zeros $\frac{1}{2} \pm i\gamma_1, \frac{1}{2} \pm i\gamma_2, \cdots$ are connected by the following mysterious "dispersion formula":

$$(7) \qquad M_t\{e^{i\lambda t}\, h(t)\} = \begin{cases} 0 \text{ unless } \lambda = \pm \gamma_n, \\ (\frac{1}{2} + i\lambda)^{-1} \text{ if } \lambda = \pm \gamma_n : \end{cases}$$

Finally, that $f$ (and, therefore, $h$) is almost periodic ($B^2$), which ensures, in particular, the existence of an asymptotic distribution function.

These facts imply, but are by no means implied by, a result of H. Cramér[12] (although the proofs require only an adaptation of Cramér's proof); a result according to which the quadratic $t$-average exists and is equal to the square sum of the amplitudes. In fact, not even the existence of the averages $M_t\{e^{i\lambda t}\, h(t)\}$ follows from this result.

Incidentally, not even the existence of all averages $M_t\{g^2\} < \infty$ and $M_t\{e^{i\lambda t}\, g(t)\}$ together assures that $g(t)$ has an asymptotic distribution function. All that is clear is that if all these averages exist, $g$ is or is not almost periodic ($B^2$) according as the sum of all squared amplitudes, $| M_t\{e^{i\lambda t}\, g(t)\} |^2$, is equal to or less than $M_t\{g^2\}$; it being understood that the set of those real numbers $\lambda$ for which $M_t\{e^{i\lambda t}\, g(t)\}$ is distinct from 0 is at most enumerable by virtue of the assumption $M_t\{g^2\} < \infty$.

In this connection, it would be interesting to know whether or not the existence of all Fourier averages $M_t\{e^{i\lambda t}\, g(t)\}$ alone, or perhaps together with the assumption of a finite average for $| g |$ (but not for $g^2$), implies that the $\lambda$-set defined by $M_t\{e^{i\lambda t}\, g(t)\} \neq 0$ is at most enumerable.

**5.** As mentioned in the second paragraph of § 4, the function $f(t)$ is now thought of as defined for negative $t$ also. This is the more necessary as the notion of almost periodicity ($B^2$) is usually referred to the limit of the *symmetric* time range $-T < t < T$. In view of (3), it is natural to extend $f(t)$ to negative $t$ by $f(t) = -f(-t)$. The effect of this extension on the asymptotic distribution can be interpreted as follows:

Let $a(t)$ be a function which is defined only $0 \leqq t < \infty$ and which has, with reference to this half-line, the asymptotic distribution function $\chi(\alpha)$, $-\infty < \alpha < \infty$. Then, if $\alpha = \alpha'$, $\alpha = \alpha''$ ($> \alpha'$) is any pair of real numbers which are not discontinuity points of $\chi(\alpha)$, the difference $\chi(\alpha'') - \chi(\alpha')$ represents the asymptotic relative amount of time spent by the curve $a = a(t)$,

---

[12] Cf. A. E. Ingham, *loc. cit.*[1], p. 106, where further references are given.

$0 \leqq t < \infty$, in the strip $\alpha' < a < \alpha''$ of the $(t, a)$-plane. Hence, if the function $a(t)$ is extended to negative $t$ by the condition $a(-t) = -a(t)$, and if the asymptotic distribution is referred to the limit of the symmetric $t$-range $-T < t < T$, the result is the same as if one would consider $a(t)$ only for positive $t$ but restrict $\alpha'$, $\alpha''$ by $\alpha'' = -\alpha'$. The existence of an asymptotic distribution function now means that there exists a monotone function $\chi(\alpha)$, $-\infty < \alpha < \infty$, which is of total variation 1 and such that, if $\beta$ is any positive number for which neither $\alpha = \beta$ nor $\alpha = -\beta$ is a discontinuity point of $\chi(\alpha)$, the difference $\chi(\beta) - \chi(-\beta)$ represents the asymptotic relative amount of time spent by the curve $a = a(t)$, where $0 \leqq t < \infty$, in the strip $-\beta < a(t) < \beta$.

It is seen from § 4 that what are usually called, under Riemann's hypothesis, the irregularities of the prime number distribution are, as a matter of fact, no irregularities at all, except when measured on the logarithmico-exponential $(\Omega, O)$-scale; a scale which is, of course, incapable of expressing the hidden almost-periodicities of anharmonic analysis. It would even be possible that $f(t)$ is not only almost periodic $(B^2)$ but is equivalent $(B^2)$ to a uniformly almost periodic function. Actually, such a uniformly almost periodic function cannot exist. More than this is implied by the following theorem:

*Under Riemann's hypothesis, the asymptotic distribution function of the reduced remainder term,*

$$h(t) = (\psi(x) - x)/x^{\frac{1}{2}}, \qquad (x = e^t),$$

*of the prime number theorem is such as to possess*

(i) *a spectrum which is unbounded in either direction of the $\alpha$-axis* (cf. the first paragraph of § 2);

(ii) *momenta of arbitrarily high order which, in addition, do not increase more rapidly than* (const. $k^2)^k$, *if $k$ is the index of the momentum* (this estimate cannot ensure that the momenta determine the distribution function uniquely).

As explained in the Introduction, (i) and (ii) can be interpreted as distributional counterparts of (I) and (II), respectively.

Since (3) is the Fourier series $(B^2)$ of $f(t)$, the function $f(t)$ has for $-\infty < t < \infty$ the same asymptotic distribution function as for $0 < t < \infty$. On the other hand, it is clear from (6) that $h(t) - \omega(t)$, where $0 < t < \infty$, has the same asymptotic distribution function as $f(t)$. Since (2), being uniformly almost periodic, is a bounded function, it follows that it is sufficient to prove (i) and (ii) for the asymptotic distribution function of $f(t)$, instead of for that of $h(t)$. This means, in the notations of § 2, that the statements (i) and (ii) are equivalent to

(i) $$[\phi] \equiv \lim_{k\to\infty} M_{2k}(\phi)^{1/2k} = \infty$$

and

(ii) $$M_k(\phi)^{1/k} = O(k^2),$$

respectively, where $\phi$ denotes the (necessarily symmetric) asymptotic distribution function of the almost periodic ($B^2$) function $f(t)$ which occurs in (4)-(5) and which has the odd Fourier series (3).

**6.** It is clear from (1) that the Dirichlet series

$$-2\sum_{n=1}^{\infty} \gamma_n^{-1} e^{-\gamma_n s}$$

represents an analytic, almost periodic function in the half-plane $\sigma > 0$. Let $f(\sigma, t)$ denote the imaginary part of this analytic function; so that

(8) $$f(\sigma, t) = -2\sum_{n=1}^{\infty} \gamma_n^{-1} e^{-\gamma_n \sigma} \sin \gamma_n t,$$

where $\sigma > 0$.

Since (8) is, on every fixed line $\sigma(>0)$, the Fourier series of the uniformly almost periodic function $f(\sigma, t)$ of $t$, and since $f(t)$ is almost periodic ($B^2$), with (3) as Fourier series, the Parseval relation is applicable to the difference $f(t) - f(\sigma, t)$. Thus

$$M_t\{[f(t) - f(\sigma, t)]^2\} = \sum_{n=1}^{\infty} \gamma_n^{-2}(1 - e^{-2\gamma_n\sigma}); \qquad (\sigma > 0).$$

Hence, it is clear from (1) that

$$M_t\{[f(t) - f(\sigma, t)]^2\} \to 0 \text{ as } \sigma \to 0.$$

This relation, when compared with the last two paragraphs of § 2, implies that, if $\phi_\sigma = \phi_\sigma(\alpha)$, $-\infty < \alpha < \infty$, denotes the asymptotic distribution function of the uniformly almost periodic function $f(\sigma, t)$ of $t$, then

(9) $$\phi_\sigma \to \phi \text{ as } \sigma \to 0.$$

Another property of the distribution functions $\phi_\sigma$ which will be used is the fact that

(10) $$[\phi_\sigma] \equiv \lim_{k\to\infty} M_{2k}(\phi_\sigma)^{1/2k} \to \infty \text{ as } \sigma \to 0.$$

It is clear from the first and the third paragraphs of § 2 that (10) is equivalent to the statement that the least upper bound of $|f(\sigma, t)|$ for $-\infty < t < \infty$ tends to $\infty$ as $\sigma \to 0$. Hence, it is sufficient to show that $|f(\sigma, \sigma)| \to$

$\infty$ as $\sigma \to 0$. Actually, it is known [13] from Littlewood's proof of the $\Omega$-theorem (cf. the Introduction) that the relation

(10 bis) $\qquad f(\sigma, \sigma) \sim C \log \sigma$, where $C = \text{const.} \neq 0$ and $\sigma \to 0$,

holds with a bounded remainder term. Without an estimate of the remainder term, the asymptotic relation (10 bis) itself, which is more than sufficient for the present purpose, can be established very easily, as follows:

If $\beta_1, \beta_2, \cdots$ is any non-decreasing sequence of numbers which tend to $\infty$ in such a way that

$$\sum_{\beta_n < r} 1 \sim rL(r), \qquad (r \to \infty),$$

holds for some logarithmic function, $L(r)$, then, according to an elementary lemma of Pólya,[14]

$$\sum_{n=1}^{\infty} F(\beta_n/r) \sim rL(r) \int_0^{\infty} F(x)\,dx, \qquad (r \to \infty),$$

holds for any continuous function $F(x)$, $0 \leqq x < \infty$, which is $O(x^{-1-\epsilon})$ as $x \to \infty$. Since (1) is equivalent to

$$\sum_{\gamma_n < r} 1 \sim \frac{1}{2\pi} r \log r, \qquad (r \to \infty),$$

it follows, by choosing $\beta_n = \gamma_n$ and $F(x) = e^{-x} \sin x/x$, that

$$\sum_{n=1}^{\infty} (r/\gamma_n) \exp(-\gamma_n/r) \sin(\gamma_n/r) \sim cr \log r, \qquad (r \to \infty),$$

where $c$ is a non-vanishing constant. Hence, (10 bis) follows by choosing $\sigma = t$ in (8) and placing $r = 1/\sigma$.

7. According to (1), the sum of the absolute values of the coefficients of the uniformly almost periodic series (8) in $t$ is obviously convergent. Hence, it is clear (without any existence theorem [15]) that the Young-Hausdorff inequalities are applicable to (8) on every fixed line $\sigma(> 0)$. Thus, if $k$ is a positive integer,

$$M_t\{f(\sigma, t)^{2k}\}^{1/2k} \leqq \{\sum_{n=1}^{\infty} (\gamma_n^{-1} e^{-\gamma_n \sigma})^{2k/(2k-1)}\}^{(2k-1)/2k}.$$

---

[13] Cf. A. E. Ingham, *loc. cit.*[1], pp. 98-99.

[14] G. Pólya, "Bemerkungen über unendliche Folgen und ganze Funktionen," *Mathematische Annalen*, vol. 88 (1923), pp. 169-193, more particularly pp. 176-177.

[15] In this regard, cf. H. R. Pitt, "On the Fourier coefficients of almost periodic functions," *Journal of the London Mathematical Society*, vol. 14 (1939), pp. 143-150.

Since $\gamma_n > 0$ and $\sigma > 0$, the series, $\{\ \}$, on the right of this inequality is majorized by the series

$$(11) \qquad C_{2k} = \sum_{n=1}^{\infty} \gamma_n^{-2k/(2k-1)},$$

which is convergent, by (1). Thus

$$M_t\{f(\sigma, t)^{2k}\}^{1/2k} < (C_{2k})^{(2k-1)/2k} \text{ for every } \sigma > 0.$$

Furthermore,[16]

$$(12) \qquad M_t\{f(\sigma, t)^{2k}\} = M_{2k}(\phi_\sigma),$$

where $M_{2k}(\phi_\sigma)$ denotes the $2k$-th momentum of the asymptotic distribution function, $\phi_\sigma$, of the uniformly almost periodic function $f(\sigma, t)$ of $t$. Consequently,

$$(13) \qquad M_{2k}(\phi_\sigma) < (C_{2k})^{2k-1} \text{ for every } \sigma > 0,$$

where $C_{2k}$ depends only on $k$.

Accordingly,

$$\limsup_{\sigma \to 0} \int_{-\infty}^{\infty} \alpha^{2k} d\phi_\sigma(\alpha) < \infty$$

for every fixed $k$. Hence,[17]

$$(14) \qquad \limsup_{\sigma \to 0} \left( \int_{-\infty}^{-r} + \int_{r}^{\infty} \right) \alpha^{2k} d\phi_\sigma(\alpha) \to 0 \text{ as } r \to \infty$$

if $k$ is arbitrarily fixed; in fact,

$$\left( \int_{-\infty}^{-r} + \int_{r}^{\infty} \right) \alpha^{2k-2} d\phi_\sigma(\alpha) \leqq r^{-2} \int_{-\infty}^{\infty} \alpha^{2k} d\phi_\sigma(\alpha).$$

**8.** At the beginning of § 6, the function $f(\sigma, t)$ was defined as the imaginary part of a function which is regular analytic in the half-plane $\sigma > 0$. Hence, $f(\sigma, t)$ is an harmonic function, and therefore its $2k$-th power is a non-negative subharmonic function, in the half-plane $\sigma > 0$. Furthermore, (8) and (1) imply that $f(\sigma, t)$, and therefore its $2k$-th power, is uniformly bounded and uniformly almost periodic in the half-plane $\sigma \geqq \epsilon$, where $\epsilon > 0$ is arbitrary. Hence, the Lemma of § 1 is applicable to $g(\sigma, t) = f(\sigma, t)^{2k}$ on any subinterval $\alpha \leqq \sigma \leqq \beta$ of the interval $0 < \sigma < \infty$.

---

[16] A. Wintner, *loc. cit.*[5], p. 271.

[17] This standard step amounts to an application of Tchebycheff's inequality. Cf., e. g., A. Kolmogoroff, " Grundbegriffe der Wahrscheinlichkeitsrechnung," *Ergebnisse der Mathematik und ihrer Grenzgebiete*, vol. 2, no. 3 (1933), pp. 37-38.

On the other hand, it is clear from (8) and (1) that

$$f(\sigma, t) \to 0 \text{ uniformly for } -\infty < t < \infty, \text{ as } \sigma \to \infty.$$

This implies that, if $k$ is arbitrarily fixed,

$$M_t\{f(\sigma, t)^{2k}\} \to 0 \text{ as } \sigma \to \infty.$$

Since, by the Lemma of § 1,

$$\operatorname*{Max}_{\alpha \leqq \sigma \leqq \beta} M_t\{f(\sigma, t)^{2k}\} = \operatorname{Max}\,(M_t\{f(\alpha, t)^{2k}\}, M_t\{f(\beta, t)^{2k}\}) > 0$$

whenever $0 < \alpha < \beta < \infty$, it follows that the function $M_t\{f(\sigma, t)^{2k}\}$ of $\sigma$ must be monotone and non-increasing on the interval $0 < \sigma < \infty$.

Consequently, on letting $\sigma \to 0$, one sees from (12) that

$$(15) \qquad M_{2k}(\phi_\sigma) \leqq \lim_{\epsilon \to 0} M_{2k}(\phi_\epsilon) \text{ for every } \sigma > 0.$$

It is understood that the limit on the right of (15) exists in the sense that it might be $+\infty$; actually, it is finite, by (13).

**9.** Since the space of the functions of $t$ which are almost periodic $(B^q)$ for some fixed $q \geqq 1$ is a complete space (with reference to the topology of the $(B^q)$-metric), the proof of the Young-Hausdorff extension of the Fischer-Riesz existence theorem can be transcribed from $(L^q)$ to $(B^q)$ without any change.[18] On the other hand, it is clear from (1) that

$$\sum_{n=1}^{\infty} \gamma_n^{-1-\epsilon} < \infty \text{ for every } \epsilon > 0.$$

Hence, there exists for every $q > 1$ a real-valued function $f^{(q)}(t)$ which is almost periodic $(B^q)$, has the Fourier expansion (3), and satisfies the inequality

$$M_t\{|\, f^{(q)}(t)\,|^q\}^{1/q} \leqq \{\sum_{n=1}^{\infty} \gamma_n^{-q/(q-1)}\}^{(q-1)/q}.$$

In particular, if $q$ is an even positive integer,

$$(16) \qquad M_t\{f^{(2k)}(t)^{2k}\} \leqq (C_{2k})^{2k-1},$$

by (11).

Since $f^{(2k)}(t)$ is almost periodic $(B^{2k})$ and has the Fourier series (3), it is clear from the last two paragraphs of § 2 that the asymptotic distribution function of the $n$-th partial sum of (3) tends, as $n \to \infty$, to the asymptotic distribution function of $f^{(2k)}(t)$. Hence, the asymptotic distribution function of $f^{(2k)}(t)$ is independent of $k$. Since $\phi$ was defined (§ 5) as the

[18] Cf. H. R. Pitt, *loc. cit.*[15], pp. 144-148.

asymptotic distribution function of the function $f(t)$ which is represented by the convergent series (3), and since this $f(t)$ is, according to § 4, almost periodic $(B^2)$ and such as to have the Fourier series (3), it follows that $\phi$ is the asymptotic distribution function of every $f^{(2k)}(t)$. But (16), where $C_{2k} < \infty$, holds for arbitrarily large values of $k$. Consequently,[19]

(17) $$M_t\{f^{(2k)}(t)^{2k}\} = M_{2k}(\phi)$$

for every $k$.

In particular

(18) $$M_{2k}(\phi) < \infty,$$

and so

(19) $$\left(\int_{-\infty}^{-r} + \int_{r}^{\infty}\right) \alpha^{2k} d\phi(\alpha) \to 0 \text{ as } r \to \infty,$$

where $k$ is arbitrary.

Let $r$ be an arbitrary positive number which is such that neither $\alpha = r$ nor $\alpha = -r$ is a discontinuity point of the monotone function $\phi = \phi(\alpha)$. Then, according to Helly's theorem on term-by term integration,

$$\int_{-r}^{r} \alpha^{2k} d\phi_\sigma(\alpha) \to \int_{-r}^{r} \alpha^{2k} d\phi(\alpha) \text{ as } \sigma \to 0,$$

by (9). Hence, on keeping $k$ arbitrarily fixed, one sees from (19) and (14) that

$$\int_{-\infty}^{\infty} \alpha^{2k} d\phi_\sigma(\alpha) \to \int_{-\infty}^{\infty} \alpha^{2k} d\phi(\alpha) \text{ as } \sigma \to 0.$$

This means that

(20) $$\lim_{\sigma \to 0} M_{2k}(\phi_\sigma) = M_{2k}(\phi)$$

holds for every $k$.

**10.** According to the first paragraph of § 2,

$$\lim_{k\to\infty} M_{2k}(\phi)^{1/2k} = [\phi].$$

But, from (15) and (20),

$$M_{2k}(\phi_\sigma) \leqq M_{2k}(\phi) \text{ for every } \sigma > 0 \text{ and for every } k.$$

Hence,

$$\lim_{k\to\infty} M_{2k}(\phi_\sigma)^{1/2k} \leqq [\phi] \text{ for every } \sigma > 0.$$

---

[19] Cf. B. Jessen and A. Wintner, *loc. cit.*[5], p. 76.

On comparing this with (10), one sees that the proof of (i), § 5 is complete.

Clearly, this proof depends essentially on the relation (15), which in turn depends on the Lemma of § 1. In this sense, (i), § 5 lies much deeper than (ii), § 5, since (ii), § 5 depends only on (13) and (20).

In fact, it is clear from (11) and (1) that, as $k \to \infty$,

$$C_{2k} = O \sum_{n=1}^{\infty} (n^{-1} \log n)^{2k/(2k-1)} = O \int_1^{\infty} (x^{-1} \log x)^{2k/(2k-1)} dx.$$

But the last integral can be written in the form

$$\int_0^{\infty} (e^{-x} x)^{2k/(2k-1)} de^x = A_k \int_0^{\infty} e^{-x} x^{2k/(2k-1)} dx \leqq A_k \int_0^{\infty} e^{-x} x^2 dx,$$

where

$$A_k = (2k-1)^{(4k-1)/(2k-1)} = O(2k-1)^2 = O(k^2).$$

Hence,

$$C_{2k} = O(k^2) \text{ as } k \to \infty.$$

Since (13) and (20) imply that

$$M_{2k}(\phi) \leqq (C_{2k})^{2k-1},$$

it is clear from the Schwarz inequality that the proof of (ii), § 5 is complete.

THE JOHNS HOPKINS UNIVERSITY.